Response Document

Referee 1 Comments and Responses

It was a pleasure to read the manuscript on "Role of hydrometeorological variables and catchment area to flood generation over the monsoonal climate region of India" by Aadhar et al.

The manuscript does an extensive investigation on the flood generating mechanisms in the Godavari River basin, focusing on the rainfall characteristics and identification of catchment area that predominantly contributes to flooding in different subbasins in the Godavari River. The authors rely on an event-based approach wherein independent flood events are identified using a POT approach, and the characteristics of rainfall and antecedent soil moisture conditions prior to the flood events are assessed. The authors identify the time window of rainfall duration which is most correlated with the flood events in each subbasin and classify the triggering rainfall event into one-day extreme, multiday extreme and moderate rainfall events. Using an EOF analysis and correlation they also identify the regions in the catchment with higher runoff coefficient or significant contribution to runoff generation. While the role of rainfall duration(Nanditha & Mishra, 2022), and catchment area on flood generation (Nanditha & Mishra, 2024) is well established in the Indian region, the second part of the study demarcating the flood-dominating catchment area in different subbasins is novel and the methodological aspect exhibits rigour and consistency. I only have a few suggestions and clarification to improve the readability and reproducibility of the manuscript.

We thank the reviewer for the valuable assessment of our work and the detailed, constructive suggestions. Below we addressed each comment individually.

Methodology

I would expect more detailing on the EOF approach employed. Please provide the relevant equations. Did the authors use accumulated rainfall and runoff for the most correlated window for each pixel for the analysis? I have some reservations on using accumulated runoff. Though runoff measured in mm (length scale) can be treated as an accumulated quantity, the way the VIC model represents it can be different. For instance, the runoff generated at timestep t may not entirely rout to the downstream cell at the end of the day due to the time lag in the routing process (the hill slope routing uses a unit hydrograph approach). So, the runoff generated for the timestep t+1 could have contribution from the previous day. I would suggest using mean runoff instead of accumulated runoff for each cell.

Similarly, peak discharge should be correlated with the mean runoff and not the accumulated runoff. I would suggest consistently use mean runoff throughout the manuscript.

Thank you for this constructive comment. We agree with the reviewer's concern regarding the use of accumulated runoff in the context of VIC routing. As highlighted, VIC's hillslope routing uses a unit hydrograph approach, meaning that runoff generated at timestep *t* may carry contributions from the previous timestep. We have therefore revised the analysis and now use mean runoff (mean daily surface runoff over the most correlated time window) consistently throughout the manuscript, for both the pixel-based EOF analysis and the correlation with peak discharge.

Moreover, we have added relevant equations of EOF analysis in the revised manuscript.

"Next, we evaluated the dominated catchment area in flood generation by examining the precipitation and runoff spatial variability (Figure 2). For that, we used the empirical orthogonal function (EOF) analysis and obtained the leading modes of spatial variability in precipitation and surface runoff (Mishra et al., 2012). EOF analysis identifies the dominant modes of spatio-temporal variability in a dataset by solving an eigenvalue problem derived from spatio-temporal covariance matrix (Lorenz, 1956; Preisendorfer, 1988; Wilks, 2006; Navarra and Simoncini, 2010; Joliffe, 1986; Hannachi, 2022). Here, we first estimated the dominant spatiotemporal modes of variability in the gridded precipitation (surface runoff) to assess its coupling with peak discharge. To do so we formed the spatial covariance matrix of N x N (where N is number of grid points) using following equation:

$$C = \frac{1}{t-1} P' P'^{T}$$

where P' is precipitation (runoff) anomaly; t is number of peak events; and T represents the transpose matrix.

