Response to Manuscript # egusphere-2022-243 "Predictability of rainfall induced-landslides: The case study of Western Himalayan Region" by Swadhi Ritumbara Das and Poulomi Ganguli

We would like to thank the reviewer for the valuable comments and for providing us with an opportunity to improve our manuscript. In this response document, first, we will provide an overall summary of the editor's comments and then address each of the comments raised by the reviewers. Our responses are embedded within the comments (in BLACK) in BLUE. The new additions to the revised manuscript are embedded below in BROWN.

Response to Reviewer 1 Comments

Comment 1. The work aims at defining the triggering conditions, expressed by rainfall thresholds, of rainfall-induced landslides in the Western Himalayan Region, using landslide data from a global catalog and rainfall data from six stations, and gridded data. The work is clear and well-presented. The literature review is not complete. The results are presented and discussed by means of some figures that could be improved.

Response: We appreciate the favorable feedback on our work from the reviewer. We have addressed each of the reviewer's concerns in the following sections. We have improve our literature review section and made revisions to the figures as suggested.

Comment 1.1. Overall, in my opinion, the work has some serious flaws both in the data and in the methodology. These flaws affect considerably the results. I think the main flaws of the manuscript lie in 1) the rainfall data, which are not complete and reliable, despite the efforts of the authors.

Response: We acknowledge that the gauge-based rainfall records are insufficient. However, we applied a robust machine learning approach to infill and reconstructed rainfall time series to make records complete and uniform in lengths across all sites for the whole analysis time period of 1970 to 2019. Further, to correct biases in rainfall imputation, we employed a statistical post-process technique using the month-wise-CDF matching approach for the daily rainfall series. We reconstructed the incomplete daily time series using the Regularized Expectation-Maximization (Reg-EM) approach and India Meteorological Department's (IMD) gridded record, which is available at a very high spatial resolution (0.25°) for an extended time period from 1901-2021 at the IMD website (https://www.imdpune.gov.in/Clim_Pred_LRF_New/Grided_Data_Download.html).

The EM algorithm is an iterative procedure to compute the maximum likelihood estimate in the presence of missing records (Dempster et al., 1977). The RegEM estimates regression coefficients by ridge regression, which is a regularized regression method in which a continuous regularization parameter controls the "noise filtering" in the time series (Schneider, 2001). The RegEM is well established in the climatic and paleo-climatic fields and the credibility of the algorithm in infilling missing gaps in rainfall series are extensively discussed in the literature (Kalteh and Hjorth, 2009; Tsidu, 2012; Feng et al., 2013; Þórðarson et al., 2021). We have demonstrated the skill of the RegEM algorithm in infilling missing gaps in rainfall time series through a detailed validation. We evaluated the model's credibility through an integrated test score as described in Beck et al. (2017). Our validation framework against daily observed rainfall time series showed the skill of the RegEM varies in the range of 0.43 (moderate) to 0.95 (excellent), which is robust in a statistical sense. The model skill in reconstructing precipitation time series is discussed in detail on page 10, lines 253-255 in the earlier version of the manuscript. However, considering the reviewer's concern, we have included the following statements on page 13, line 340:

"The lack of adequate spatiotemporal coverage of gauge-based observations is one of the challenging aspects of the present study. However, we overcome the limitation using a robust machine learning method, the regularized EM (Schneider, 2001), which proved to be very skillful in capturing seasonal precipitation variability and its extremes at individual station locations, as demonstrated through the integrated skill score from the literature (Beck et al., 2017)."

Comment 1.2. The method used for selecting landslide-triggering and non-triggering events, which is highly questionable and far from having a physical justification.

Response: We appreciate reviewer's feedback. Here, we would like to point to the reviewer that the review of the literature suggests that an old rainfall event typically displays less impact in triggering landslides than the recent one (Vallet et al., 2015). To avoid any further misconceptions and subjectivity associated with the terms, 'triggering' and 'non-triggering', we have revised the terms as "antecedent" and "precedent" rainfall events. While "antecedent" rainfall corresponds to temporal clustering of rain events preceding a landslide considering a long (several weeks) time window of up to a 30-days, the "precedent" rainfall events correspond to a short-term period, preferably a rainy day before the occurrence of landslides. The above typology is physically consistent since the two periods do not exclusively overlap

and the 'antecedent' period follows the 'precedent' period. Second, the choice of a 30-day time window is based on an antecedent soil moisture condition (AMC) prevalent over a catchment that may trigger flash floods followed by landslides, owing to compound occurrences of extreme rain on already saturated soil (Bertola et al., 2021). The review of the literature shows efforts to calibrate antecedent rainfall in triggering landslides and found that antecedent rainfall conditions up to 30-day are sufficient to trigger rainfall-induced landslides (Marques et al., 2008; Khan et al., 2012; Lee et al., 2014; Johnston et al., 2021), while for deep-seated landslides antecedent precipitation of 90-120 days are often considered (Bevacqua et al., 2021). Since in our case, most of the events are associated with shallow landslides (Kirschbaum et al., 2015; Table 1 on page 11 of this response letter), the choice of 30-day antecedent time window are sufficient to consider AMC condition triggering landslides.

Comment 1.3. The method used for calculating the rainfall thresholds, which is particular and do not result in proper thresholds (the curves are best-fit curves).

Response: We agree. As suggested in the following section of the review and the available literature (Guzzetti et al., 2007; Khan et al., 2012; Giannecchini et al., 2012) based on daily rainfall records from a quality-controlled database (https://dsp.imdpune.gov.in/), first, we develop a power-law relationship between mean rainfall intensity (I) and the corresponding event duration (d). Despite considerable scatter and clustering of points at a specific duration, we find that with an increase in duration, the average intensity likely to trigger slope failures often decreases linearly for several gauges in a log-log plot.

Following Brunetti et al. (2010), we adopted a frequentist method to determine rainfall thresholds for landslides in the revision. The methodology adopted is summarized in Figure 1 (Figure A1 in Appendix section of the revised manuscript). First, we fit the best fit line based on linear regression estimates between mean intensity of rainfall preceded up to 1-month from the date of occurrence of landslides versus the event duration. Next, we determine the PDF of the residual using empirical Kernel density function. This is because the residual errors with constant variance, $N\left(0,\sigma^2\right)$ (where $N\left(\bullet\right)$ denotes normality assumption of residuals with 0 mean and constant variance, σ^2) is rarely satisfied in case of models dealing extremes – typically, large prediction errors are associated with larger storms (Schoups and Vrugt, 2010). Finally, we determine the shift of each threshold from the best fit line corresponds to x% exceedance probability levels, where x, indicates shift of the exceedance line with 10%, 15%,

20%, 30% and 50% exceedance probabilities for at-site estimates (**Table 6** in the subsequent section of this response document), and 1%, 5%, 10%, 20%, 30% and 50% exceedance probability levels at regional scale (**Table 7**). The associated threshold equations are presented in **Table 6** and **Table 7** of this response document. The above methodology based on empirical rainfall ID thresholds has been successfully implemented before in south Asia and elsewhere (Harilal et al., 2019; Maturidi et al., 2020; Larsen and Simon, 1993).

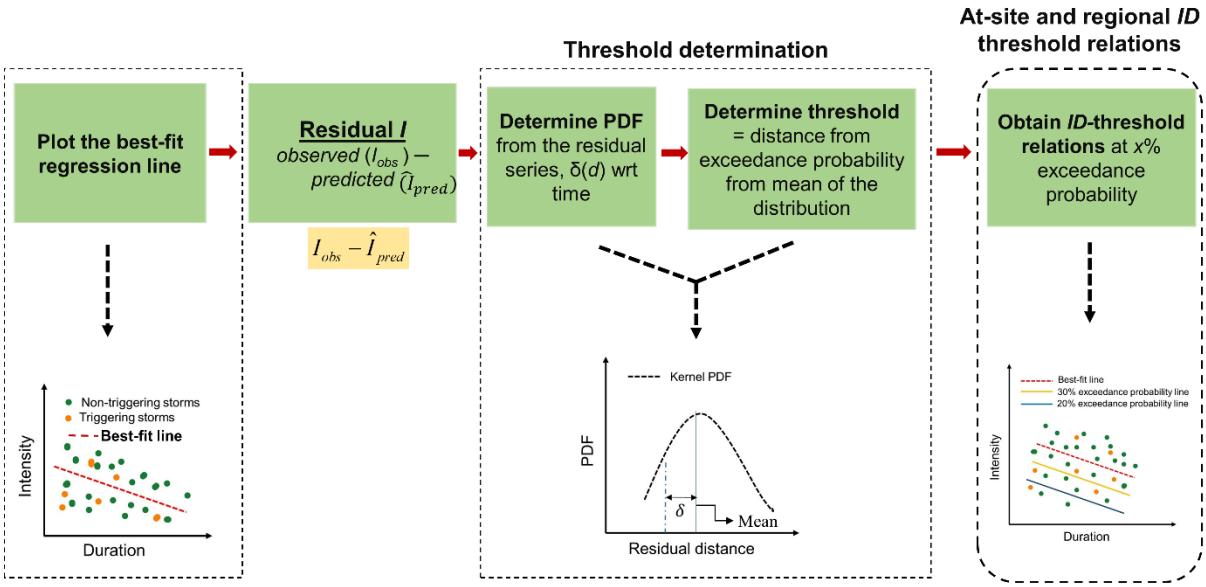

Figure 1 (Fig. A1 in revision): Detailed process flow to obtain rainfall thresholds responsible for landslides from the power-law relations of mean rain intensity (*I*) versus event duration.

Comment 1.4. The obtained results, which need some clarifications.

Response: Thanks for the suggestions. We appreciate the reviewer's comments and have revised our method to develop empirical rainfall ID threshold relationships for at-site and regional scale as suggested in the literature (Harilal et al., 2019; Maturidi et al., 2020; Larsen and Simon, 1993).

Comment 1.5. Moreover, it seems that the Authors do not know well the current literature on rainfall thresholds: this issue spoils the discussion of the results.

Response: We agree. We have revised the literature review as suggested. We have included additional studies focusing on rainfall thresholds in the revised manuscript.

Comment 1.6 Finally, the structure of the paper needs a revision. There is not a clear distinction among method, results and discussion. Some methodological points are reported in the result and discussion section.

Response: Agreed and incorporated. The methodological issues reported are now removed and revised as suggested.

Comment 1.7 Overall, the manuscript is a good attempt to study the critical conditions that lead to the initiation of landslides in a Himalayan region. Unfortunately, it has severe methodological issues which do not allow its acceptance in the present form. In the following, I report a list of comments and suggestions that could be useful for addressing these issues and improving the quality of the work. In my opinion, the manuscript does not meet the standards for being considered for publication. It can be reconsidered after major revisions.

Response: Agreed and we have revised as suggested. We appreciate the reviewer's concern and addressed each of his/her comments in details.

Comment 2. Line 19: I would not say "long-term climatology".

Response: Agreed and incorporated. We have revised it in the following way:

"Further, we show an increase in rainfall in recent decades (2007-16) over low-elevated areas of the WHR compared to the climatology (1988-2006), revealing intensification of rain events, which could potentially amplify landslide occurrences."

Comment 3. L31: These numbers deserve a reference.

Response: Reference has been added.

"In India, out of the 0.42 million km^2 (12.6%) that are prone to landslide, 0.32 million km^2 falls in the Himalayan range (www.gsi.gov.in) (Hindustan Times, 2016)".

Comment 4. L49: The paper Gariano & Guzzetti 2016 does not deal with probabilistic approaches in rainfall thresholds.

Response: The paper has been removed from the revision.

Comment 5. L51: It is worth mentioning also the frequentist method (Brunetti et al. 2010), which seems to be the most adopted method worldwide, with applications also in the Himalayan Area.

E.g. 1) https://link.springer.com/article/10.1007/s10064-018-1415-2;

2) https://www.mdpi.com/2076-3263/9/7/302;

3) https://www.mdpi.com/2073-4441/11/8/1616/htm;

4) https://link.springer.com/article/10.1007/s11069-020-04407-9

Response: We agree. We have incorporated the suggested literature in the revision:

"Brunetti et al. (2010) compiled a catalogue comprising 753 rainfall events that have resulted in landslides in Italy. They have identified the minimum rainfall intensity responsible for rain-induced landslides in central Italy based on mean rainfall intensity versus event duration. They have evaluated two independent statistical methods, Bayesian inference and a frequentist approach. Their analysis suggested that landslides in Italy can be triggered by less severe rainfall intensity than previously established. Further, the frequentist method implemented in this study allowed the identification of multiple minimum rainfall thresholds based on different exceedance probabilities. Gariano et al. (2019) developed empirical accumulated event rainfall versus event duration thresholds for the possible initiation of landslides in southwestern Bhutan, using daily rainfall measurements obtained from only three rain gauges. For this, they have applied a frequentist method and used an automated tool, namely 'Calculation of Thresholds for Rainfall-Induced Landslides-Tool (CTRL-T)', to identify rainfall conditions responsible for landslides and reported thresholds correspond to different exceedance probabilities. Teja et al. (2019) demonstrated the application of CTRL-T tool to determine the relationship between accumulated rainfall versus event duration across the Kalimpong region of Darjeeling Himalayas of eastern Himalaya. Their results showed rainfall event of 48 hrs duration with a cumulated rainfall event of ~37 mm might cause landslides in the area. Dikshit et al. (2019) proposed cumulated event rainfall versus duration thresholds for Chukha area in southwest Bhutan for the analysis period, 2004-2014. Their analysis further showed a 10-day antecedent rainfall would be adequate to trigger landslides in the area. Recently, Mandal and Sarkar (2021) analyzed monsoon-induced landslides on the National Highway-10 (NH10) area of Darjeeling Himalaya. They have developed a power-law relationship for mean intensity-versus-duration. They determined rainfall threshold based on a 1% exceedance probability from the shift in residual of the best-fit line. Their results showed that areas with slate lithology

is more vulnerable to landslides than other rock types. However, to date most of the analyses focused on the eastern Himalayan range of India and a few areas of Nepal Himalaya. Recently, using remotely-sensed rainfall records, Martha et al. (2021) showed the variations in rainfall resulting in landslides in a pan-Indian domain. Their study revealed that rainfall thresholds responsible for triggering landslides in the northwestern Himalayas are typically larger than the northeastern Himalayan range. Further, their study showed that the occurrence of landslides across the northeastern Himalayan range does not follow any distinct range of rainfall value.

Comment 6. L54-55: Please add references.

Response: References are added.

"Globally, several studies have reported rain-induced shallow to deep landslides and debris flow, and derived rainfall Intensity-Duration threshold curve (Guzzetti et al., 2007; Brunetti et al., 2010; Staley et al., 2013; Giannecchini et al., 2012; Kanungo and Sharma, 2014)."

Comment 7. L56-57: I do not understand what you mean with this sentence.

Response: We apologize. We have revised the sentence as below:

"Due to a lack of standardized sources of rainfall and geomorphological information, an inconsistency may exist in the global landslide catalogue; not all local-level event information is available in the global archive."

Comment 8. L68: I think you missed two papers that deserve to be read before writing an article dealing with rainfall thresholds in the Himalayas.

1) https://link.springer.com/article/10.1007/s10346-018-0966-4;

2) https://www.mdpi.com/2076-3417/10/7/2466

Response: We thank the reviewer for pointing these two important articles. We have added the following sentences in the revised manuscript on page 3 in line 63:

"Although several studies have analyzed rainfall thresholds over Europe and Asia (Segoni et al., 2018), especially focusing on eastern (Dikshit et al., 2019; Sengupta et al., 2010; Mandal and Sarkar, 2021; Abraham et al., 2022) and western Himalayan regions (Kanungo and Sharma, 2014; Gariano et al., 2019; Martha et al., 2015), most of these studies relied on single or very few numbers of rain gauges (Dikshit et al., 2019; Sengupta et al., 2010; Gariano et al.,

2019; Harilal et al., 2019; Abraham et al., 2022). Further, no studies have so far investigated the regionalization of empirical rainfall *ID* threshold relationships. While the primary strength of station-based rainfall records is their ability to capture local-scale processes, such as extreme precipitation events precisely, their spatiotemporal coverage in complex topographical terrain often limits a credible estimate of *ID* threshold relation. However, no attempt has been made so far to reconstruct daily or hourly rainfall time series with sparser temporal coverage for predicting landslides, although an equivalent analysis was carried out by developing rainfall *intensity-duration-frequency* curves using reconstructed ground-based hourly rainfall time series for urban infrastructure management (Ganguli and Coulibaly, 2017).

Comment 9. L92: Don't you think that only six stations are few for making reliable analyses?

