



Using a Bayesian joint probability approach to improve the skill of medium-range forecasts of the Indian summer monsoon rainfall

Nibedita Samal^a, R. Ashwin^a, Akshay Singhal^a, Sanjeev Kumar Jha^{a,*}, David E. Robertson^b

^a Indian Institute of Science Education and Research, Bhopal, Madhya Pradesh 462066, India

^b Commonwealth Scientific and Industrial Research Organization, Clayton South Victoria 3169, Australia

ARTICLE INFO

Keywords:

Bayesian Post-processing
Precipitation forecasts
Indian summer-monsoon
Ensembles
NCMRWF
Sub-basin scale

ABSTRACT

Study region: Ganga, Mahanadi, Godavari, Narmada, and Tapti River basins of India.

Study focus: The manuscript focuses on improving skills of the Indian summer-monsoon precipitation forecasts obtained from National Center for Medium-Range Weather Forecasting (NCMRWF) at both sub-basin and gridded scale. A well-established Bayesian Joint Probability (BJP) based statistical post-processing approach, operational in Australia, is used for the first time in India throughout diverse geographical extent. The work evaluates how the post-processor can be used in a summer-monsoon dominated region like India. The study informs whether annual or seasonal precipitation forecasts should be used as the length of data will play crucial role in both the cases. The spread-skill of the ensembles obtained from BJP approach and the NCMRWF is explored.

New hydrological insights for the region: Introduction of the BJP-based post-processing approach in India could help the forecast community to implement more robust approach in improving the skills of the forecasts. Our results show that instead of using the data of whole year, only monsoonal precipitation forecasts are adequate to setup the BJP approach. The calibrated forecasts obtained using three years of hindcast and observations data at grids and at the centroid of 177 sub-basins are found to be more skillful. The calibrated forecasts can discriminate between extreme and low precipitation events, and have appropriate ensemble spread to capture precipitation peaks. This study presents a guideline for water managers and forecasters to apply BJP approach to improve the forecasts.

1. Introduction

Rainfall exhibits high variability in magnitude, frequency and distribution across space and time (Singhal et al., 2022a; Singhal and Jha, 2021a). India is largely dependent on precipitation for its economic growth and sustenance. Hence, for a country like India, the availability of precipitation forecasts is vital. Each year, rainfall events of varying intensities cause floods in large part of the country during the monsoon months of June, July, August and September (JJAS), leading to numerous casualties, displacement of people and loss of property (Ray et al., 2021; Roxy et al., 2017). Reliable precipitation forecasts can help in issuing floods related warnings, effective management of water resources and consequently, inform better decision making (Singhal et al., 2022b).

* Corresponding author.

E-mail address: sanjeevj@iiserb.ac.in (S.K. Jha).

<https://doi.org/10.1016/j.ejrh.2022.101284>

Received 17 May 2022; Received in revised form 21 November 2022; Accepted 29 November 2022

Available online 5 December 2022

2214-5818/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The Quantitative Precipitation Forecasts (QPFs), which are obtained from Numerical Weather Prediction (NWP) models, are primary estimations of a possible precipitation event in a region a few days or weeks in advance (Ahmed et al., 2014; Froude, 2010; Tian et al., 2019). The deterministic QPFs obtained from the NWP models are single-valued forecasts which represent a single best possible state of future weather. Deterministic QPFs often contain errors since single-valued forecasts are unable to capture the uncertainty involved with weather patterns (Jha et al., 2018). On the other hand, probabilistic forecasts (with several forecast members) have the potential to represent multiple possible states of weather and thus, capture the weather uncertainty. However, the disadvantage of producing ensemble forecasts using NWP models is that they are computationally very expensive when required for large areas and at long lead times, and also, the spread of ensemble is generally too narrow (Ramos et al., 2013; Schepen et al., 2018; Shrestha et al., 2015). Moreover, the raw probabilistic forecasts are unsuitable to be used since they contain inherent systematic bias due to uncertainty in both the assimilation and physical parameterization processes of NWP models (Buizza, 2018; Maraun et al., 2015; Saminathan et al., 2021). The NWP models are also influenced by the formulation, horizontal-vertical resolutions, and the initial conditions of the model which influences the quality of the QPFs (Bowler et al., 2008; Nipen and Stull, 2011). Therefore, it is essential to post-process the raw forecasts with statistical models before using them for operational purposes.