Further, spatial covariance matrix (C) was decomposed as:

$$Ce_k = \lambda_k e_k$$

where e_k and λ_k are the k^{th} spatial EOF and its associated eigenvalue, respectively (Lorenz, 1956). The fraction of total variance explained by each mode is $\lambda_k / \Sigma\lambda_j$. We also obtained the Principal component (PC) time series by projecting the anomaly field onto each EOF:

$$a_k(t) = \sum_{i=1}^{N} P'_{i,t} \cdot e_{k,i}$$

Finally, coupling between each PC time series and the peak discharge anomaly was quantified using the Pearson correlation coefficient. In the EOF analysis, we used the accumulated precipitation and mean surface runoff at each pixel for the time windows found most correlated with peak discharge, according to the findings in step 1 above."

Figures

The caption of Figures 4-6 is disordered. The top panel in the current figure is explained as the last panel.

Thank you for carefully reviewing the figure captions. We apologise for this oversight. We have carefully revised the captions of Figures 4, 5, and 6, as well as a few supplementary figures, in the revised manuscript to ensure that the panel descriptions follow the correct order.

Other suggestions

Line 145. No need of Moreover here.

Thank you. Removed.

Line.295 - Is this surface runoff alone? Did you consider the 75th percentile of entire timeseries or seasonality is taken into account?

Thank you for your comment. Yes, this is surface runoff alone, as simulated by the VIC model (baseflow is not included). The 75th percentile threshold for soil moisture was estimated by considering all occurrences of the same calendar day across the study period (1967–2019) that coincided with high streamflow events. For instance, if a high streamflow event occurred on 23 July in a given year, all 23 July values from 1967 to 2019 were considered in computing the soil-moisture threshold.

Line 299: What is the thickness of soil moisture considered?

Thanks. We have mentioned the soil-layer thickness in the revised manuscript. It is 60 cm.

95th percentile is considered as the flood peak; this may include flows which are not exactly extreme. Could you please conduct a sensitivity analysis by considering other POT thresholds, say 99 percentile and see if it affects the findings of the manuscript.

Thank you for your concern and constructive suggestion. We have conducted the analysis considering the annual block maximum and different threshold values of high discharge time series and observed the similar results and conclusion. For instance, we estimated the occurrence of high-intensity precipitation (P99) and streamflow (Q99) events during period 1967-2019 (Figure R1) and found the similar pattern compared to 95th percentile threshold (Figure S3). The number of independent extreme events is lower under the 99th percentile threshold compared to 95th percentile threshold, but the key conclusions are qualitatively unchanged.