Response: We agree. We tried to include the maximum available station-based rainfall records in our analysis. We have used the quality-controlled rainfall records from the IMD (https://dsp.imdpune.gov.in/). These are the best available station-based records available for this area, which are prone to landslides. We argue that the previous assessments of local *ID* thresholds analysis often relied on single or very few numbers of rain gauges typically ranging from one to three (Dikshit et al., 2019; Sengupta et al., 2010; Gariano et al., 2019; Harilal et al., 2019; Abraham et al., 2022). We point to the reviewer that the primary strength of point rain gauge measurement lies in their ability to capture local scale processes precisely, which often satellite or gridded records fail to achieve. This is because satellite rainfall estimates have embedded uncertainties since satellites do not measure rainfall by itself, and should be related to precipitations based on one or multiple surrogate variables. These uncertainties, may cascade in the process of temporal samplings, error from algorithms and satellite instruments itself (Gebremichael et al., 2005; Toté et al., 2015; Wu et al., 2012).

Comment 10. L92-93: I would mention here that these are Indian states. For an immediate understanding.

Response: Agreed and incorporated.

"We selected six rain gauge stations across the WHR which are spatially distributed across three Indian states namely Jammu and Kashmir (Banihal and Katra), Himachal Pradesh (Mandi and Solan) and Uttarakhand (Dehradun and Joshimath). Figure 1 shows the spatial distributions of all six stations within the WHR".

Comment 11. L95: Actually, only one station covers the period 1970-2019. For all the others, the measurement range starts from 1985. It's a difference of 15 years, so it is worth mentioning it. Moreover, the stations Joshimath and Katra have very few data, so I would remove them from the analyses. At the end, only Dehradun and Banihal have a sufficient amount of data.

Response: Although the period of record availability varies from site to site (Table 1 in the earlier version of the manuscript), we maintained a uniform record lengths from 1970 to 2019 across all sites. We agree that Joshimath and Katra have very few records, but we have reconstructed the daily rainfall time series using regularized Expectation Maximization (RegEM) algorithm considering forcing from the daily high-resolution (at 0.25°) gridded rainfall records provided by the IMD, archived at the Climate Data Service Portal (https://cdsp.imdpune.gov.in/home_gridded_data.php). The RegEM is well established in the climatic field, and the credibility of the algorithm in reconstructing precipitation series is extensively discussed in the literature (Kalteh and Hjorth, 2009; Tsidu, 2012; Feng et al., 2013; Þórðarson et al., 2021). While spatiotemporal coverage of gauge-based products in complex topographical terrain often limits their applicability in deriving rainfall *ID* thresholds, the added value of our analysis is that we have reconstructed the daily rainfall series from 1970-2019 to maintain the uniformity of records across the spatial and temporal scales. We have already discussed the method of reconstruction of daily rainfall time series at individual point rain gauge locations on page 5 (Method section) and addressed the credibility of the proposed algorithm against observed ground-based time series in detail at the Result and Discussion section on page 10, lines 250-263 in the earlier version of the manuscript. We further argue that the integrated test score, proposed by Beck et al. (2017) showed a reasonably well fit (with skill scores vary from moderate to excellent) between observed versus reconstructed time series. Although based on meteorological record availability, we have *attempted six available spatially distributed locations across landslide-vulnerable areas of the western Himalayas, which none of the existing studies have attempted so far*, we plan to extend our analysis to broader geographical domain as part of the future assessment.

Comment 12. L98: For which locations are the gridded daily rainfall obtained? The same of the stations or for the entire areas of the considered states? And for which period? Please specify.

Response: The high resolution (0.25°×0.25°) gridded daily rainfall data from IMD (Srivastava et al., 2009) was obtained for the whole WHR domain from the period from 1970 to 2019.

Further, to reconstruct the precipitation series, we selected neighboring rain grids within a radius, r = 50 km around each station, which resulted in ~ 10-12 neighboring grids across each rain gauge. While a smaller radius is preferable to avoid smoothing of data from different precipitation events and orographic facets, we argue that we have not implemented any smoothing algorithm to infill the gaps in the time series. We employ a robust machine learning algorithm, RegEM (Schneider, 2001) to reconstruct the incomplete rainfall time series and to maintain an uniform temporal coverage across sites for the analysis period, 1970-2019. Since a smaller search radius would results in an inadequate number of neighboring rain grids, to increase the skill of RegEM algorithm, we increase the search radius to up to 50 km, which is consistent with earlier assessments in areas of complex topography (Bellido-Jiménez et al., 2021; Livneh and Rajagopalan, 2017; Wahl et al., 2015).

We have added the following texts on page 5, line 114:

"We obtain the high resolution (0.25°×0.25°) gridded daily rainfall data from the IMD (Srivastava et al., 2009; Climate Data Service Portal: https://cdsp.imdpune.gov.in/home_gridded_data.php) for the whole WHR region for the period 1970 to 2019. Further, to reconstruct the precipitation series, we selected rain grids within an r = 50 km radius of each station location, which resulted in around 10-12 neighboring grids around each rain gauge. We employ a robust machine learning algorithm, RegEM (Schneider, 2001) to reconstruct the incomplete rainfall time series and to maintain an uniform temporal coverage across all sites for the analysis period (1970-2019). Since a smaller search radius would results in an inadequate number of neighboring rain grids, to increase the skill of RegEM algorithm, we set the search radius to 50 km, which is consistent with earlier assessments in areas of complex topography (Bellido-Jiménez et al., 2021; Livneh and Rajagopalan, 2017; Wahl et al., 2015)."

Comment 13. L100-101: I would add some details on the landslide inventory. How many landslides? What information includes? What's the temporal range? What the temporal and spatial accuracy... and so on.

Response: Agreed and incorporated. We have added the following texts in the manuscript and present Table A1 showing locations of landslides and location accuracy (if available), around each station.

The global landslide database consists of more than eleven thousand events from 1988-2016 (Juang et al., 2019). It comprises various types of information such as location (Latitude and Longitude), type and size of landslide, fatality, date of landslide, temporal resolution recorded as HH:MM (24-hour clock in local time), and the main trigger of landslide (See Table 2 in Juang et al., 2019 and Table 1 at Kirschbaum et al., 2015 for details). However, the global landslide catalog often does not include local-level event information. Therefore, we have also included the events available at Geological Survey of India's (GSI) landslide inventory catalogue (GSI Bhukosh, 2022) for Himachal and Uttarakhand states, available at the GSI Bhukosh archive (https://bhukosh.gsi.gov.in/Bhukosh/). The GSI's landslide inventory database (GSI Bhukosh, 2022), offer wealth of information including landslide locations, activity level, geology type, triggering drivers and dimension of slides. However, the GSI Bhukosh archive does not provide sufficient information on the exact date of occurrence of landslide. Therefore based on available synopsis of the event and 'Entry date' information, we have considered the date of entry of the event as a proxy date for the landslide occurrence. From both these archive, we have compiled the events that belong to the WHR area with rainfall as the main trigger for landslides. Table A1 lists the number of landslides associated for each sites, temporal range, and locational accuracy.

Table 1*. Details of landslides inventory records as adopted in this study

Stations	Distance (km)	Number of events	Classification of landslides[1]				Locational accuracy according to NASA COOLR database (in km)
			Small	Medium	Large	Very large	
Banihal	6-22	20	1	20	0	0	1-25
Katra	3-24	21	1	20	0	0	5-25
Mandi	2-22	15	2	13	0	0	1-50
Solan	7-22	16	0	16	0	0	5-25
Dehradun	1-20	12	4	8	0	0	25-50
Joshimath	1-24	16	2	13	0	1	1-50

*Table A1 in the appendix section of the revised manuscript.[1]Landslide classifications are based on the literature (Kirschbaum et al., 2015; Juang et al., 2019), and were determined from the compiled database from NASA COOLR and GSI Bhukosh.

Comment 14. L104: How many gaps are present in each rainfall series? You could add this information in table 1, by adding e.g. the percentage of missing data in the time span of each station.

Response: Agreed and incorporated column number 9 in Table 1 of revised manuscript.

Table 2. Details of fraction of missing records for each site

Station Name	Missing gaps (in %)
Banihal	0.6
Katra	9.8
Mandi	40
Solan	8.5
Dehradun	0.2
Joshimath	2.2

Comment 15. L144: This paragraph is not clear at all. It starts talking about triggering rainfall events that cause landslides, it continues talking about weighted average rainfall and it ends talking about thresholds not better specified. I can't follow what the authors want to describe. It is very unclear to me how the triggering rainfall events were reconstructed and how they are linked to the landslides that are supposed to have triggered. I would suggest a rewording of the section.

Response: We apologize. We have reworded the section heading in the revised version of the manuscript to, **Identification of Rain Events**. Here, we discarded the rain events from the time series that have significantly less magnitude and can be considered as trace (or drizzle) values. According to the IMD, the definition of a rainy day is typically based on a non-negative constant threshold. For example, IMD uses 2.5 mmd^{-1} as a threshold to define a rainy day. However, this definition does not account for seasonal heterogeneity of rainfall, despite the fact that rainfall in India show high seasonality due to the influence of southwest monsoon and western disturbances (Soman and Kumar, 1990; Battula et al., 2022). Therefore, following Ratan and Venugopal (2013), we use a seasonally varying threshold to define a rainy day. This is achieved by setting a certain threshold that varies across seasons, keeping the seasonal heterogeneity in view. To this end, we divide the seasons into high flow and low flow seasons based on seasonality of rainfall distributions. For instance, while for five stations, Katra, Mandi, Solan, Dehradun, and Joshimath, the high flow season is during the southwest monsoon

(June-September) season, for Banihal, the apparent high flow season is during the winter months (January-March), owing to the effect of western disturbances (Battula et al., 2022). For remaining seasons (i.e., low flow months), we consider 1 mm rain magnitude as a threshold to consider as a rainy day; below this threshold is considered as a drizzle. For high flow seasons, we calculated the threshold as 1% of mean seasonal rainfall, which we obtain by seasonal average of rain event magnitude with the duration of events as a weighing factor. For the high flow season, the threshold to define a rain event is approximately 3 mm for each site, which is close to the standard threshold to define a rainy day, set by the IMD (Ratan and Venugopal, 2013).

Accordingly, we have revised the section as below:

2.2.2 Identification of Rain Events

To identify rain events from the reconstructed precipitation time series (1970-2019), we followed the definition of a rainy day as per the IMD, which is typically based on a non-negative constant threshold. According to the IMD, a (light) rainy day has been defined as a day with rainfall of magnitude 2.5 mm to 7.5 mm (see page 4 in the IMD weather glossary: https://www.imdpune.gov.in/Weather/Reports/glossary.pdf). However, this definition does not account for the spatial and seasonal heterogeneity of rainfall. Therefore, following Ratan and Venugopal (2013), we use a seasonally varying threshold to define a rainy day. This is achieved by setting a certain threshold that varies across seasons and sites, keeping the spatial and seasonal heterogeneity in view. To this end, we divide the seasons into high flow and low flow seasons based on rainfall seasonality distributions. To this end, we divide the seasons into the high flow-(HFS) and low flow seasons (LFS) based on the seasonality of rainfall distributions. The HFS consists of monsoon season across all sites as the southwest monsoon season is the primary contributor to total annual rainfall in India (Soman and Kumar, 1990), except for high altitude areas of Jammu and Kashmir state, where Banihal is located. For Banihal, the apparent HFS occurs during the winter (refer figure 4) due to the influence of western disturbances (Battula et al., 2022). For LFS, we consider the minimum threshold to define a rainy day is 1 mm. For HFS, we fix the threshold as 1% of seasonal mean rainfall (mm), weighted by the event duration (in days):

$$mean\ seasonal\ rainfall = \frac{sum(Rain\ sum_{seasons} \times length_{seasons})}{sum(length_{seasons})} \quad (1)$$

In Eq. 1, the variable $Rain\ sum_{seasons}$ indicates sum of the rain events with a magnitude greater than 1 mm during a site-specific HFS season, whereas $length_{seasons}$ denotes the number of events with rain magnitude of more than 1 mm. We consider a threshold of 1% of the mean seasonal rainfall for the HFS. To maintain the consistency in seasonal rainfall thresholds across the sites, we consider the spatial average (50^{th} percentile) of the threshold values for the monsoon season across all sites. The rainfall threshold in HFS is set as approximately 3 mm for all six site based on mean seasonal rainfall (Eq. 1). Among six stations, Banihal showed the highest seasonal average rainfall magnitude in the winter (December-March) months, unlike other stations. Therefore, we consider the rainfall threshold during winter (HFS) for Banihal as 3 mm, while the rain threshold is set as 1 mm for the other three seasons (LFS). Our criteria to define seasonally varying rain events are consistent with the operational definition of a rainy day set by the IMD and the existing literature (Ratan and Venugopal, 2013).

Comment 16. L146: If I have understood well, you considered a rainfall event a series of rainfall measurements equal or greater than 1mm, regardless of the duration. I suppose that you obtained several events with very short durations. If it's so, how did you handle this issue?

Response: We would like to point to the reviewer that we fix a seasonally varying threshold to define a rainy day, which is consistent with the available literature (Ratan and Venugopal, 2013). A rainy day for LFS season is set to rainfall threshold of 1 mm or larger. However, since most of the landslides events are associated with HFS episode, we do not have chances of encountering several events with very short duration as pointed by the reviewer.

Comment 17. L155-156: This sentence is not clear to me.

Response: We have reword this sentence for better readability:

"We consider a threshold of 1% of the mean seasonal rainfall for the HFS episode to define a rainy day to account for seasonal variability in defining rain events."

Comment 18. Don't you think that a distance of 50 km between the landslides and the rainfall is too much for associate a rainfall event to a landslide? And to say that those rainfall events have triggered those landslides? I think so. Not to mention a distance of 100 km! In particular

in a mountain environment where rainfall is affected by a high spatial variability (see e.g. https://www.sciencedirect.com/science/article/abs/pii/S0022169415007696) It would be interesting to know what's the mean distance between stations and landslides considered in this study. Most of the studies regarding relationships between rainfall and landslides - particularly rainfall thresholds - do not consider distance longer than 15-20 km. I think this is a great flaw of the article.

Response: Agreed; as suggested we have reduced the radial distance between stations and landslide locations to d = 25 km, which is slightly larger than as pointed by the reviewer. To this end, we would like to state that the Global landslide catalogue does not provide enough landslide information to such a small distance. Therefore, we have also considered GSI landslide catalogue (Bhukosh, 2022) to incorporate additional landslides for states of Himachal and Uttarakhand. Although we acknowledge that the choice of d is slightly larger in this case, which could impart an additional uncertainty in the analysis, with a smaller radius, we have very few samples to robustly identify any specific trend in the event rainfall responsible for landslides. We have incorporated Table A1 in the Appendix section of the revised manuscript that shows the range of distances from the station to landslide locations.

Comment 19. However, how many events were associated to each station? Please add this information.

Response: Agreed and incorporated as Table A1 (Table 3 in the response document) in the revised manuscript. The third column of Table A1 contains the number of events associated within 25 km radius of each station:

Table 3*: The great circle radial distance and the number of landslides within 25 km radius of each station with the locational accuracy of each station.

Stations	Distance (km)	Number of events	Classification of landslides[1]				Locational accuracy according to NASA COOLR database (in km)
			Small	Medium	Large	Very large	
Banihal	6-22	20	1	20	0	0	1-25
Katra	3-24	21	1	20	0	0	5-25
Mandi	2-22	15	2	13	0	0	1-50
Solan	7-22	16	0	16	0	0	5-25
Dehradun	1-20	12	4	8	0	0	25-50
Joshimath	1-24	16	2	13	0	1	1-50

*Table A1 in the appendix section of the revised manuscript.[1]Landslide classifications are based on the literature (Kirschbaum et al., 2015; Juang et al., 2019).

Comment 20. L169: This can be reasonable if the landslides are deep-seated; in case of shallow landslides it seems a very long period. Do you have information on the depth of the landslides?

Response: The classification of landslides are available based on landslide volume information available at NASA COOLR repository (Table 2 in Juang et al., 2019):

- Small landslides: Volume - <10 cubic meters
- Medium landslides: Volume - 10 to <1000 cubic meters
- Large landslides: Volume - 1000 to <100,000 cubic meters
- Very large landslides: Volume - 100,000 to <1 million cubic meters

We have included this information in the revised version of the manuscript.

Comment 21. L172: Considering the triggering rain event as the one that "produces the maximum rain intensity at most 30 days before the landslide" is highly questionable. so, if you have a rainfall event with the maximum intensity occurred from 25 to 23 day before the landslide and then you have another event that occurred immediately before the landslide (with a lower intensity), you consider the first one? This has not physical justification.

Response: We have revised the result section dealing with rainfall ID relations as suggested. Instead of considering the maximum intensity, we have now considered mean rainfall intensity for all events 30-days before landslide versus the event duration in the revised analysis.

Comment 22. L179-180: Usually the exponent of an ID threshold is negative because the intensity becomes lower with higher duration, as you can see from several papers dealing with it. Your result is strange.