Statistical post-processing approaches are computationally inexpensive, quick and easy to apply, and capture the weather uncertainty quite well (Shrestha et al., 2015; Verkade et al., 2013; Yagli et al., 2020). However, statistical post-processing of rainfall forecasts still remains a challenging task. Several statistical post-processing approaches have been developed and implemented in the past to improve the skill of the raw forecast, such as the analog method (Hamill and Whitaker, 2006), linear and non-linear regression (Clark and Hay, 2004; Verkade et al., 2013), logistic regression (Hamill et al., 2006; Medina et al., 2019; Saminathan et al., 2021), extended logistic regression (Roulin and Vannitsem, 2012), Bayesian theory and fuzzy probability (Cai et al., 2019), ensemble model output statistics (EMOS) approach (Gneiting et al., 2005; Li et al., 2019), and copula-based post-processing techniques (Li et al., 2021). However, problems arising due to the irregular nature of precipitation, large values of zero rainfall, reduction of forecast skill with the increase in lead times and seasonal variations in rainfall patterns need to be addressed.

An approach based on joint probability of forecasts and long historical observations has been developed, which has gained credible recognition in the field of statistical post-processing. The basic aim of this approach is to produce calibrated forecasts by establishing a joint probability distribution between forecasts (as predictors) and the observation data (as predictands) (Robertson et al., 2013; Wang et al., 2009). Previous studies suggest that the joint probability based distribution approaches offer advantage over the other post-processing approaches (Li et al., 2017). However, to the best of our knowledge, the potential of this approach to obtain calibrated forecasts is yet to be explored over India. More recent analysis suggests that neglecting rainfall seasonality in post-processing methods can produce forecasts that contain seasonal biases and display poorer skill (Wang et al., 2019). In India, the strong seasonality of precipitation due to the monsoon adds an additional challenge to the post-processing of rainfall. Many previous applications of post-processing methods, where short records of NWP is used, neglect the effects of rainfall seasonality (Shrestha et al., 2016). However, in case where the forecast data is available only for a small period, it becomes a very interesting and challenging exercise to check if a post-processor can still improve the forecasts. The Bayesian modelling based post-processing approaches have been demonstrated to successfully generate skillful forecasts with shorter records of data. For instance, (Robertson et al., 2013) used the rainfall post-processing approach over Murray-Darling basin, Australia to obtain calibrated forecasts using 20 months of data. Shrestha et al. (2015) used 21 months of data to produce calibrated forecasts for different catchments of Australia. Jha et al. (2018) used 3 years of data for obtaining calibrated forecasts for a Canadian basin. Hence, the RPP model used in this manuscript has already been proven to be effective in producing skillful forecasts by obtaining parameters even from shorter data records. Moreover, the results presented in the manuscript incorporated sampling variability while computing the verification measures. This is done by applying bootstrapping approach with 5000 times sampling of the data records, while calculating the uncertainties. This gives a confidence about the robustness of the calibrated forecasts and justifies its ability to calibrate forecasts from operational agencies.

This study sets up the Bayesian Joint Probability (BJP) model-based rainfall post-processing approach developed by Robertson et al., (2013) over five major river basins of India. The main aim of the study is to post-process the deterministic raw QPFs obtained from the National Centre for Medium-Range Weather Forecasting (NCMRWF) to produce ensembles calibrated QPFs for five daily lead-times at sub-basin scale. The study has also been extended to establish the RPP model using the gridded NCMRWF forecasts. Applying RPP at the grid-scale helps us to understand the influence of spatial resolution, data length and lead times on the performance of the model.

NCMRWF is the national weather forecast center of India which performs medium-range global assimilation and rainfall forecasts in real-time at the spatial scale of $0.18^\circ \times 0.12^\circ$. The calibrated QPFs are subsequently verified with the satellite-derived Integrated Multi-satellite Retrievals for Global Precipitation Measurement (IMERG) precipitation product (Das et al., 2022; Kumar Singh et al., 2019; Li et al., 2018; Prakash et al., 2018). Prior to post-processing, the historical IMERG rainfall are interpolated to same spatial scale as the NCMRWF forecasts which is $0.18^\circ \times 0.12^\circ$. For the application of RPP on sub-basin level, the forecasts and observation data are interpolated to the centroid of each of the 177 sub-basins following the sub-basin average method. Two experiments are performed at each sub-basin and gridded level to evaluate the performance of the approach in incorporating the seasonal behavior of rainfall in India. In the first experiment, we apply the post-processor to calibrate the annual precipitation forecasts containing both the dry and wet season (dry season mostly contains zero rainfall), while in the second experiment, we focus on calibrating the forecasts of the wet season (JJAS) alone. Please note that although the input data provided to the model is annual, we extract and present the results for the monsoon season only. The study is important as it provides a reliable avenue to obtain daily calibrated QPFs both at the sub-basin and the gridded scale which can help in streamflow forecasting, flood warning, agro-met advisories and other related decision-making.