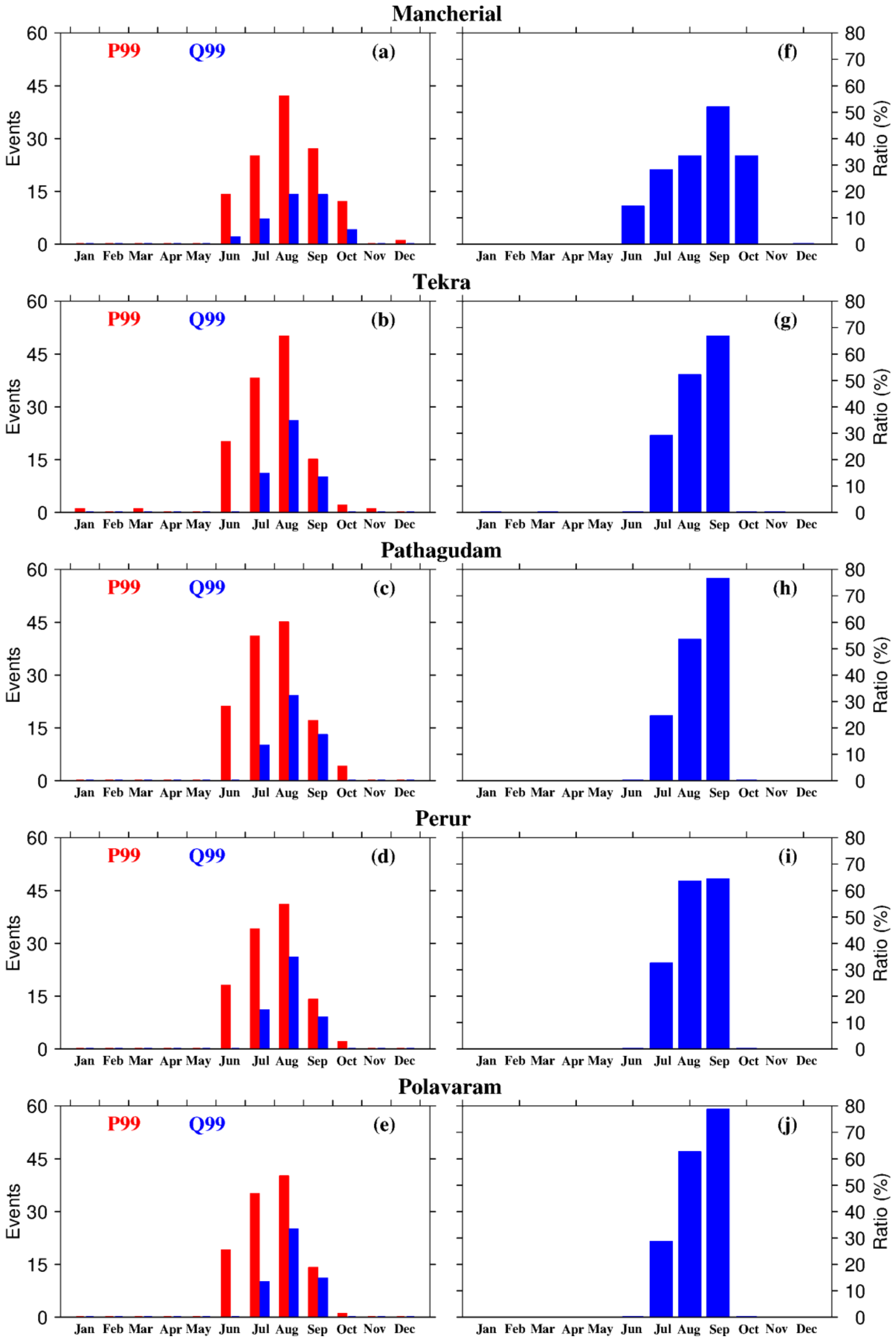

Figure R1. (a-e) Occurrence of high-intensity precipitation (P99) and streamflow (Q99) events (above 99th percentile) during 1967-2019 in each month at (a) Mancherial, (b) Tekra, (c) Pathagudam, (d) Perur, and (e) Polavaram stations. (f-j) Ratio of the number of extreme *streamflow and precipitation events in each month at Mancherial, Tekra, Pathagudam, Perur, and Polavaram stations, respectively.*

Additionally, we also evaluated the correlations between simulated streamflow peak and basin-averaged accumulated observed precipitation (prior to the flood peak) for different durations (days) and lags (days) and observed the same outcome. The high discharge events are associated with short-duration (2-days) rainfall.

Table R1. Correlations between simulated streamflow peak (above 99th percentile events) and basin-averaged accumulated observed precipitation (prior to the flood peak) for different durations (days) and lags (days). The analysis is for high streamflow events at the Mancherial station for the period 1969-2019. Highlighted number is the best correlation value.