Response: We agree. As responded earlier, we have considered mean rainfall intensity for all events 30-days before landslide versus event duration in the analysis (Figure 2; Figure 10 in the revised manuscript). Further, we have discarded events with very small intensities, such as rain intensity lesser than (0.5 mm/hr) and associated durations. Our revised analysis showed, a negative scaling exponents for derived power-law relationships for some stations, however, yet positive exponents for a few stations. Furthermore, we would like to point to the reviewer that the regional ID threshold curves for region 2 (Figure 3; Figure 11 in the revised manuscript), shows a negative exponent.

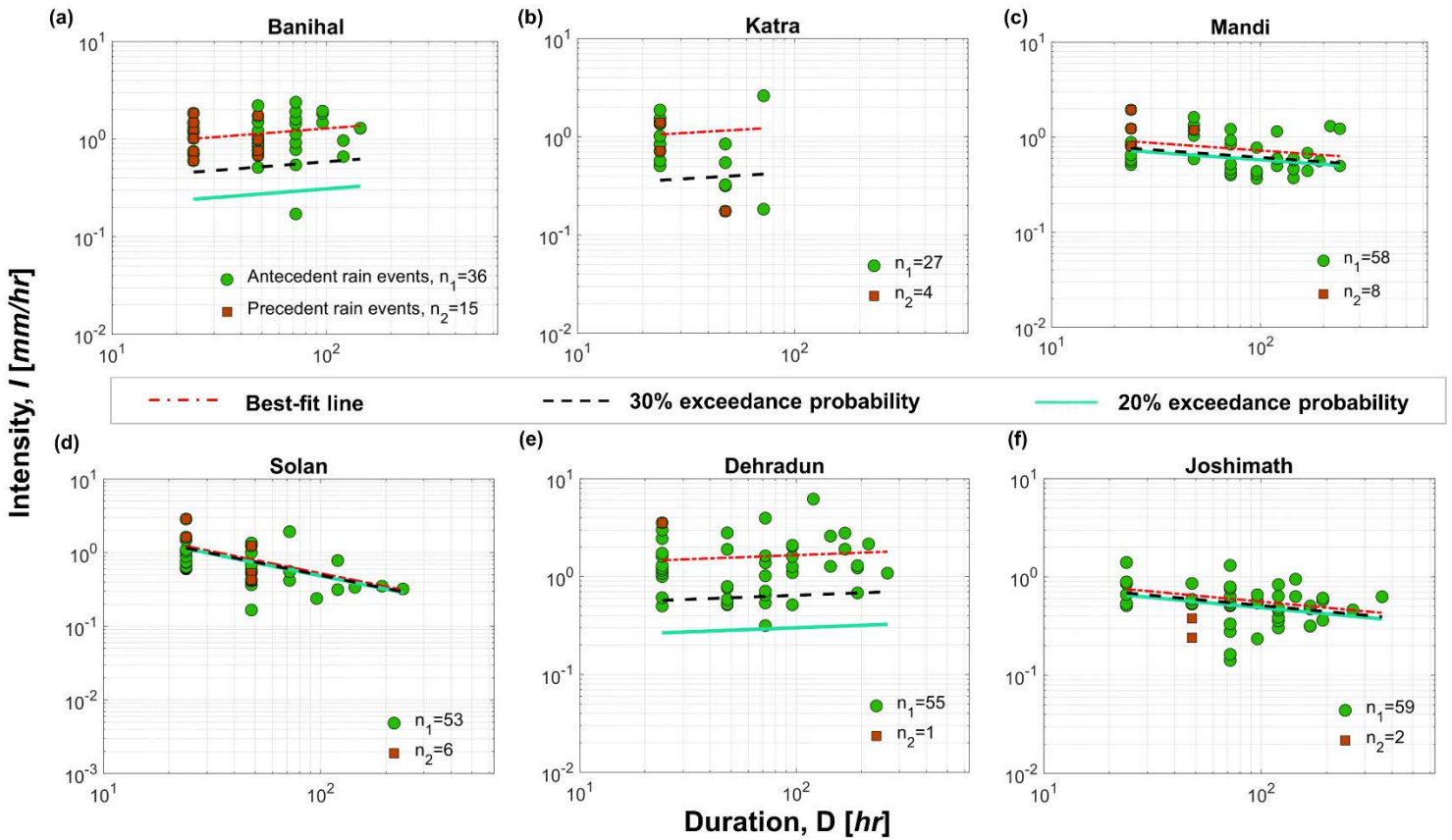

Figure 2. (Figure 10 in revision) Relation between average rainfall intensity (in mm) -versus- duration (in hours) for storms dating from 2007 to 2019 in the WHR. The red dashed line represents the best fit line. The circles indicate individual storm events; the circle in green shows antecedent rain events preceding 30-day before the landslides, whereas squares in brown show precedent rain events immediately before the day of landslides. The dotted-black and cyan line represents threshold curves corresponding to 30% and 20% exceedance probabilities respectively. For Katra, only 30% exceedance probability curve is shown due to limited number of records available for this station. The equations for empirical *ID* threshold curves at 10%, 15%, 20%, 30% and 50% exceedance probabilities are presented in **Table 6** for all at-site locations.

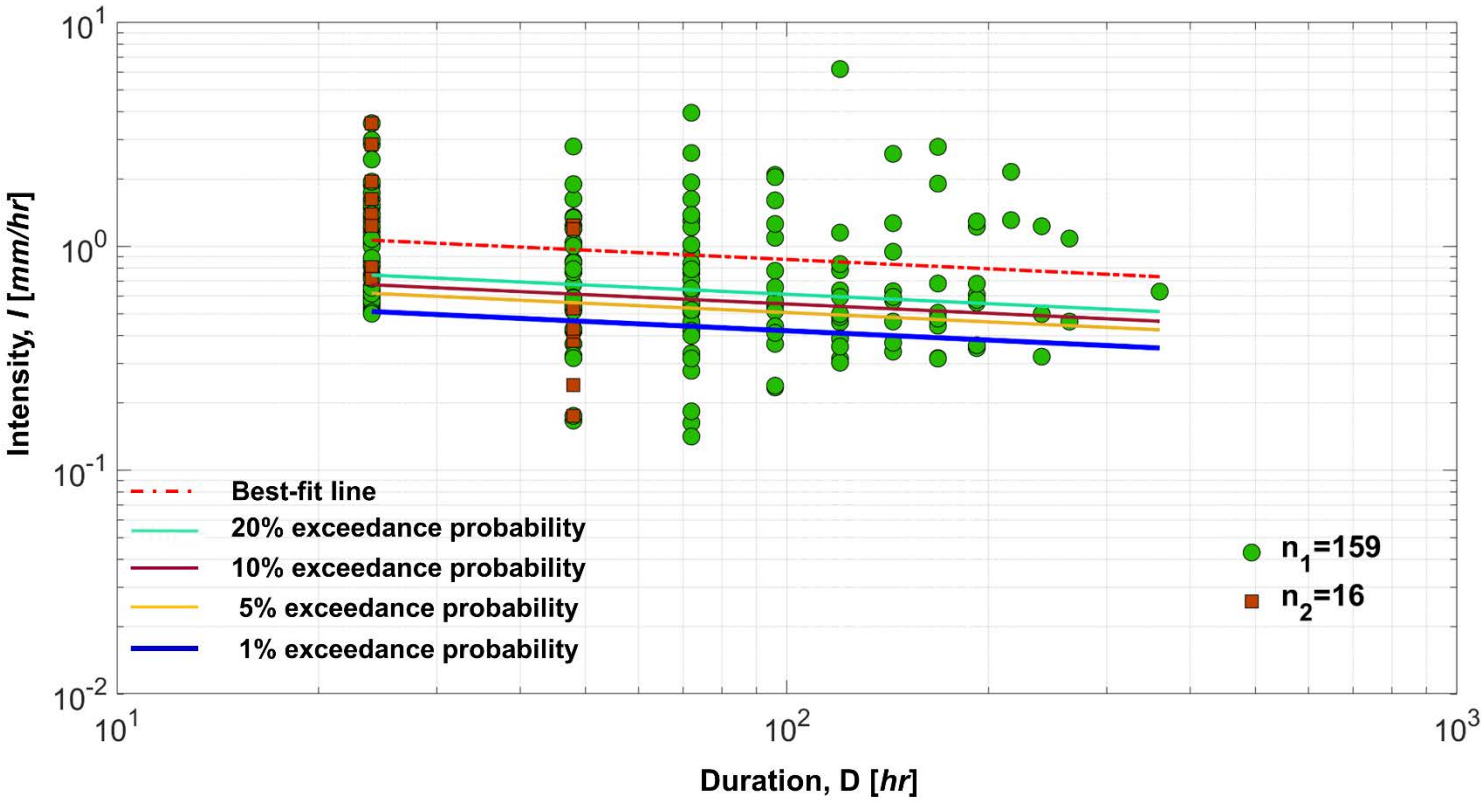

Figure 3. (Figure 11 in revision) Regional Intensity (in mm)-versus- Duration (in hours) curve for WHR. The red dashed line represents the best fit line. The circles indicate individual storm events; the circle in green shows antecedent rain events preceding 30-day before the landslides, whereas squares in brown show precedent rain events before the day of landslides. The lines parallel to the best-fit line show the rainfall intensity thresholds triggering landslides correspond to different exceedance probabilities, such as 1%, 5%, 10% and 20% levels. The associated equations for *ID* threshold equations are presented in **Table 7.**

Comment 23. L182: There are some recent examples in the rainfall threshold literature that advice against using mean intensity and duration when calculating thresholds, because the two variables are dependent on each other, as you explained here. Conversely, in the case of thresholds defined using cumulated rainfall and duration (ED), the two variables are not dependent on each other. For this reason, it is preferable to define ED thresholds, in which the two variables measure independent quantities.

Response: We agree. We point to the reviewer that although cumulated rainfall versus duration (ED) for analyzing shallow landslides have been recently shown in several literature, the review of the literature showed varied accumulation lengths that potentially trigger landslides around the globe typically varies from *n* = 1-, 3-,7-,10-,15-,18-, 30-day to 3-4-month (Kim et al., 1992; Pasuto and Silvano, 1998; Chleborad, 2003; Heyerdahl et al., 2003; Cardinali et al., 2006; Guzzetti et al., 2007; Bevacqua et al., 2021), where *n* denotes the number of days of accumulated periods, often ranging from days to months. Further, a few assessment studies could not showed any evidence of antecedent rainfall for initiation of landslides (Brand et al., 1984; Corominas and Moya, 1996; Aleotti, 2004; Corominas, 2000) since an old rainfall event typically displays less impact in triggering landslides than the recent one. Further, it has been shown that antecedent soil moisture conditions have profound impact in controlling rainfall-triggered landslides in temperate climate (Campbell, 1975; Keefer et al., 1987; Wieczorek, 1987), but in humid climate their role is elusive (Larsen and Simon, 1993). Furthermore, antecedent moisture conditions can be investigated from ranges of environmental variables and finding an optimal time lag, responsible for landslides, such as, average soil moisture conditions (Merz and Blöschl, 2003; Zink et al., 2017) and accumulated rainfall amount of *n*-day (Kim et al., 1992; Pasuto and Silvano, 1998; Chleborad, 2003; Heyerdahl et al., 2003; Cardinali et al., 2006; Guzzetti et al., 2007; Bevacqua et al., 2021). The relation between flow rate of water and initial saturation level of soil is often nonlinear, whilst the effect of antecedent soil moisture strongly depends on soil characteristics and geology. A detailed investigation of *n*-day accumulated rainfall control, a proxy for antecedent soil moisture on landslides require in-depth analyses, such as the role of soil moisture and extreme precipitations, with their relative contributions in initiating landslides. Further, a recent study has shown that if the seasonality in rainfall intensity and the seasonality of antecedent catchment wetness are in phase together, their resonance could lead to extreme basin wetness, resulting in flash floods and landslides (Manoj J et al., 2022; Macdonald et al., 2022). A detailed assessment of antecedent wetness and their synchronous and/or asynchronous (i.e., lead or lag) relation with

precipitation seasonality in controlling landslides can be considered as a potential future direction.

We have discussed these issues on page 14, in line 353 in the revised manuscript.

Comment 24. L188-190: I would suggest adding some details on the methods used for regionalization.

Response: We have regionalize rainfall ID threshold based on seven environmental attributes that include geographical and climatic features (latitude, longitude of rain gauges, elevation, mean seasonal rainfalls for all four seasons, i.e., winter, summer, monsoon and fall rainfall). First, we normalize the environmental attributes using the following equation:

$$Normalisation = \frac{X_{(i,j)} - Min(X_j)}{Max(X_j) - Min(X_j)} \quad (2)$$

Where *i* represents each station and *j* represents each attribute.

Next, we apply the Principal Component Analysis (PCA) to reduce the dimension of the feature space while preserving the information content of the dataset. In our analysis, we reduce the dimension of the feature space from seven to five (i.e, we get 5 PCs). Out of five PCs, three PCs together explains ~ 95% variability of the regional characteristics. Hence we use the first three PCs for regionalization. The PCA bipot of top two PCs show the weightage of each variable towards the PCs. We find that winter rainfall has the largest contribution to the first PC, followed by the elevation and summer rainfall, while the monsoon rainfall and longitude have notable contributions to the second PC. Finally, we apply Fuzzy c-means clustering method (Ross, 2005; Bezdek, 1980; Sadri and Burn, 2011), which minimizes the inter-cluster variance and group the sites by assigning the membership to each feature vector with respect to the Euclidean distance between the feature vector and the cluster centre. Using the PCA and Fuzzy c-means clustering method, we find two distinct regions based on similar climatological and physiographical properties across the WHR.

We have discussed these in detail in L188-195 in the revised version of the manuscript.

Comment 25. How we can see that the region 1 is related to Banihal from figure 3?

Response: We regionalize WHR region by applying the PCA followed by the Fuzzy c-means clustering method (Ross, 2005; Bezdek, 1980; Sadri and Burn, 2011). The PCA reduce the dimension of the feature space while preserving the information content of the dataset. We find

that, on applying the PCA, the dimension of the feature space was reduced to five (i.e, we get 5 PCs) from seven. Three PCs together explains ~ 95% variability of the regional characteristics. The PCA bi-plot (Figure 3 in response document and the manuscript revision) of top two PCs shows the weightage of each variable towards the PCs. We find that winter rainfall has the largest contribution to the first PC, followed by the elevation and summer rainfall, while the monsoon rainfall and longitude have notable contributions to the second PC. Keeping in view of the reviewer's concern, we have revised the Figure as shown below:

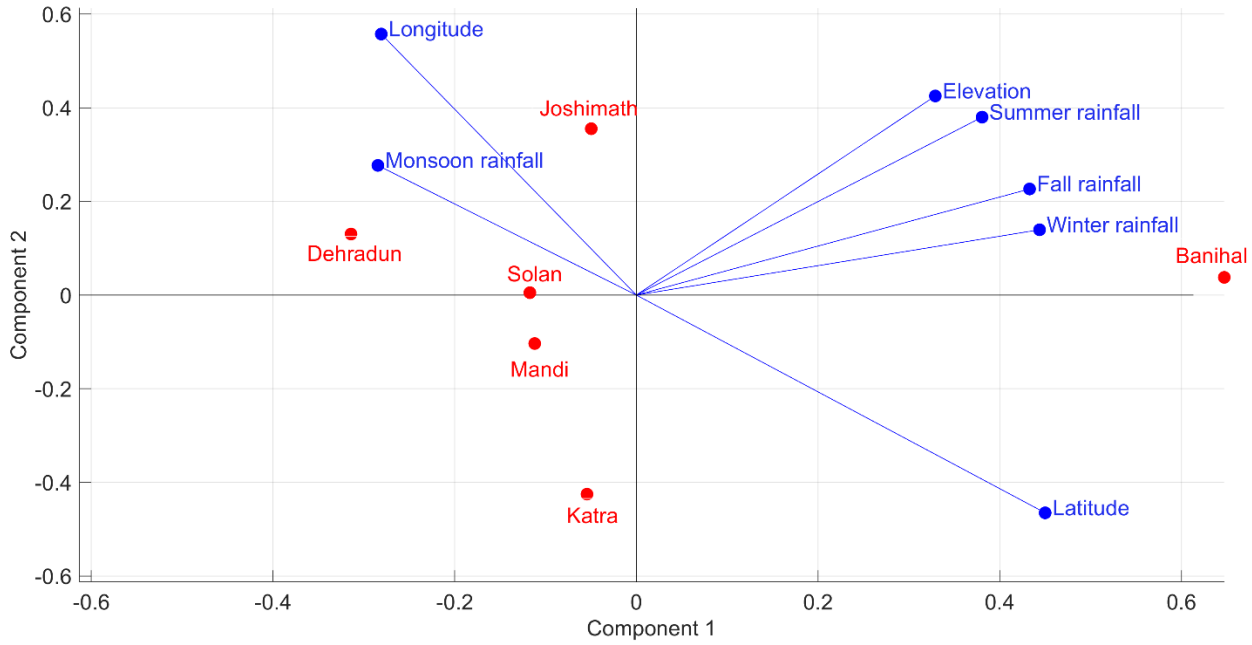

Figure 4. (Figure 3b in revision): Biplot of two primary principal components

Comment 26. L203-204: "The seasonal rainfall variability across the gauged sites shows the occurrence of peak rainfall during the monsoon season for most of the locations except Banihal". This is a very high difference, in particular comparing to the Katra station which is located 50 km far from Banihal and with an elevation difference of about 800 m. This deserves a better explanation.