The specific objectives of the study are: (1) to set up a BJP model-based post-processing approach over the five river basins of India to produce sub-basin averaged calibrated QPFs at 5-day lead times; (2) to assess the performance of the post-processing approach in

producing the annual calibrated QPFs (ACQ) using the corresponding raw deterministic forecast; (3) to evaluate the performance of post-processing in producing seasonal calibrated QPFs (SCQ) using the corresponding raw deterministic forecast; (4) to investigate the quality of the forecasts at varying rainfall intensities for both ACQ and SCQ; (5) to investigate the potential of calibrated QPFs to reliably discriminate the precipitation events (and non-events) for both ACQ and SCQ; (6) to extend the study to gridded level and evaluate the gridded calibrated QPFs based on accuracy and skill of the forecasts; (7) to examine the accuracy and spread of the post-processed ensemble forecasts by comparing them to the available perturbed ensemble forecasts from NCMRWF for both ACQ and SCQ.

The remainder of the paper is structured as follows. Section 2 describes the details related to the study area and datasets. Section 3 describes the methodology implemented to post-process the precipitation forecasts. Section 4 presents the results of post-processing along with their evaluation. Section 5 discusses the results, and Section 6 deals with conclusions and the future scope of this study.

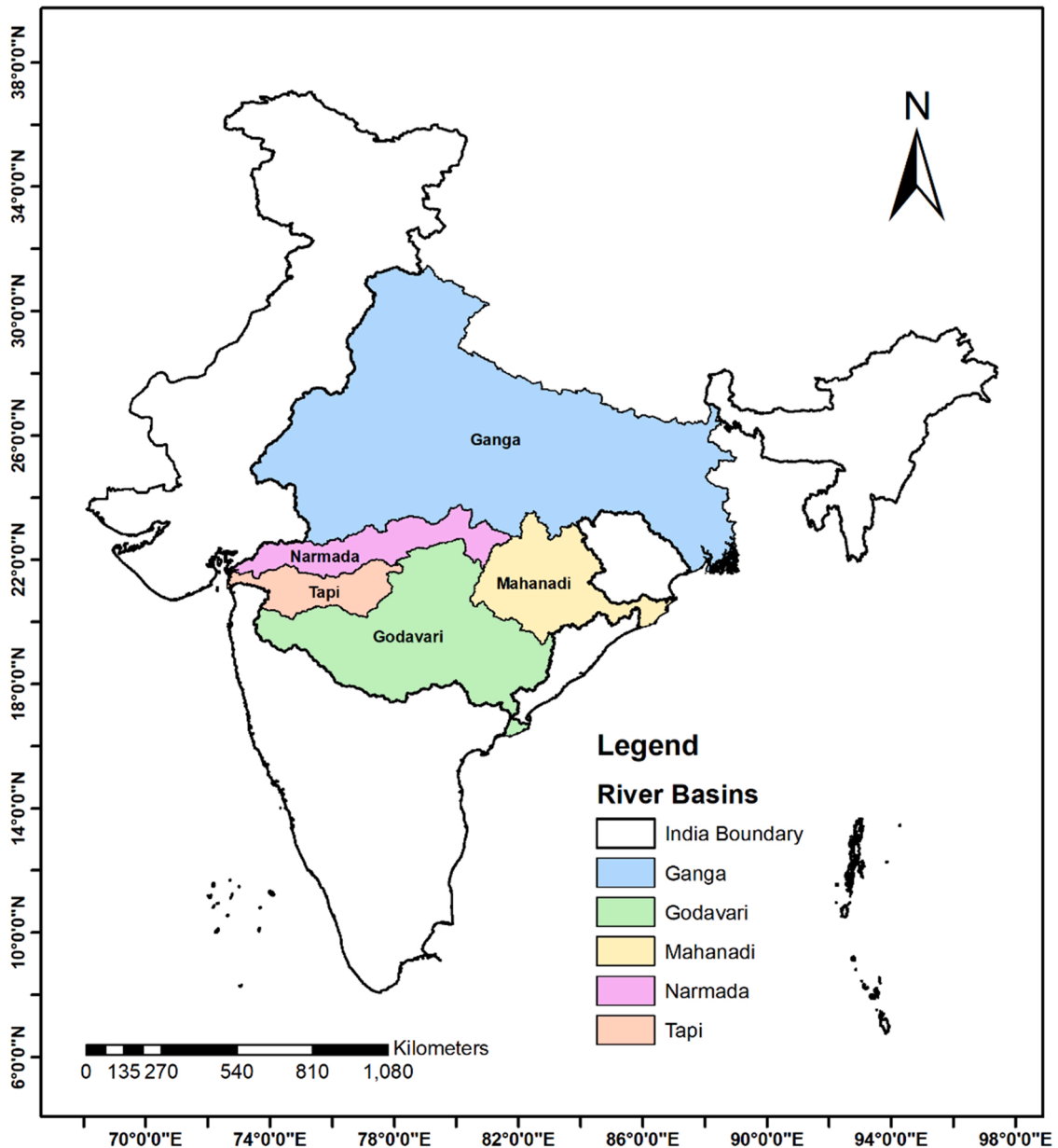


Fig. 1. Location of the five major River basins of India selected for this study.

2. Study area and datasets

2.1. Study area

The study area includes five major river basins of India which are Ganga, Godavari, Mahanadi, Narmada, and Tapti as depicted in Fig. 1. These rivers cover approximately 60 % of the total geographical area of India and contribute to the major water availability in the country. The Ganga River originates in the high elevation region of Uttarakhand and confluences into the Bay of Bengal in the east with elevation reaching up to 7000 m. The river makes the northern part of India a highly fertile zone, due to which agriculture becomes the predominant source of livelihood for most of the population. Further, the Narmada, Tapti and Mahanadi basins are located in the central part of the country having lower elevation ranges while the Godavari River basin is located in peninsular India characterized by some areas of higher elevation. The basins receive most of the precipitation during the Indian summer monsoon season (Singh and Jha, 2021). Previous studies suggest that frequencies of higher magnitude precipitation in the region are increasing (Ghosh et al., 2012; Guhathakurta and Revadekar, 2017) which have led to frequent floods (Lakshmi et al., 2019). The distribution of the 177 sub-basins are as follows: Ganga comprises of the maximum sub-basins (88), followed by Godavari (50), Mahanadi (20), Tapti (10) and Narmada (9). More details about the various river basins are presented in Table 1.