Lag / Duration	0	1	2	3	4	5	6	7
1	0.049	0.194	0.724	0.324	0.211	-0.009	-0.101	0.049
2	0.173	0.656	0.766	0.339	0.136	-0.061	0.010	0.173
3	0.642	0.719	0.759	0.310	0.070	0.003	-0.033	0.642
4	0.702	0.714	0.734	0.254	0.102	-0.029	-0.008	0.702
5	0.694	0.692	0.692	0.267	0.066	-0.010	-0.014	0.694
6	0.669	0.655	0.692	0.229	0.070	-0.014	0.003	0.669
7	0.631	0.664	0.649	0.220	0.060	0.000	0.040	0.631
8	0.639	0.632	0.606	0.207	0.065	0.033	0.056	0.639
9	0.607	0.597	0.585	0.200	0.088	0.047	0.049	0.607
10	0.578	0.577	0.570	0.212	0.099	0.042	0.012	0.578
11	0.561	0.568	0.566	0.217	0.094	0.009	0.030	0.561
12	0.553	0.567	0.566	0.211	0.064	0.026	0.031	0.553
13	0.553	0.567	0.560	0.184	0.077	0.028	0.035	0.553
14	0.554	0.564	0.545	0.193	0.076	0.031	0.060	0.554
15	0.552	0.554	0.548	0.187	0.078	0.056	0.074	0.552
16	0.542	0.561	0.538	0.187	0.098	0.070	0.069	0.542
17	0.547	0.554	0.529	0.202	0.109	0.066	0.062	0.547

18	0.541	0.544	0.522	0.207	0.103	0.059	0.044	0.541
19	0.531	0.538	0.509	0.198	0.097	0.042	0.042	0.531
20	0.526	0.524	0.494	0.191	0.080	0.040	0.045	0.526

Line 377: is located instead of reside

Thanks. Done. We have updated it in the revised manuscript.

Lines 384-386: The correlation is estimated here between the peak discharge and accumulated rainfall in the relevant regions.

Thank you for your comment. Here, we estimated the correlation (as mentioned in the manuscript) between the first principal component (PC-1) of the EOF analysis of spatially distributed accumulated precipitation (or mean runoff) over the sub-basin and the observed peak discharge at the respective gauging station.

Lines 405-409: The duration mentioned here is different for different subbasins right? If the rainfall occurs for multiple day duration on dry soil vs wet soil it is expected to have difference in runoff coefficient. Are you here comparing different duration scenarios for the same subbasins, if so, multiday duration can increase runoff coefficient compared to single day duration rainfall for dry conditions. I feel there is some lack of clarity here.

Thank you for your comment. No, we have used three durations (2, 7, and 8 days) in Figure 6 (identified from our temporal analysis across all sub-basins; Table 2). We compute the runoff coefficient (mean surface runoff / mean precipitation) over each of these three fixed durations consistently across all sub-basins, and compare them under both dry and wet antecedent soil moisture conditions.

For multiday events on dry soil, a longer accumulation window should have a higher runoff coefficient than a single day event because prolonged rainfall progressively saturates the soil and increases runoff generation. However, here, we showed the variation in runoff coefficient across durations under dry and wet conditions.

Line 432: Initial hydrologic conditions instead of pre

Done.

Lines 474-476: Please explain this is not clear to me. In this manuscript, the authors have not considered the spatial pattern of soil moisture wetness as far as my understanding. the soil moisture pattern will also be related to the rainfall pattern in a catchment, and I do not understand how this conclusion of "soil moisture pattern may not affect the pre-event hydrometeorological patterns" is arrived at. What exactly the authors imply by hydrometeorological patterns?

Thank you for your comment. We have removed these lines from the revised manuscript.

Line 481: Did you mean forest to crop land conversion? If so, please mention it explicitly.

Thanks. We have updated it in the revised manuscript.

A possible extension is to explore if the identified flood-dominating area is also associated with the spatial rainfall pattern anomalies. That is, to explore if the region is receiving higher intensity rainfall than other parts of the catchment which leads to infiltration excess runoff generation. I would like to see such an analysis if it won't take much time and effort to corroborate the conclusion of the manuscript that the catchment hydrologic partitioning is the cause for the increase runoff contribution and to rule out that it is not related to the spatial heterogeneity in rainfall distribution.

Thank you for your comment and suggestion. We have already estimated the dominant rainfall pattern over the Godavari River Basin using EOF analysis to identify the high rainfall occurrence regions. Please check Figure 4, the leading mode of variability in the accumulated precipitation during the high streamflow events for the period 1967-2019.