Response: We appreciate the feedback. We have incorporated the following explanations in the revised manuscript:

"Although Katra and Banihal are located at a distance of around 50 km, an elevation difference of more than 790 m above the Mean Sea Level exists between two stations. While over India, approximately 80% of annual average rainfall occurs during the southwest monsoon season from June-September months (Soman and Kumar, 1990), the precipitation climatology of high

elevation areas of the Himalayas differs from the rest of the Indian subcontinent (Barros and Lang, 2003). Due to its location and elevation, its mountain chain acts as a barrier, and the evolution of Western Disturbances from the Mediterranean sea (Hunt et al., 2018) is the major contributor to snowfall/rainfall during winter. Recently, (Battula et al., 2021) showed a positive precipitation anomaly at lower elevations of the WHR during southwest monsoon months, but higher elevations in winter exhibit intense western disturbances, enabling orographic ascent of tropical moisture at higher elevations."

Comment 27. However, looking at table 1, I see a lot of differences in the annual and seasonal average rainfall. Katra station has very few data so I would not include it in the analyses, because it could affect too much negatively the comparison.

Response: We appreciate the feedback. We would like to point to the reviewer that we are reconstructing the daily rainfall for the whole analysis period (1970-2019) using regularized EM algorithm and high-resolution gridded IMD rainfall products. The skill of precipitation reconstruction show reasonably well fit of observed versus reconstructed time series, which are discussed in detail in the earlier version of the manuscript. To summarize, despite low sparse temporal coverage, our machine learning-based algorithm prove to be efficient in infilling gaps and reconstructing the precipitation series, which aids in further analysis of rainfall *ID* threshold relation.

Comment 28. L204: I'm not sure that you have described in the text how you have built this graph shown in Figure 4. Perhaps a clarification would be useful.

Response: Figure 4 represents the seasonal variation in rainfall climatology from 1985-2017, which is common for all stations. To investigate seasonal variations in at-site rainfall distributions, we calculate the standardized anomaly, which we determine by dividing the anomalies by the climatological standard deviation (Eq. 3). This provides more information about the magnitude of anomalies since the influence of dispersion have been removed. Here we find the rainfall anomaly $(X - \bar{X})$ of each month from 1985-2017 and then standardize it using the following equation:

$$standardized\ anomaly = \frac{(X-\bar{X})}{\sigma_X} \quad (3)$$

Where X stands for sum of monthly rainfall; $\bar{X}$ stands for mean of X and σ_X stands for the standard deviation of X. We incorporate this information in the main text as well as in figure 4 caption in the revised manuscript.

Comment 29. L212: Actually, I don't see this strong synchronicity. Perhaps it would be better to calculate an index for justify this link, as e.g. the Kendall rank correlation coefficient or the Pearson or the Nash-Sutcliffe (1970) coefficient.

Response: Agreed, we have reported the Kendall's τ rank correlation between frequency (number) of landslides and monthly rainfall in the revised manuscript:

Region 1 = 0.35 with p-value of 2.2×10^{-6}

Region 2 = 0.48 with p-value of 1.06×10^{-12}

The statistical significance value (p-value) of the Kendall's test suggests a strong correlation between the monthly rainfall and the frequency of landslides with p-value < 0.05.

Comment 30. L218: Actually, you didn't identify these landslide events. You used a global catalog.

Response: Agreed. We have revised the statement as below:

Across the WHR, we have filtered out more than 500 landslide events resulting from heavy rainfall, tropical cyclone-induced storm, and snowfall-snowmelt events, archived at the global landslide catalogue maintained by NASA COOLR (Kirschbaum et al., 2015; Juang et al., 2019) from 2007 to 2016. However, we could not point out any landslides from 1988-2006 in the global landslide catalogue for the region, possibly due to the lack of credible records.

Comment 31. L220: If you could not identify any landslides in the period 1988-2006, I would not use rainfall data before 2007 I the aim of the work is to study the "predictability of rainfall-induced landslides" as stated in the title.

Response: The rainfall data before 2007 was used to investigate changes in regional rainfall distribution against long term climatology. For investigating rainfall climatology (i.e., long-term average) of the area as we need a minimum of 30 years rainfall records.

Comment 32. L222: I would not say around 2010-2013. Actually, the events are clustered in 2010 and 2013.

Response: Agreed and incorporated in the text.

"Record breaking landslide events are typically clustered in 2010 and 2013, out of which 2013 was the most devastating in terms of fatality (Figure 6b)."

Comment 33. L232: Figure 7 is not very clear. I would suggest removing the landslides and the donut charts from panel b, for allowing comparison among (a) and (b). The donut charts could be moved in figure 6 (perhaps in a new panel with the map with the landslide types). Figure 6 is related to landslides, while figure 7 is related to rainfall. So please avoid confusion.

Response: Agreed and incorporated (Figures 5-6 in the response document and Figures 6-7 in the revision).

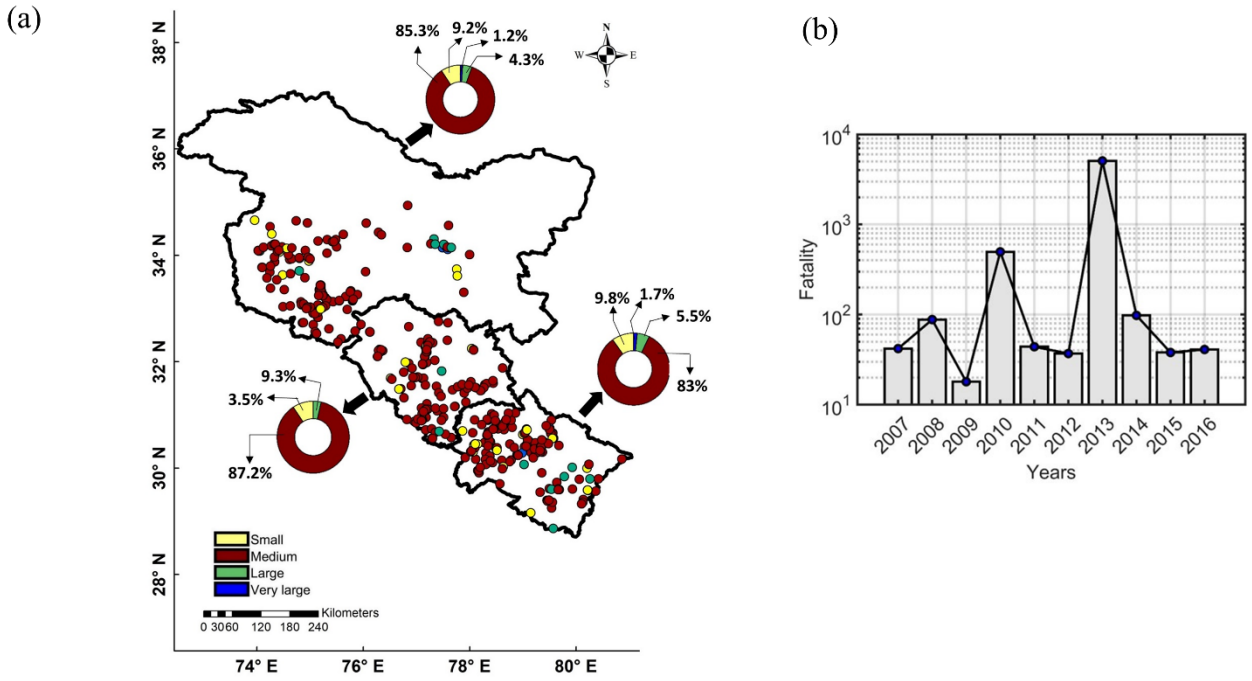

Figure 5. (Figure 6 in revision) Spatial distribution of historical landslides versus the mortality. **(a)** Landslide inventory map for the period 2007-2016 and **(b)** temporal distribution of fatality.

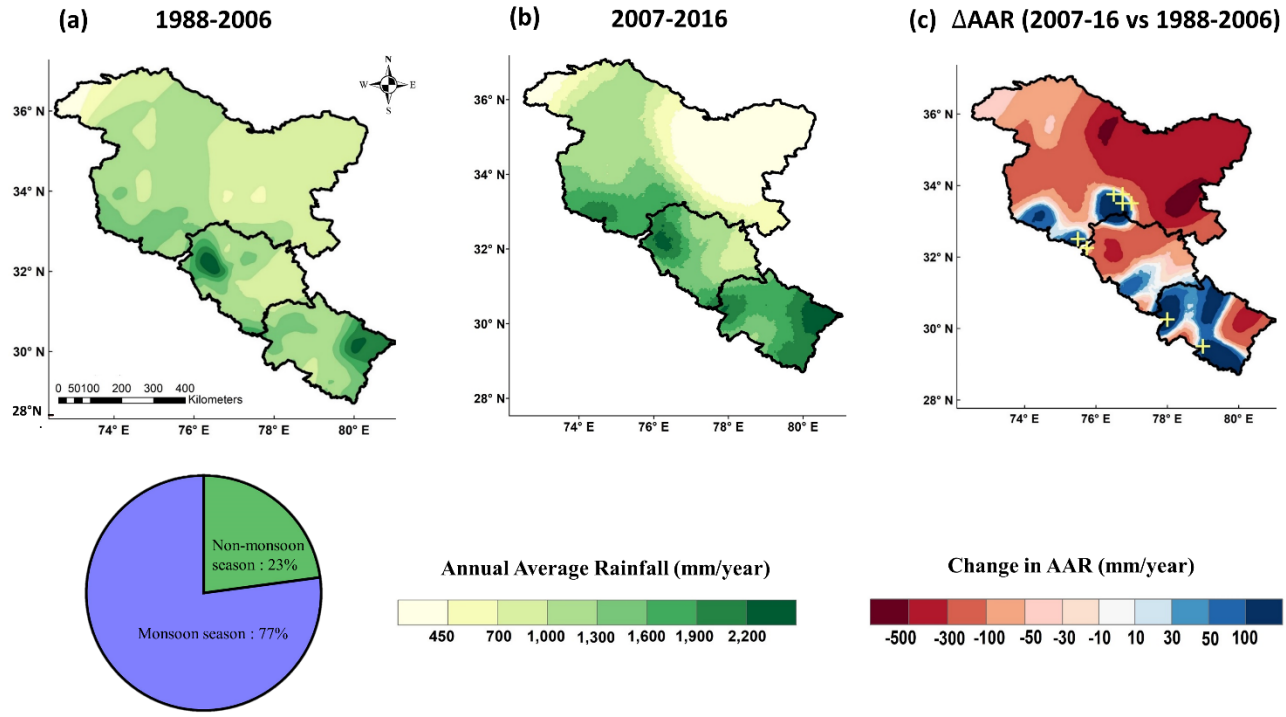

Figure 6 (Fig 7. in revision): Spatial distributions of rainfall over the WHR. Spatial distributions of rainfall during (a) 1988-2006 (b) 2007-2016 and (c) changes in annual rainfall climatology during the recent (2007-2016) versus retrospective (1988-2006) periods. The rain grids with significant positive median changes (2007-2016 versus 1988-2006) are marked in '+' symbol.

Comment 34. L235: Again about figure 7, the six stations are localized all in the southern part of the states. So I think it's a bit difficult to extrapolate the rainfall maps for all the northern, mountainous parts unless you used gridded data for the whole area. But this is not specified in the data section.

Response: The spatial distribution of rainfall maps are developed using high-resolution gridded rainfall products with spatial resolution 0.25°×0.25° archived at Climate Data Service Portal of the IMD (https://cdsp.imdpune.gov.in/home_gridded_data.php). No attempt was made to extrapolate station-based rainfall records to avoid any error associated with interpolation. We have included the related information in the revised manuscript's figure caption.

Comment 35. L240: You write about an increase but you don't have landslide data for the previous period.

Response: We agree. We have revised the sentence as below:

"Comparative assessment of spatial trends in rainfall patterns with few localized increases in rainfall distribution (Figure 7c) in the recent (2007-2016) versus the past decades (1988-2006) suggests an increase in rainfall in recent years, especially over the low elevated terrains, could potentially link to amplifications in landslide frequency (Figure 7b). Although no attempt has been made to relate the observed trends in present-day landslide frequency (2007-2016) against the past decades (1988-2006) due to a lack of credible observational records, a notable increase in the number of landslides in recent days is reported by several studies, which were linked to extreme and prolonged rainfall at seasonal and inter-annual time scales (Kirschbaum et al., 2020, 2012; Stanley et al., 2020). Recently, Stanley et al. (2020) have shown that landslides tend to become frequent in the Himalayas after local precipitation exceeds historical maximum accumulation. They have reported a positive but highly variable relationship between the number of landslides and an increase in precipitation. Furthermore, Kirschbaum et al. (2020) reported a projected increase in high-intensity (> 20 mm/day) daily precipitation over high-mountain Asian regions during the southwest monsoon season."

Comment 36. L241: How these landslide sizes were determined? What do small, medium, large, and very large mean?

Response: The landslide size is determined by the volume of landslide as per the standard typology (See Table 2 in Juang et al., 2019):

- Small landslides: Volume - <10 cubic meters
- Medium landslides: Volume - 10 to <1000 cubic meters
- Large landslides: Volume - 1000 to <100,000 cubic meters
- Very large landslides: Volume - 100,000 to <1 million cubic meters

We have incorporated these information in Data section and Figure 6 caption.

Comment 37. L244-245: What were the results of this Wilcoxon rank-sum test?

Response: Agreed, we have revised this section and added a new Table A2 listing the results of Wilcoxon rank-sum test for individual grid points. The table shows the grid points which has positive changes in median of rainfall of the two periods and p-value<0.1.

"To further investigate whether the 'local' (i.e., at individual rain grids) changes in rainfall pattern are statistically significant, we perform the non-parametric Wilcoxon rank-sum test, that compare the similarity of distributions between two independent samples of unequal sizes.

Using the rank-based Wilcoxon rank-sum test, we examine whether the rainfall climatology of the recent (2007-2016) versus past (1988-2006) time window is statistically significant at a 10% significant level. Stipples ('x' in yellow) in Figure 7c indicates rain grids where the differences in annual average rainfall are locally significant. Table A2 show changes in annual average rainfall for these identified grids and p-values obtained from the Wilcoxon rank-sum test. The significant grid points are identified when p-value is less than 0.10."

Table 4*: Rain-grids with significant positive changes in recent (2007-2016) versus past (1988-2006) rainfall magnitude

Longitude	Latitude	Difference in annual average rainfall during 2007-2016 versus 1988-2006	p-value
73	29	3.51	0.018903*
73	29.25	5.15	0.000574*
73	29.5	3.97	0.024672*
73	29.75	3.88	0.014107*
73.25	29	4.95	0.006219*
73.25	29.25	4.55	0.006848*
73.25	29.5	3.00	0.015452*
73.25	29.75	3.96	0.007458*
73.25	30	4.30	0.004701*
73.5	29	4.35	0.009955*
73.5	29.25	2.91	0.007298*
73.5	29.5	5.45	5.51E-05*
73.5	29.75	6.37	0.005262*
73.5	30	6.45	5.27E-05*
73.5	30.25	6.58	1.77E-05*
73.75	29	3.82	0.001695*
73.75	29.25	3.59	0.001666*
73.75	29.5	5.79	9.54E-06*
73.75	29.75	6.47	5.46E-05*
73.75	30	6.54	9.45E-05*
73.75	30.25	3.61	0.014701*
73.75	30.5	3.62	0.097975
74	29	5.27	0.000306*
74	29.25	6.96	0.000137*
74	29.5	8.62	4.36E-07*
74	29.75	5.59	0.010609*
74.25	29	5.08	0.006364*
74.25	29.25	8.18	1.18E-05*
74.25	29.5	8.76	8.88E-07*
74.25	29.75	7.43	0.000315*
74.5	29	6.69	0.002613*
74.5	29.25	8.53	5.69E-06*

74.5	29.5	5.17	0.003131*
74.5	29.75	5.90	0.000865*
74.5	30.75	4.26	0.044295*
74.75	29	5.31	0.013756*
74.75	29.25	7.95	0.000175*
75	29	2.16	0.021106*
75	29.25	4.55	0.019964*
75	30.25	2.27	0.05732
75.25	29.75	5.89	0.096492
75.5	32.25	25.47	0.005603*
75.5	32.5	23.87	0.047637*
75.75	30.25	5.57	0.018998*
75.75	30.5	5.20	0.018181*
75.75	32.25	17.27	0.025829*
76	30	9.35	0.000153*
76	30.75	3.72	0.060989*
76.25	30.75	9.80	0.039569*
76.5	30.75	15.81	0.00195*
76.5	33.75	26.32	0.010673*
76.75	33.5	23.68	0.013975*
76.75	33.75	18.53	0.009713*
77	33.5	15.26	0.095401
78	30.25	8.29	0.076716
79	29.5	14.44	0.045953*

*significant at a 10% significance level with p-value < 0.10. The p-values are obtained from Wilcoxon rank sum test. The null hypothesis is that rainfall time series of 2007-16 versus 1988-2006 are samples from continuous distributions with equal medians.