2.2. Datasets

The deterministic precipitation forecast obtained from NCMRWF and satellite-derived data from IMERG are used in this study. The QPF from NCMRWF is used at the daily temporal resolution, available at lead times of 1–5 days, while the spatial resolution of the data is 0.18° x 0.12°. Moreover, IMERG-V6 dataset is currently available at a spatial resolution of 0.1° x 0.1° and a half-hour temporal resolution. Both the observation and forecast datasets are used over a time period of July 2018 to September 2021. More details of the datasets are provided in Table 2. The half-hour period 3IMERGHH data is aggregated to obtain the daily data and upscaled to the same resolution as NCMRWF i.e., 0.18° x 0.12°. Previous studies have found that upscaling of rainfall datasets do not significantly influence their characteristics or lead to possible loss of information due to the change in resolution (Imhoff et al., 2020; Nicótina et al., 2008). The final data set for the observed rainfall contains 101 (points in x-direction) × 142 (points in y-direction) with a total of 14,342 grids over the domain. Further, both the datasets are interpolated at the centroid of each sub-basin following the weighted area-average method (Singh et al., 2021; Singhal and Jha, 2021b). Consequently, a time series of rainfall at each sub-basin is obtained after interpolation. We also use the available perturbed ensemble forecasts from NCMRWF (which comprises 11 ensemble members) for comparison with the post-processing-generated forecast.

3. Methods

In this section, we first provide a brief description of the post-processing approach, followed by the utilization of the Schaake Shuffle technique and details about the metrics to verify the calibrated QPFs. Please note that each of the steps involved here is applied to obtain both the ACQ and SCQ in this study.

3.1. The post-processing approaches

The post-processing approach, developed by Robertson et al., (2013), is a BJP model-based statistical rainfall post-processor that establishes a joint probability distribution between the raw QPFs and the corresponding observation data to produce calibrated QPFs. The working principle of the post-processing approach is based on three steps: (1) transform the non-normal distribution of the precipitation data (forecast and observation) into a bivariate normal distribution using the log-sinh transformation (Eqs. 1 and 2); (2) infer the relevant parameters to model the joint probability of raw QPFs and observations which maximizes the likelihood of posterior parameter distribution and (3) back-transform the forecast values to their original space.

$$x_{fst}^{\hat{}} = \frac{1}{\beta_{fst}} (\sinh(\alpha_{fst} + \beta_{fst} \cdot x)) \tag{1}$$

$$y_{obs}^{\hat{}} = \frac{1}{\beta_{obs}} (\sinh(\alpha_{obs} + \beta_{obs} \cdot y)) \tag{2}$$

Table 1
Details of the basins and its sub-basins present in the study area.