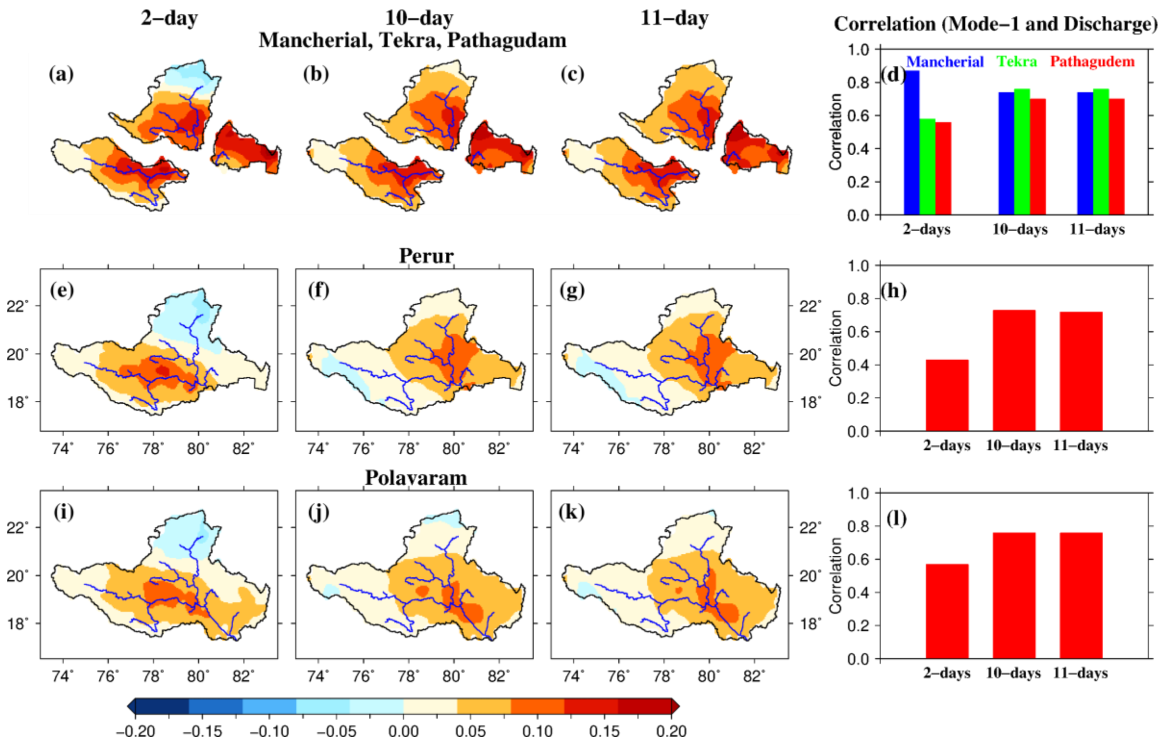

Figure 4: *The leading mode of variability in the accumulated precipitation during the high streamflow events for the period 1967-2019 using the method of empirical orthogonal function (EOF). (a-c) The first leading EOF mode of (a) 2-days, (b) 10-days, and (c) 11-days accumulated precipitation for Mancherial (left), Tekra (middle), and Pathagudam (right) sub-basins. (d) Correlation between first leading mode and flood peak at Mancherial (left), Tekra (middle), and Pathagudam (right) stations. (e-h) same as (a-d) but for Perur sub-basin. (i-l) same as (a-d) but for Polavaram sub-basin.*

References

Nanditha, J. S., & Mishra, V. (2022). Multiday Precipitation Is a Prominent Driver of Floods in Indian River Basins. *Water Resources Research*, *58*(7), e2022WR032723. https://doi.org/10.1029/2022WR032723

Nanditha, J. S., & Mishra, V. (2024). Projected increase in widespread riverine floods in India under a warming climate. *Journal of Hydrology*, *630*, 130734. https://doi.org/10.1016/J.JHYDROL.2024.130734

Thanks. We have cited the above relevant papers in the revised manuscript.