Comment 38. L257: You write “little to no improvement”. Actually, it seems there's a downgrade e.g. for Mandi station.

Response: Agreed. We have added a sentence here:

“In few cases, we observe downgrade in performance while applying the SPP-QM. For example, in Mandi station.”

Comment 39. L259: “Sparse temporal coverage” indeed, too few measurements.

Response: Agreed. Therefore, we are not using SPP-QM, which may impart an additional uncertainty in reconstructing rainfall series due to parameterization involve in Gamma probability density function.

Comment 40. L263-264: Again, this is highly questionable. Very often, it is not the event with the higher intensity that could be responsible for landslide initiation.

Response: Agreed and revised to mean intensity.

"We develop a regression relation between average rainfall intensity and duration for each site. We consider rain events of at most 30 days preceding landslide events to derive the power-law relation."

Comment 41. L267-278 "Whereas the other associated events with modest to low intensity are considered non-triggering events". This should be better justified.

Response: We have revised the analysis for mean intensity instead of maximum intensity. New results are completely different from the previous version of the manuscript. The related explanations have been changed.

Comment 42. L269-270: Something strange in Figure 9, in my opinion. How can you have a triggering event without a landslide (e.g. 2007 or 2016 in panel a) or after (e.g. 2015) a landslide? How can you have more triggering events that landslides in a period (e.g. 2007 in panel b)? Again, usually the triggering events are the ones that can be linked to the landslides, so most often are the ones immediately before the landslides.

Response: We appreciate the reviewer for pointing this issue. The figure is now revised. We present the figure for high flow seasons for the two regions. The red boxes show 'event with maximum rainfall intensity', identified 30-day before the day of landslides, the blue box indicates 'event with *NOT* a maximum rainfall intensity' and the red box with stipple indicates 'the day of landslide'. In earlier version, in panel (a) the landslide events were shown until the end of March. Hence, we could not point the day of landslide that occurred in April 4 (2016). The day of landslide is now added in the revised figure. Likewise, for year 2015, the day of landslide occurred on 30 April, which we have included in the revision. Panel (b) shows the region 2, which constitutes five rain gauge stations and compares temporal distribution of the maximum rain intensity versus landslide occurrences.

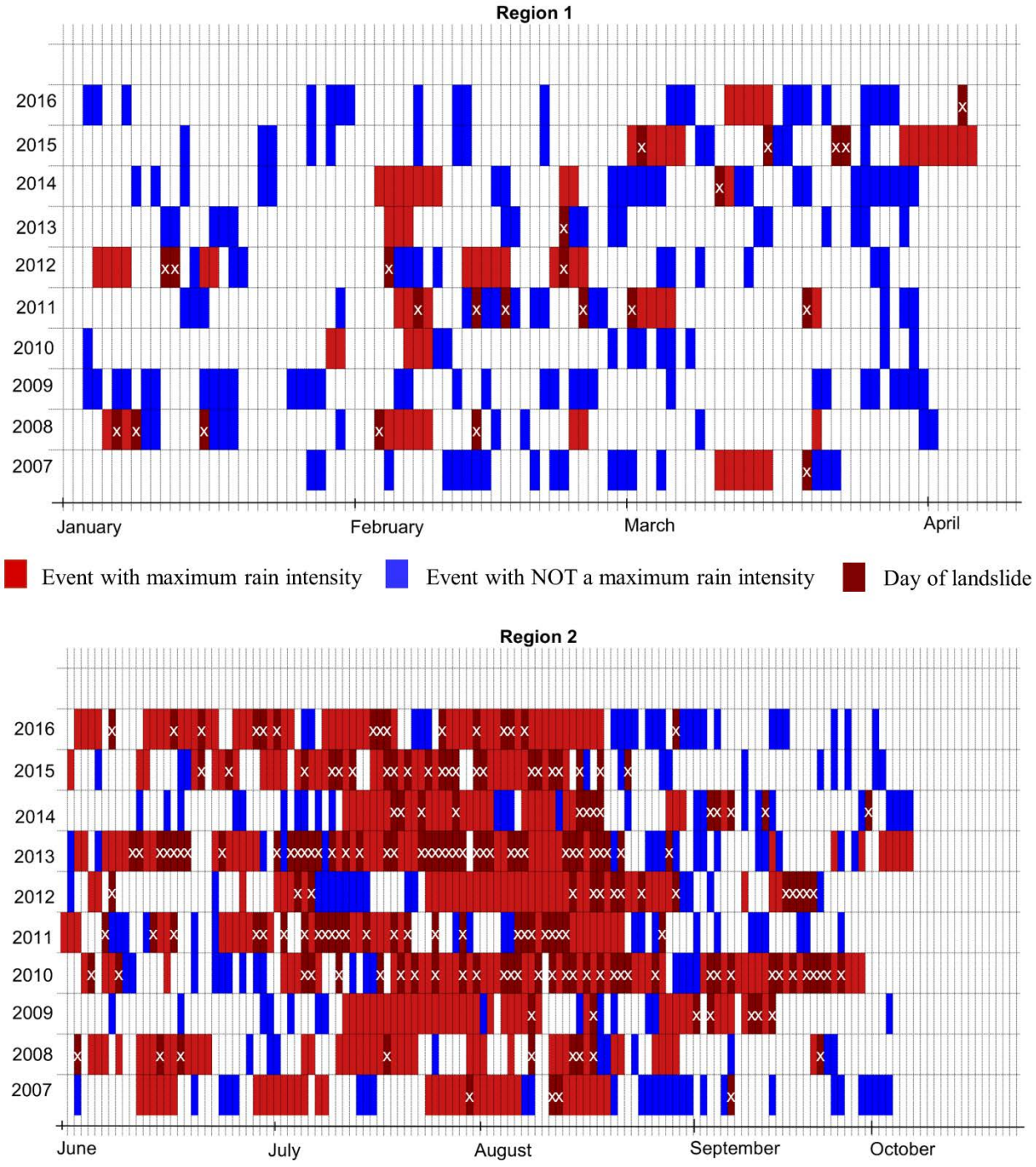

Figure 7 (Fig 9. in revision): Temporal contiguity of rain events with the maximum versus *NOT* a maximum rain intensity followed by landslide occurrence for identified regimes: (a) Region 1 showing winter (January-March) season considering the site Banihal, (b) Region 2, constituting five rain-gauge stations, depicting southwest monsoon (June-September) season. The day of landslides, unique for each station are marked in 'x'.

Comment 43. L275-276: This sentence is not clear.

Response: This statement suggests while earlier assessments (Mandal and Sarkar, 2021) showed 8-17% landslides occur within 3-5 days of heavy rainfall, we show that 22-26% landslides occur within a week of heavy rainfall events. We have revised the statements in the new version of the manuscript.

Comment 44. : L277: This overlap among triggering and non-triggering events can be due to the fact that the selection of the triggering events was made - in my opinion - in a wrong way.

Response: Agreed and now we have considered mean rainfall intensities instead of the maximum rain intensity to derive *ID* threshold curves. We have presented the revised figures as Figures 4-5 in the response letter.

Comment 45. L279: This intensity value is a bit strange. Do you think that such a low intensity is able to trigger a landslide in the area? Probably your data would need a post processing and analysis (allowing the elimination of very low values) before being use in threshold calculation.

Response: As suggested, in the revised analysis, we have eliminated the short durations low intensity values (0.5 mm/hr occurring in 24-48 hrs).

Comment 46. L281: 1.2mm – very low.

Response: Agreed. We have eliminated the very low rainfall magnitude. After elimination of lowest values, we get the minimum value as 8 mm.

Comment 47. L282: Is not the number of landslide events that decrease... is the number of rainfall conditions associated (or not) to landslide initiation.

Response: Agreed and incorporated.

"Further, the number of rain events tends to decrease with an increase in duration."

Comment 48. L282-283: "This implies landslides are triggered by short-duration high intensity rainfall". I would not make this general and strong statement.

Response: Agreed and incorporated.

Possibly the landslides are triggered by short-duration high intensity rainfall.

Comment 49. L284-285: It seems that you have not mentioned how many triggering events were associated to each station. This is crucial for threshold calculation.

Response: We agree. We would like to point to the reviewer that to avoid the subjectivity associated with the terms, 'triggering' and 'non-triggering', we have defined the "antecedent" and "precedent" rainfall events as suggested in the literature (Vallet et al., 2015). While "antecedent" rainfall corresponds to temporal clustering of rain events preceding a landslide considering a long (several weeks) time window of up to 30-days, the "precedent" rainfall events correspond to a short-term period, preferably a rainy day immediately before the occurrence of landslides. The above typology is physically consistent since the two periods do not exclusively overlap and the 'antecedent' period follows the 'precedent' period. Second, the choice of a 30-day time window is based on an antecedent soil moisture condition prevalent over a catchment that may trigger flash floods followed by landslides owing to compound occurrences of extreme rain on already saturated soil (Bertola et al., 2021). Further, in revision, we have we incorporated the mean rainfall intensity instead of the maximum intensity. In addition, we have separately shown preceding events immediately before landslides (see Figure 4-5 on pages 17-18 of this response document). Further, following Brunetti et al. (2010), we have adopted a frequentist procedure to derive rainfall *ID* thresholds for landslides.

Comment 50. L287-288: That's fine. However, such curves are not thresholds. They are best-fit curves of the point distribution. Please read some papers dealing with threshold calculation and define proper thresholds.

Response: We fully agree. We have completely revised our analysis. Following Brunetti et al. (2010), we have adopted a frequentist procedure to derive rainfall *ID* thresholds for landslides (See Figure 1 of this response document and Figure A1 in the revised manuscript for details). Following Brunetti et al. (2010), we calculate the residuals (i.e., observed records – predicted data points from the best-fit line). The Shaprio-Wilk's test suggests, except Banihal (p-value = 0.07), Katra (p-value = 0.075) and Joshimath (p-value = 0.013), the residuals are not normally distributed for remaining sites (p-value ranges between 1.5×10^{-5} and 4.8×10^{-5}). The heteroscedastic and non-Gaussian nature of residuals has often been the case for extremes, in which large prediction errors are associated with larger storms (Schoups and Vrugt, 2010). Hence, we model the residuals using empirical density function using Kernel density estimators with a normal kernel. We show the goodness-of-fit of the model using Kolmogorov-Smirnov

test (K-S test) in Table 5 (Table A3 in the appendix section of revised manuscript). Further, we compare the observed versus empirical distributions in figure 9 (Figure A2 in the revised manuscript). We then determine the distance, δ from the mean of the residual distributions fitted with kernel density function with normal kernel. Next, we obtain the shift of the threshold lines correspond to 20% and 30% exceedance probabilities (Abraham et al., 2022). The distance of exceedance probabilities from the mean is considered as the intercept of the threshold curves, while the slope remains the same as the best-fit line (Brunetti et al., 2010). Next, we determine regional ID curves corresponds to 1%, 5%, 10%, 20%, 30% and 50% exceedance probabilities. The detailed flowchart of this method is provided in the figure below (Figure 8; Figure A1 in the revised manuscript). Figure 10 shows the probability density function of the residual distributions depicting difference between observed and predicted rainfall intensity (Figure A3 in the revised manuscript). The Table 6 shows the threshold curves for each station, whereas Table 7 presents derived threshold curves for the whole WHR considering regionalization of rain gauges based on similarity measures of climatological attributes.

Figures 2 – 3 of this response letter show at-site and regional *ID* threshold relations. While at-site *ID* threshold curves are developed considering local/point-scale rainfall measurements, the regional relationship is developed by pooling point rainfall measurements of stations with similar hydroclimatic processes as determined by the Fuzzy c-means clustering algorithm. Figures 2 and 3 show thresholds fitted to the lower boundary of points reflecting approximate minimal rainfall intensity necessary to trigger landslides. As observed from the figures, most precedent rain events (*i.e.*, events immediately before the day of the landslide) are associated with short-duration and low-intensity rainfall, which trigger shallow landslides with small to medium-sized events. This is because the short-duration and low-intensity rainfall causes an excess pore pressure in shallow soil zones, leading to shallow landslides (Larsen and Simon, 1993). Our findings are further confirmed by the landslide size distribution in Table 3. Although for at-site, $T_{\text{min (at-site)}}$ = 20%- and for regional, $T_{\text{min (regional)}}$ = 1%- exceedance probability levels are the minimum thresholds below which no landslides are expected to occur, we find that except Banihal, Mandi and Solan, for remaining sites a few precedent *outlying* events that lie below the design threshold lines. This could be attributed to the uncertainty associated with the timing of the landslides (*e.g.*, 'entry date' instead of the actual timing of slide information in the Bhukosh archive) and the uncertainty involved in estimating missing rainfall records. Furthermore, the frequency (number) of landslide information within the

shortest radii around each station is also limited, which could have imparted an additional uncertainty to the analyses. For regional estimate (Figure 3), the individual at-site uncertainty in *ID*-threshold estimates are propagated. Nevertheless, we find more than 83% of precedent rain events lie above the minimum threshold line, suggesting the validity of the obtained ID threshold curves at the regional level.

Table 5*. Goodness-of-fit assessment for non-parametric Kernel density estimates in fitting residual error of linear regression fit. The goodness-of-fit is assessed via Kolmogorov-Smirnov (K-S) goodness-of-fit test

Stations	K-S distance (mm)	p-value
Banihal (Region 1)	0.1062	0.012
Katra	0.0981	0.1600
Mandi	0.0992	0.0120
Solan	0.0897	0.0220
Dehradun	0.0896	0.0380
Joshimath	0.0748	0.1100
Region 2	0.0539	0.212

*Table A3 in the appendix section of the revised manuscript. P-value > 0 indicates validity of the proposed distribution. Region 2 is derived by pooling rainfall information from climatologically similar stations.

Table 6: Intensity-Duration threshold relationships correspond to different exceedance probabilities.

Stations	10%	15%	20%	30%	50%
Banihal	-*	$I_{15} = 0.06D^{0.17}$	$I_{20} = 0.14D^{0.17}$	$I_{30} = 0.26D^{0.17}$	$I_{50} = 0.584D^{0.17}$
Katra	-	-	$I_{20} = 0.003D^{0.13}$	$I_{30} = 0.24D^{0.13}$	$I_{50} = 0.7D^{0.13}$
Mandi	$I_{10} = 1.09D^{-0.16}$	$I_{15} = 1.143D^{-0.16}$	$I_{20} = 1.18D^{-0.16}$	$I_{30} = 1.25D^{-0.16}$	$I_{50} = 1.48D^{-0.16}$
Solan	$I_{10} = 7.34D^{-0.59}$	$I_{15} = 7.43D^{-0.59}$	$I_{20} = 7.5D^{-0.59}$	$I_{30} = 7.64D^{-0.59}$	$I_{50} = 7.98D^{-0.59}$
Dehradun	-	$I_{15} = 0.07D^{0.08}$	$I_{20} = 0.203D^{0.08}$	$I_{30} = 0.438D^{0.08}$	$I_{50} = 1.126D^{0.08}$
Joshimath	$I_{10} = 1.16D^{-0.21}$	$I_{15} = 1.213D^{-0.21}$	$I_{20} = 1.253D^{-0.21}$	$I_{30} = 1.318D^{-0.21}$	$I_{50} = 1.445D^{-0.21}$

*'-' shows that threshold relations could not be computed due to the lack of sufficient observations

Table 7: Regional Intensity-Duration threshold relationships correspond to different exceedance probabilities for the whole WHR

Regions	1%	5%	10%	20%	30%	50%
Region 1*	-	-	$I_{10} = 0.066D^{0.17}$	$I_{20} = 0.14D^{0.17}$	$I_{30} = 0.26D^{0.17}$	$I_{50} = 0.584D^{0.17}$
Region 2	$I_1 = 0.795D^{-0.14}$	$I_5 = 0.9585D^{-0.14}$	$I_{10} = 1.0471D^{-0.14}$	$I_{20} = 1.1563D^{-0.14}$	$I_{30} = 1.241D^{-0.14}$	$I_{50} = 1.655D^{-0.14}$

* Region 1 contains only one station, i.e., Banihal. '-' shows that threshold relations could not be computed due to the lack of sufficient observations

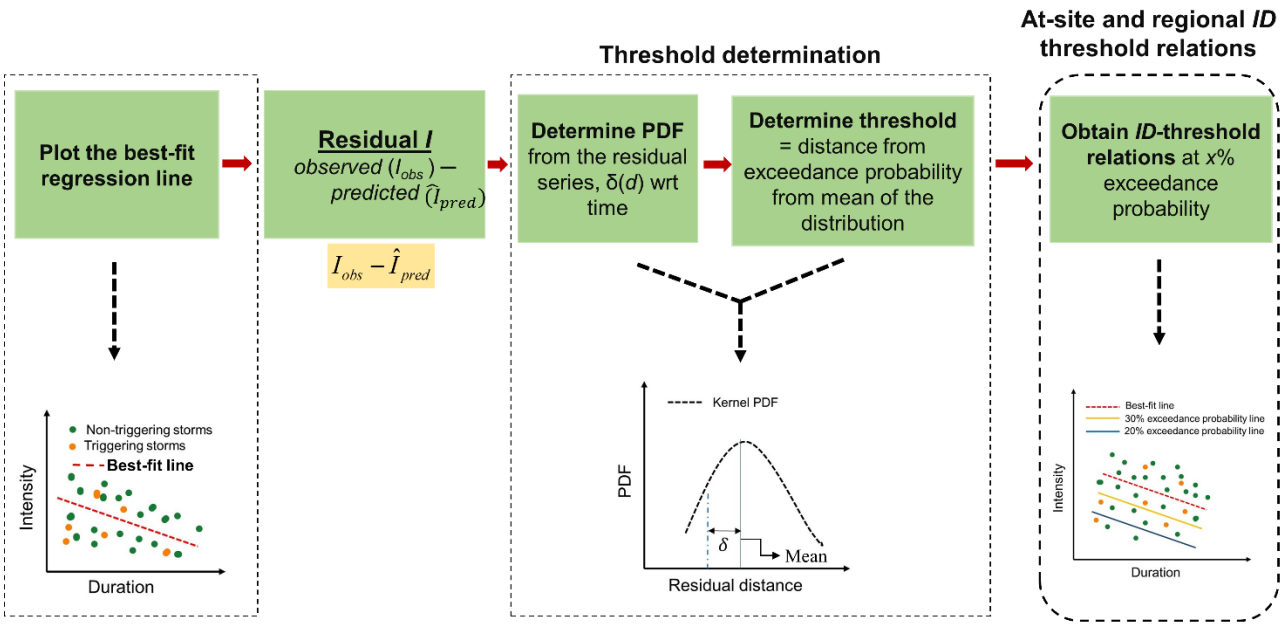

Figure 8 (Fig. A1 in revision): Detailed flowchart of the method to derive threshold curves from the best-fit lines.

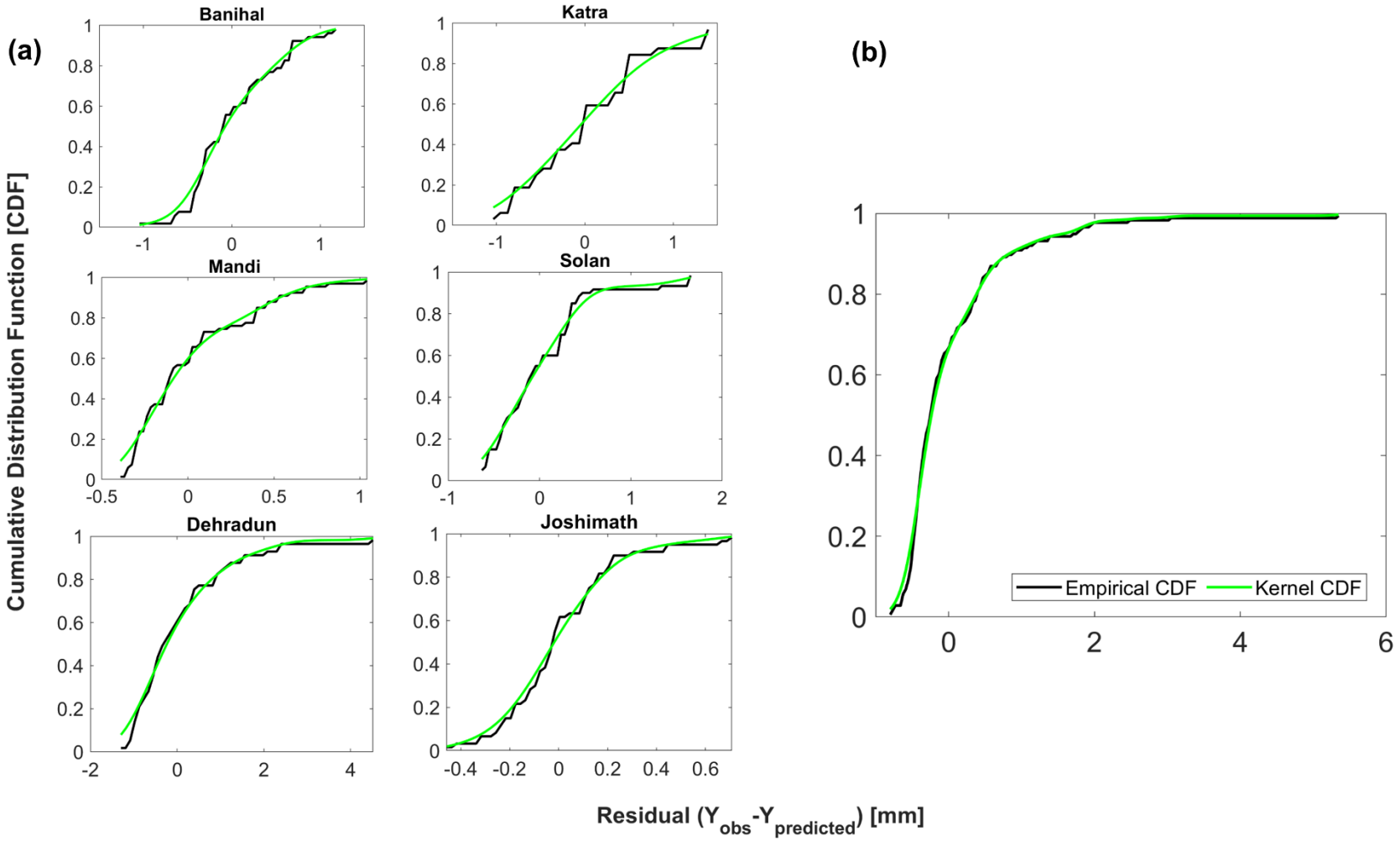

Figure 9 (Fig A2. in revision): Empirical Distribution functions (CDFs shown in black lines) versus the non-parametric Kernel distributions (CDFs shown in green lines) of residual rain intensity (observed records ─ predicted data points from the best-fit line) for (a) individual station locations (b) region 2, obtained by pooling rainfall measurements from climatologically similar locations. Empirical CDFs are determined using Weibull's plotting position formula.

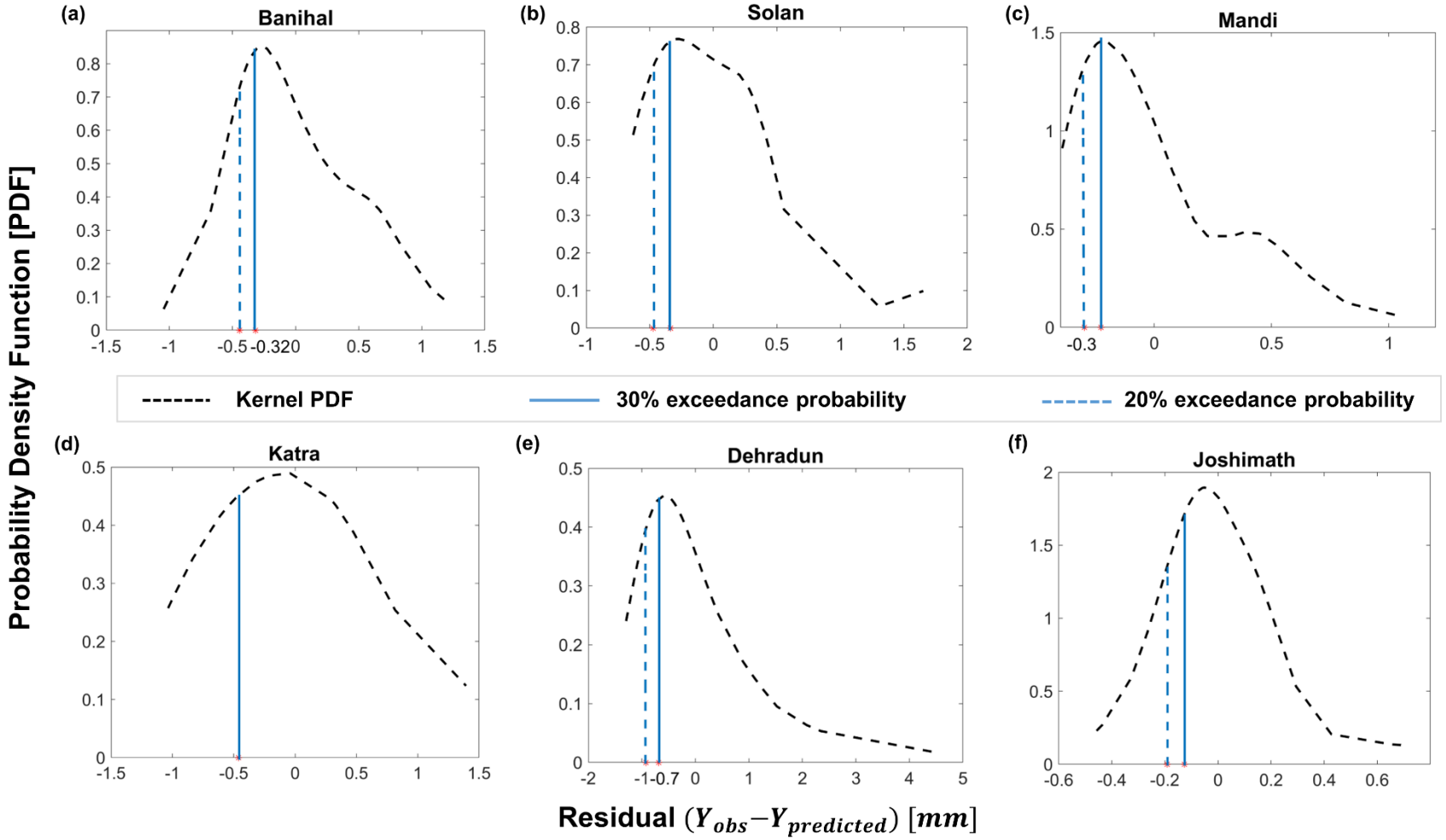

Figure 10 (Fig A3. in revision): Empirical Probability density functions (PDFs shown in dashed black lines) of residual rain intensity (observed records – predicted data points from the best-fit line) for each station. The vertical straight lines in solid and in dashed blue show the shift of the threshold lines with 20% and 30% exceedance probability from the best fit line. For Katra, only a 30% exceedance probability line is shown due to the limited number of observation points, resulting in negative intercept values at lower threshold levels.

Comment 51. L289: These values of the intercept are very low, even more if considering that the thresholds are not thresholds but actually best-fit curves. By making some calculations using eq. 3, one can obtain very low values of intensity as threshold values. Probably some low events should be removed from the analysis before calculating the thresholds. However, I still think that the problem lies in the method for the definition of landslide-triggering rainfall events.

Response: Agreed and the low intensities shorter duration events are removed. In addition, we have revised our method to compute the *ID* threshold relation using mean intensity.

Comment 52. L316: Only one station have this 51-year operational period.

Response: Here we would point to the reviewer that after reconstructing daily rainfall series, all sites are of uniform length and the rainfall time series varies from 1970-2019.

Comment 53. Table 1: I would use only one decimal place when referring to rainfall, unless you know there is this high measurement accuracy.

Response: Agreed and incorporated.

Comment 54. Figure 1: The quality of this figure could be improved. As an example, the map with the indication of the three states within the Indian continent should be placed as an inset of the other map. The North arrow and the scale bar could be placed within the map with the indication of the rain gauges.

Response: Agreed and revised as follows:

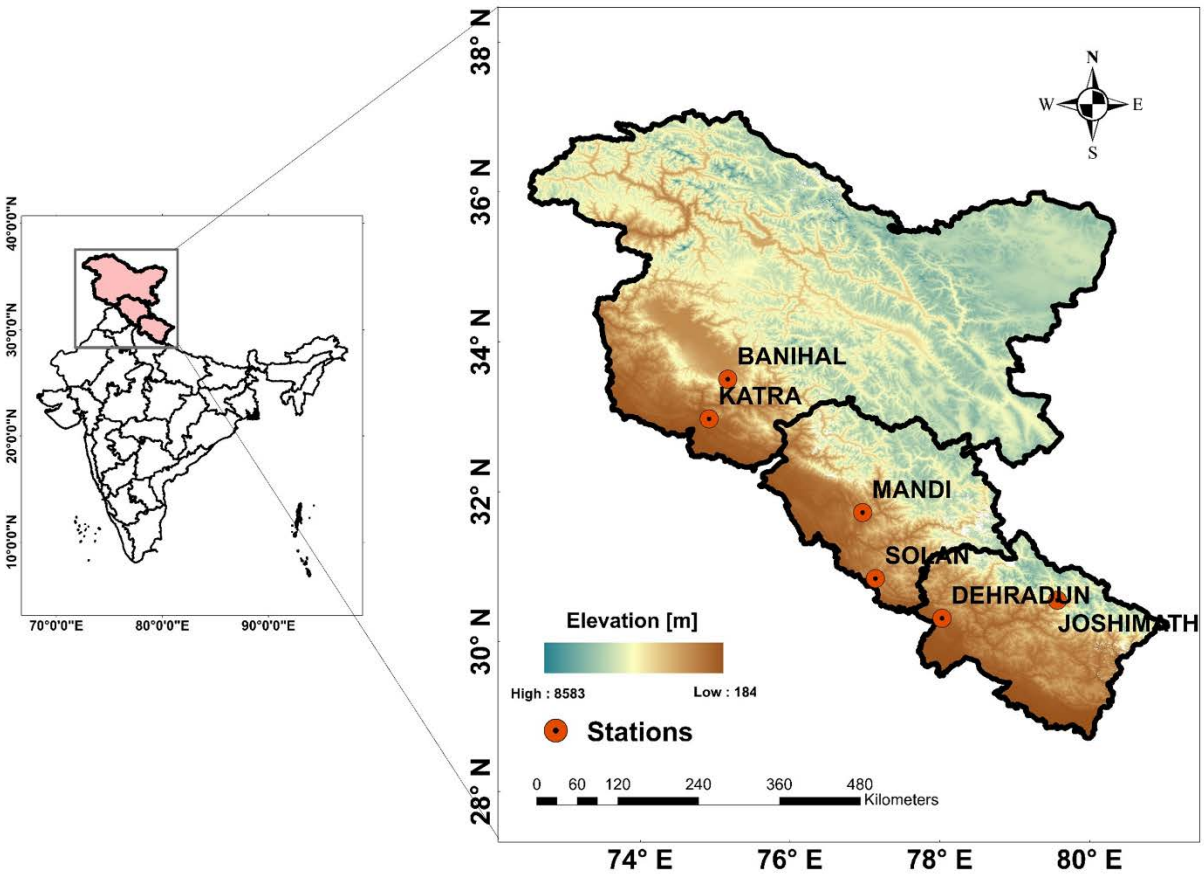

Figure 11 (Fig. 1 in revision): Spatial locations of rain gauges across the WHR (See Table 1 for details). The elevational profile shows the locations of high hills across the northern part of the WHR, which gradually decreases from the north to the south. The digital elevation model of 1-Arc second (approximately 30 m) spatial resolution was derived from the SRTM-1 Arc Second Global data product archived at USGS Earth explorer (https://earthexplorer.usgs.gov/). The elevation map of India is projected using spatial analysis software Arc GIS Desktop version 10.8.1. The inset shows the location of the WHR over the Indian subcontinent.

Comment 55. **Figure 3b**: I would add the name of the sites in the plot, for a better understanding.

Response: Agreed and incorporated as follows:

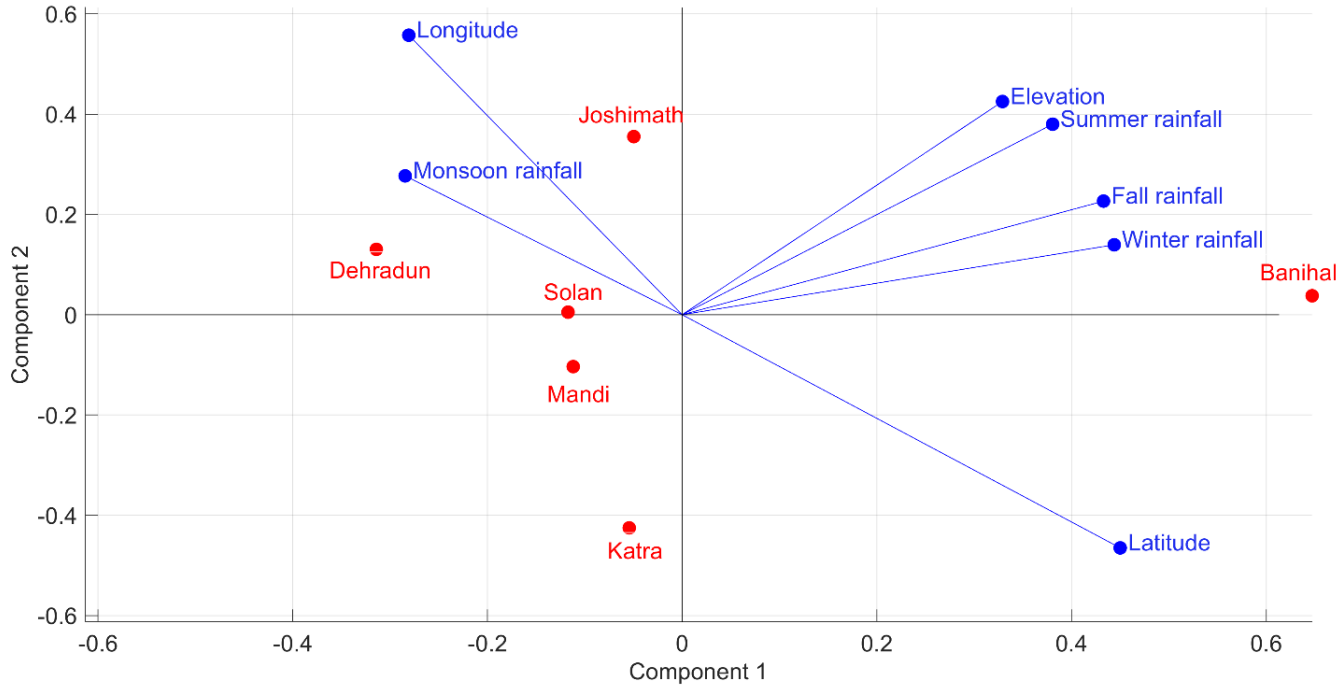

Figure 12 (Fig. 3b in revision): Biplot of the PCs. Red color dots indicate the clusters of the selected sites. Blue arrows indicate the weightage of different attributes in the PCA space.

Comment 56. Figure 5: I would suggest coloring in blue the bars of monthly precipitation to be in agreement with the scale label. I don't see the need for adding the smoothing. Please add justification. However, in the legend a "W" is missing.

Response: Agreed and revised as follows. Here we do not agree with the reviewer. Without smoothing, the trend line would appear as a jagged shape. The LOWESS regression smoothing of the scatter point of landslide events aids in deciphering the year-to-year variability/trend in the record. We have incorporated the word, 'W' in the revised figure.

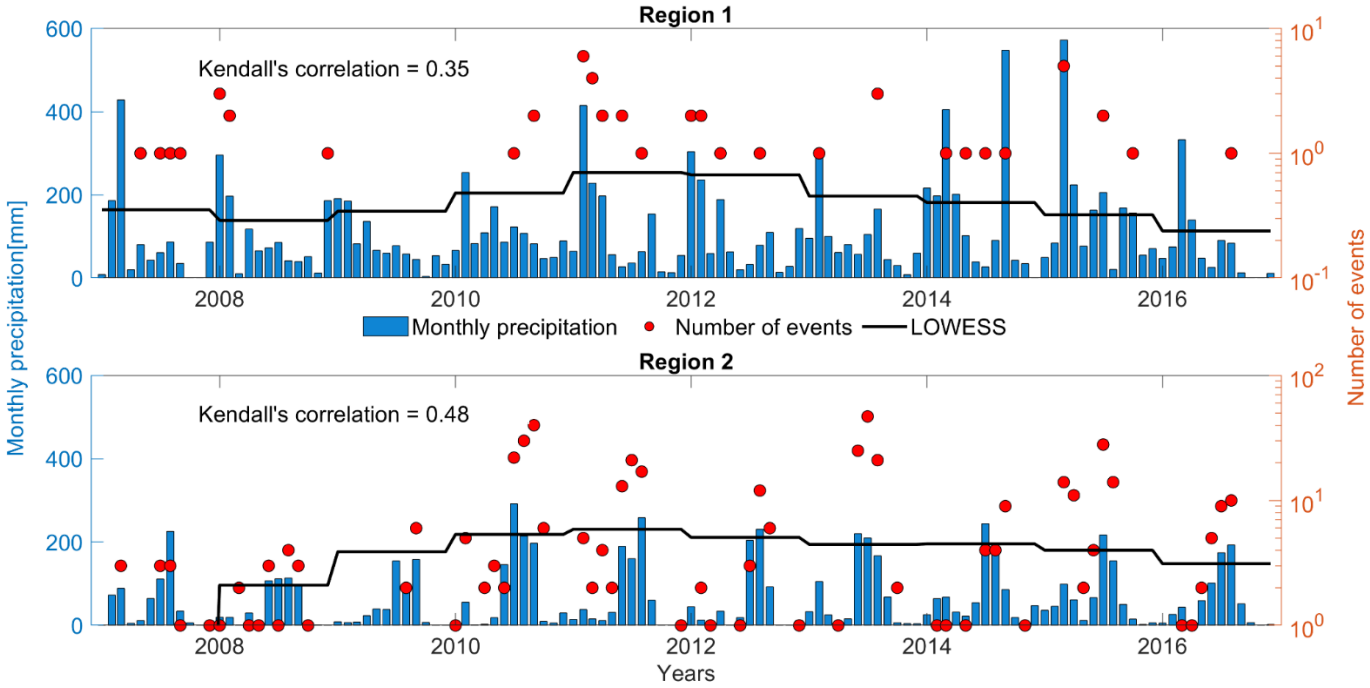

Figure 13 (Fig. 5 in revision): Temporal evolution of monthly precipitation versus the number of landslides for (a) Region 1 (b) Region 2. The landslide frequency (the number of events) are smoothed using the Locally weighted Scatter plot smoothing (LOWESS) regression, shown in solid thick lines in black with a span length of 0.75.

Comment 57. Figure 6. The North arrow and the scale bar could be placed within the map. The little map with the indication of the three Indian states could be removed as it is already shown in figure 1. However, this map is not clearly visible. I would remove the DEM in background. The legend could be placed inside the map.

Response: Agreed and revised accordingly.

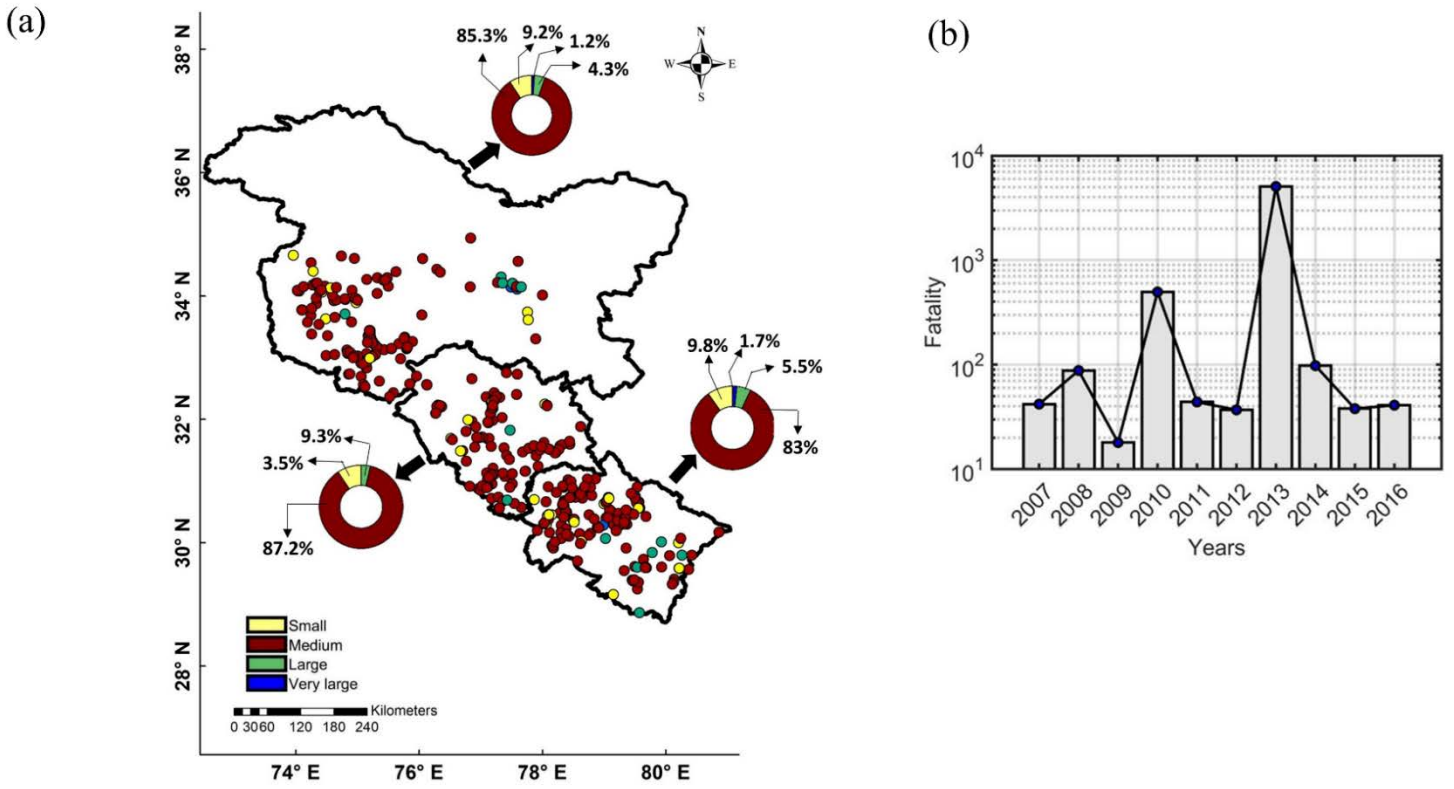

Figure 14 (Fig. 6 in revision): Spatial distribution of historical landslides versus the mortality. (a) Landslide inventory map for the period 2007-2016 and (b) temporal distribution of fatality.

Comment 58. Figure 7: The pie chart (monsoon vs. non-monsoon) should be described also in the caption. This figure 7a is not very clear. I would suggest removing the landslides and the donut charts from panel b, for allowing comparison among (a) and (b). The donut charts could be moved in figure 6 (perhaps in a new panel with the map with the landslide types). Figure 6 is related to landslides, while figure 7 is related to rainfall. So please avoid confusion. Finally, please describe the acronym AAR in the caption.

Response: Agreed and revised.

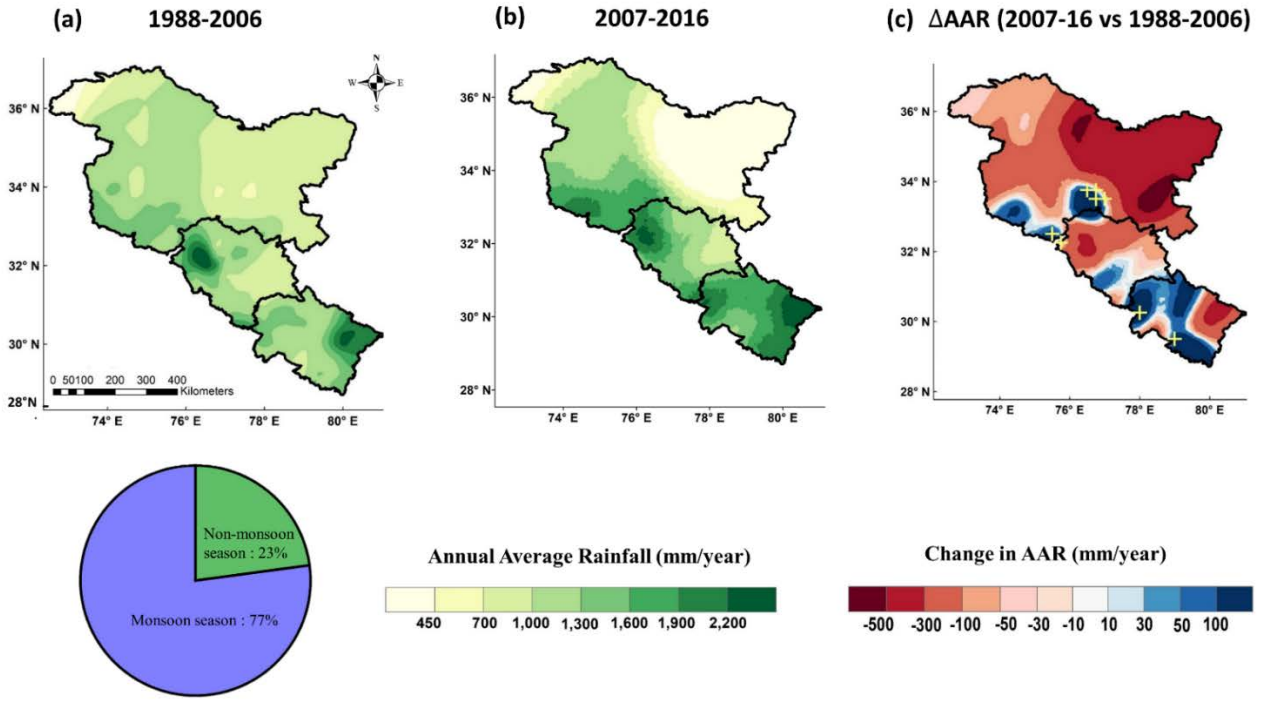

Figure 15 (Fig. 7 in revision): Spatial distributions of rainfall over WHR. Spatial distribution of rainfall during (a) 1988-2006 (b) 2007-2016 and (c) Changes in annual rainfall climatology during the recent (2007-2016) versus retrospective (1988-2006) periods. The grids with significant positive median changes (2007-2016 vs 1988-2006) are marked in '+' symbol.

Comment 59. Figure 10. I would leave only the names of the stations and remove the codes. Moreover, how we can understand which events are triggering and which non-triggering in these graphs? A legend and a description in the caption are needed.

Response: Agreed and revised:

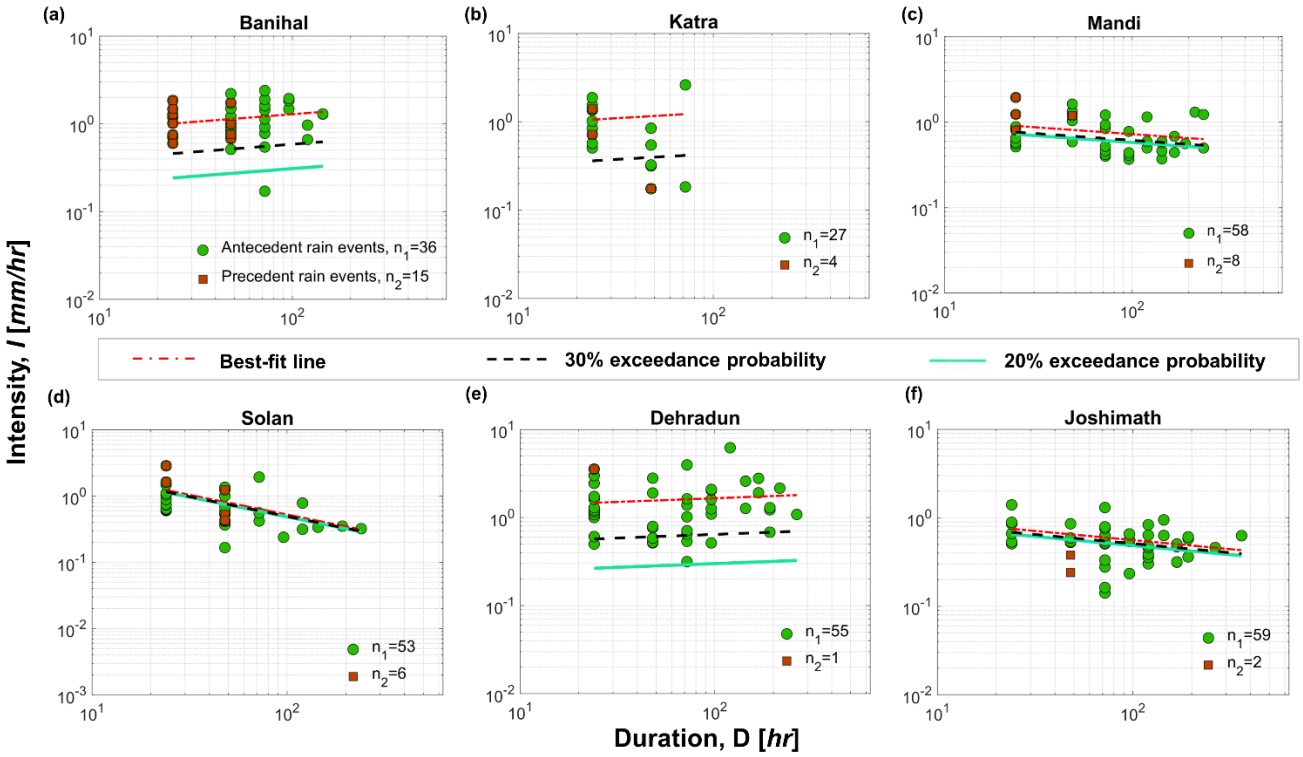

Figure 16 (Fig. 10 in revision) Relation between average rainfall intensity (in mm) -versus- duration (in hours) for storms dating from 2007 to 2019 in the WHR. The red dashed line represents the best-fit. The circles indicate individual storm events, the circle in green shows antecedent rain events preceding 30-day before the landslides, whereas squares in brown show precedent rain events before the day of landslides. The dotted-black and cyan line represent threshold curves corresponding to 20% and 30% exceedance probabilities respectively.

References

Abraham, M. T., Satyam, N., Pradhan, B., Segoni, S., and Alamri, A.: Developing a prototype landslide early warning system for Darjeeling Himalayas using SIGMA model and real-time field monitoring, Geosciences Journal, 26, 289–301, 2022.

Aleotti, P.: A warning system for rainfall-induced shallow failures, Engineering geology, 73, 247–265, 2004.

Barros, A. P. and Lang, T. J.: Monitoring the monsoon in the Himalayas: Observations in central Nepal, June 2001, Monthly Weather Review, 131, 1408–1427, 2003.

Battula, S. B., Siems, S., and Mondal, A.: Dynamical and thermodynamical interactions in daily precipitation regimes in the Western Himalayas, International Journal of Climatology, 2021, https://doi.org/10.1002/joc.7511, 2021.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Bellido-Jiménez, J. A., Gualda, J. E., and García-Marín, A. P.: Assessing Machine Learning Models for Gap Filling Daily Rainfall Series in a Semiarid Region of Spain, Atmosphere, 12, 1158, 2021.

Bertola, M., Viglione, A., Vorogushyn, S., Lun, D., Merz, B., and Blöschl, G.: Do small and large floods have the same drivers of change? A regional attribution analysis in Europe, Hydrology and Earth System Sciences, 25, 1347–1364, https://doi.org/10.5194/hess-25-1347-2021, 2021.

Bevacqua, E., De Michele, C., Manning, C., Couasnon, A., Ribeiro, A. F. S., Ramos, A. M., Vignotto, E., Bastos, A., Blesić, S., Durante, F., Hillier, J., Oliveira, S. C., Pinto, J. G., Ragno, E., Rivoire, P., Saunders, K., van der Wiel, K., Wu, W., Zhang, T., and Zscheischler, J.: Guidelines for Studying Diverse Types of Compound Weather and Climate Events, Earth's Future, 9, e2021EF002340, https://doi.org/10.1029/2021EF002340, 2021.

Bezdek, J. C.: A convergence theorem for the fuzzy ISODATA clustering algorithms, IEEE transactions on pattern analysis and machine intelligence, 1–8, 1980.

Brand, E. W., Premchitt, J., and Phillipson, H. B.: Relationship between rainfall and landslides in Hong Kong, in: Proceedings of the 4th International Symposium on Landslides, 276–84, 1984.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Natural Hazards and Earth System Sciences, 10, 447–458, https://doi.org/10.5194/nhess-10-447-2010, 2010a.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Natural Hazards and Earth System Sciences, 10, 447–458, 2010b.

Campbell, R. H.: Soil slips, debris flows, and rainstorms in the Santa Monica Mountains and vicinity, southern California, US Government Printing Office, 1975.

Cardinali, M., Galli, M., Guzzetti, F., Ardizzone, F., Reichenbach, P., and Bartoccini, P.: Rainfall induced landslides in December 2004 in south-western Umbria, central Italy: types, extent, damage and risk assessment, Natural Hazards and Earth System Sciences, 6, 237–260, 2006.

Chleborad, A. F.: Preliminary evaluation of a precipitation threshold for anticipating the occurrence of landslides in the Seattle, Washington, Area, US Geological Survey open-file report, 3, 39, 2003.

Corominas, J.: Landslides and climate, in: Proceedings of the 8th International Symposium on Landslides, 1–33, 2000.

Corominas, J. and Moya, J.: Historical landslides in the Eastern Pyrenees and their relation to rainy events, in: Landslides, Proceedings of the Eight International Conference and Field Trip on Landslides. Balkema, Rotterdam, 125–132, 1996.

Dempster, A. P., Laird, N. M., and Rubin, D. B.: Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society: Series B (Methodological), 39, 1–22, 1977.

Dikshit, A., Sarkar, R., Pradhan, B., Acharya, S., and Dorji, K.: Estimating Rainfall Thresholds for Landslide Occurrence in the Bhutan Himalayas, Water, 11, 1616, https://doi.org/10.3390/w11081616, 2019.

Feng, S., Hu, Q., Wu, Q., and Mann, M. E.: A Gridded Reconstruction of Warm Season Precipitation for Asia Spanning the Past Half Millennium, Journal of Climate, 26, 2192–2204, https://doi.org/10.1175/JCLI-D-12-00099.1, 2013.

Ganguli, P. and Coulibaly, P.: Does Nonstationarity in Rainfall Requires Nonstationary Intensity-Duration-Frequency Curves?, Hydrol. Earth Syst. Sci. Discuss., 2017, 1–31, https://doi.org/10.5194/hess-2017-325, 2017.

Gariano, S. L., Sarkar, R., Dikshit, A., Dorji, K., Brunetti, M. T., Peruccacci, S., and Melillo, M.: Automatic calculation of rainfall thresholds for landslide occurrence in Chukha Dzongkhag, Bhutan, Bull Eng Geol Environ, 78, 4325–4332, https://doi.org/10.1007/s10064-018-1415-2, 2019.

Gebremichael, M., Krajewski, W. F., Morrissey, M. L., Huffman, G. J., and Adler, R. F.: A detailed evaluation of GPCP 1 daily rainfall estimates over the Mississippi River Basin, Journal of Applied Meteorology and Climatology, 44, 665–681, 2005.

Bhukosh: https://bhukosh.gsi.gov.in/Bhukosh/MapViewer.aspx, last access: 16 July 2022.

Giannecchini, R., Galanti, Y., and D'Amato Avanzi, G.: Critical rainfall thresholds for triggering shallow landslides in the Serchio River Valley (Tuscany, Italy), Natural Hazards and Earth System Sciences, 12, 829–842, https://doi.org/10.5194/nhess-12-829-2012, 2012.

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorology and atmospheric physics, 98, 239–267, 2007.

Harilal, G. T., Madhu, D., Ramesh, M. V., and Pullarkatt, D.: Towards establishing rainfall thresholds for a real-time landslide early warning system in Sikkim, India, Landslides, 16, 2395–2408, 2019.

Heyerdahl, H., Harbitz, C. B., Domaas, U., Sandersen, F., Tronstad, K., Nowacki, F., Engen, A., Kjekstad, O., Devoli, G., and Buezo, S. G.: Rainfall induced lahars in volcanic debris in Nicaragua and El Salvador: practical mitigation, in: Proceedings of international conference on fast slope movements—prediction and prevention for risk mitigation, IC-FSM2003. Patron Pub, Naples, 275–282, 2003.

Hindustan Times: More than 12% of India's land mass is prone to landslides: GSI, 2016.

Hunt, K. M. R., Turner, A. G., and Shaffrey, L. C.: The evolution, seasonality and impacts of western disturbances, Quarterly Journal of the Royal Meteorological Society, 144, 278–290, https://doi.org/10.1002/qj.3200, 2018.

Johnston, E. C., Davenport, F. V., Wang, L., Caers, J. K., Muthukrishnan, S., Burke, M., and Diffenbaugh, N. S.: Quantifying the Effect of Precipitation on Landslide Hazard in Urbanized and Non-Urbanized Areas, Geophysical Research Letters, 48, e2021GL094038, https://doi.org/10.1029/2021GL094038, 2021.

Juang, C. S., Stanley, T. A., and Kirschbaum, D. B.: Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR), PloS one, 14, e0218657, 2019.

Kalteh, A. M. and Hjorth, P.: Imputation of missing values in a precipitation–runoff process database, Hydrology Research, 40, 420–432, 2009.

Kanungo, D. P. and Sharma, S.: Rainfall thresholds for prediction of shallow landslides around Chamoli-Joshimath region, Garhwal Himalayas, India, Landslides, 11, 629–638, 2014.

Keefer, D. K., Wilson, R. C., Mark, R. K., Brabb, E. E., Brown III, W. M., Ellen, S. D., Harp, E. L., Wieczorek, G. F., Alger, C. S., and Zatkin, R. S.: Real-time landslide warning during heavy rainfall, Science, 238, 921–925, 1987.

Khan, Y. A., Lateh, H., Baten, M. A., and Kamil, A. A.: Critical antecedent rainfall conditions for shallow landslides in Chittagong City of Bangladesh, Environ Earth Sci, 67, 97–106, https://doi.org/10.1007/s12665-011-1483-0, 2012.

Kim, S. K., Hong, W. P., and Kim, Y. M.: Prediction of rainfall-triggered landslides in Korea, in: International symposium on landslides, 989–994, 1992.

Kirschbaum, D., Adler, R., Adler, D., Peters-Lidard, C., and Huffman, G.: Global Distribution of Extreme Precipitation and High-Impact Landslides in 2010 Relative to Previous Years, Journal of Hydrometeorology, 13, 1536–1551, https://doi.org/10.1175/JHM-D-12-02.1, 2012.

Kirschbaum, D., Stanley, T., and Zhou, Y.: Spatial and temporal analysis of a global landslide catalog, Geomorphology, 249, 4–15, https://doi.org/10.1016/j.geomorph.2015.03.016, 2015.

Kirschbaum, D., Kapnick, S. B., Stanley, T., and Pascale, S.: Changes in Extreme Precipitation and Landslides Over High Mountain Asia, Geophysical Research Letters, 47, e2019GL085347, https://doi.org/10.1029/2019GL085347, 2020.

Larsen, M. C. and Simon, A.: A rainfall intensity-duration threshold for landslides in a humid-tropical environment, Puerto Rico, Geografiska Annaler: Series A, Physical Geography, 75, 13–23, 1993.

Lee, M. L., Ng, K. Y., Huang, Y. F., and Li, W. C.: Rainfall-induced landslides in Hulu Kelang area, Malaysia, Nat Hazards, 70, 353–375, https://doi.org/10.1007/s11069-013-0814-8, 2014.

Livneh, B. and Rajagopalan, B.: Development of a gridded meteorological dataset over Java island, Indonesia 1985–2014, Scientific data, 4, 1–10, 2017.

Macdonald, E., Merz, B., Guse, B., Wietzke, L., Ullrich, S., Kemter, M., Ahrens, B., and Vorogushyn, S.: Event and Catchment Controls of Heavy Tail Behavior of Floods, Water Resources Research, 58, e2021WR031260, https://doi.org/10.1029/2021WR031260, 2022.

Mandal, P. and Sarkar, S.: Estimation of rainfall threshold for the early warning of shallow landslides along National Highway-10 in Darjeeling Himalayas, Nat Hazards, 105, 2455–2480, https://doi.org/10.1007/s11069-020-04407-9, 2021.

Manoj J, A., Guntu, R. K., and Agarwal, A.: Spatiotemporal dependence of soil moisture and precipitation over India, Journal of Hydrology, 610, 127898, https://doi.org/10.1016/j.jhydrol.2022.127898, 2022.

Marques, R., Zêzere, J., Trigo, R., Gaspar, J., and Trigo, I.: Rainfall patterns and critical values associated with landslides in Povoação County (São Miguel Island, Azores): relationships with the North Atlantic Oscillation, Hydrological Processes: An International Journal, 22, 478–494, 2008.

Martha, T. R., Roy, P., Govindharaj, K. B., Kumar, K. V., Diwakar, P. G., and Dadhwal, V. K.: Landslides triggered by the June 2013 extreme rainfall event in parts of Uttarakhand state, India, Landslides, 12, 135–146, 2015.

Martha, T. R., Roy, P., Jain, N., Khanna, K., Mrinalni, K., Kumar, K. V., and Rao, P. V. N.: Geospatial landslide inventory of India—an insight into occurrence and exposure on a national scale, Landslides, 1–17, 2021.

Maturidi, A., Kasim, N., Taib, K. A., Azahar, W., and Tajuddin, H. A.: Empirically based rainfall threshold for landslides occurrence in Cameron highlands, Malaysia, Civil Eng Architect, 8, 1481–1490, 2020.

Merz, R. and Blöschl, G.: A process typology of regional floods, Water Resources Research, 39, https://doi.org/10.1029/2002WR001952, 2003.

Pasuto, A. and Silvano, S.: Rainfall as a trigger of shallow mass movements. A case study in the Dolomites, Italy, Environmental Geology, 35, 184–189, 1998.

Ratan, R. and Venugopal, V.: Wet and dry spell characteristics of global tropical rainfall, Water Resources Research, 49, 3830–3841, https://doi.org/10.1002/wrcr.20275, 2013.

Ross, T. J.: Fuzzy logic with engineering applications, John Wiley & Sons, 2005.

Sadri, S. and Burn, D. H.: A Fuzzy C-Means approach for regionalization using a bivariate homogeneity and discordancy approach, Journal of Hydrology, 401, 231–239, https://doi.org/10.1016/j.jhydrol.2011.02.027, 2011.

Schneider, T.: Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values, Journal of climate, 14, 853–871, 2001.

Schoups, G. and Vrugt, J. A.: A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, Water Resources Research, 46, https://doi.org/10.1029/2009WR008933, 2010.

Segoni, S., Piciullo, L., and Gariano, S. L.: A review of the recent literature on rainfall thresholds for landslide occurrence, Landslides, 15, 1483–1501, https://doi.org/10.1007/s10346-018-0966-4, 2018.

Sengupta, A., Gupta, S., and Anbarasu, K.: Rainfall thresholds for the initiation of landslide at Lanta Khola in north Sikkim, India, Natural hazards, 52, 31–42, 2010.

Soman, M. and Kumar, K. K.: some aspects of daily rainfall distribution over India during the south-west monsoon season, https://doi.org/10.1002/JOC.3370100307, 1990.

Stanley, T., Kirschbaum, D. B., Pascale, S., and Kapnick, S.: Extreme Precipitation in the Himalayan Landslide Hotspot, in: Satellite Precipitation Measurement: Volume 2, edited by: Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., and Turk, F. J., Springer International Publishing, Cham, 1087–1111, https://doi.org/10.1007/978-3-030-35798-6_31, 2020.

Teja, T. S., Dikshit, A., and Satyam, N.: Determination of Rainfall Thresholds for Landslide Prediction Using an Algorithm-Based Approach: Case Study in the Darjeeling Himalayas, India, Geosciences, 9, 302, https://doi.org/10.3390/geosciences9070302, 2019.

Toté, C., Patricio, D., Boogaard, H., Van der Wijngaart, R., Tarnavsky, E., and Funk, C.: Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique, Remote Sensing, 7, 1758–1776, 2015.

Tsidu, G. M.: High-Resolution Monthly Rainfall Database for Ethiopia: Homogenization, Reconstruction, and Gridding, Journal of Climate, 25, 8422–8443, https://doi.org/10.1175/JCLI-D-12-00027.1, 2012.

Vallet, A., Varron, D., Bertrand, C., and Mudry, J.: Hydrogeological threshold using support vector machines and effective rainfall applied to a deep seated unstable slope (Séchilienne, French Alps), in: Engineering Geology for Society and Territory-Volume 2, Springer, 2143–2146, 2015.

Wieczorek, G.: In central Santa Cruz Mountains, California, Debris flows/avalanches: process, recognition, and mitigation, 7, 93, 1987.

Wu, H., Adler, R. F., Hong, Y., Tian, Y., and Policelli, F.: Evaluation of global flood detection using satellite-based rainfall and a hydrologic model, Journal of Hydrometeorology, 13, 1268–1284, 2012.

Zink, M., Kumar, R., Cuntz, M., and Samaniego, L.: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty, Hydrology and Earth System Sciences, 21, 1769–1790, https://doi.org/10.5194/hess-21-1769-2017, 2017.

Þórðarson, A. F., Baum, A., García, M., Vicente-Serrano, S. M., and Stockmarr, A.: Gap-Filling of NDVI Satellite Data Using Tucker Decomposition: Exploiting Spatio-Temporal Patterns, Remote Sensing, 13, 4007, https://doi.org/10.3390/rs13194007, 2021.