Basin	No. of Sub-Basins	Total Area Coverage (km ²)	Sub-basin with minimum Area (km ²)	Sub-basin with Maximum Area (km ²)
Ganga	88	1086,000	20.4	68,436.5
Mahanadi	20	1,41,589	238.1	23,014.1
Godavari	50	3,12,812	69.5	51,574.2
Narmada	9	97,410	4214.2	33,757.4
Tapti	10	65,145	1231.7	18,631.1

Table 2
Description of precipitation datasets used in the study.

Data Type	Data Source	Ensembles	Time Period		Lead Time	Temporal resolution	Spatial resolution	Forecast hours
			ACQ	SCQ				
Forecast	NCMRWF	Control forecast	July 2018-Sep 2021	2019–2021 (JJAS)	5 days	daily	0.18° x 0.12°	UTC 00:00:00, 12:00:00
Forecast	NCMRWF	Perturbed forecast (11 members)	2019–2021 (JJAS)		5 days	daily	0.18° x 0.12°	UTC 00:00:00, 12:00:00
Observation*	IMERG	GPM satellite retrievals	July 2018-Sep 2021		–	daily	0.18° x 0.12°	–

*The original spatial resolution of IMERG (0.1° x 0.1°) is upscaled to 0.18° x 0.12° for this study.

where (x_{fcst}, y_{obs}) represent the forecast and observation in transformed space. The parameters involved in this transformation are represented by $(\alpha_{fcst}, \beta_{fcst})$ and $(\alpha_{obs}, \beta_{obs})$ for the forecast and observation, respectively.

After transformation, (x_{fcst}, y_{obs}) is assumed to follow a bivariate normal distribution given as:

$$\begin{bmatrix} x_{fcst} \\ y_{obs} \end{bmatrix} \sim N(\mu, \Sigma)$$

$$p(x_{fcst}, y_{obs}) \sim N(\mu, \Sigma)$$

where Σ and μ are defined as:

$$\mu = \begin{bmatrix} \mu_{x_{fcst}} \\ \mu_{y_{obs}} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_{x_{fcst}}^2 & \rho_{x_{fcst}, y_{obs}} \sigma_{x_{fcst}} \sigma_{y_{obs}} \\ \rho_{x_{fcst}, y_{obs}} \sigma_{x_{fcst}} \sigma_{y_{obs}} & \sigma_{y_{obs}}^2 \end{bmatrix}$$

$(\mu_{x_{fcst}}, \sigma_{x_{fcst}}^2)$ and $(\mu_{y_{obs}}, \sigma_{y_{obs}}^2)$ represent the mean and standard deviation of the observation and forecast, respectively, and $\rho_{x_{fcst}, y_{obs}}$ is the correlation coefficient between x_{fcst} and y_{obs} . The set of nine parameters used in modelling the joint distribution is the following:

$$S = (\alpha_{fcst}, \beta_{fcst}, \mu_{x_{fcst}}, \sigma_{x_{fcst}}^2, \alpha_{obs}, \beta_{obs}, \mu_{y_{obs}}, \sigma_{y_{obs}}^2, \rho_{x_{fcst}, y_{obs}})$$

The set of nine parameters is estimated using the Shuffled Complex Evolution Algorithm (Duan et al., 1994) based on maximizing the likelihood of the posterior distribution. Once the parameters are estimated, the bivariate normal distribution conditioned on the raw QPFs is used to estimate the forecasts. The forecast ensembles generation is done by randomly sampling from the conditional distribution. The back-transformation of the forecast values is done by using the inverse of x_{fcst} and y_{obs} (inverse of Eqs. 1 and 2).

3.2. Cross-validation of calibrated forecasts

In order to evaluate the performance of the post-processor, we adopt a leave-one-month-out cross-validation procedure. In this procedure, the data, except for the left-out month, is used to estimate the parameters of the BJP model. Using the estimated parameters, the precipitation for all the days of the “left-out” month is estimated and verified with the corresponding observation data. The procedure is repeated until all the available raw QPFs are calibrated.

3.3. Schaake shuffle

The Schaake Shuffle is a technique that is applied to regain spatial co-variability (between neighboring locations) and temporal persistence (in predicted precipitation) that are lost during the post-processing of data from the NWP models (Schaake et al., 2007). The steps involved in Schaake Shuffle are as follows: - (i) a forecast date is chosen, and a sample of observation data is selected. The observation sample and the ensemble have the same size selected from an observation period in the past; (ii) the observation sample chosen above in Step 1 is ranked (in ascending order). The ranking procedure is repeated similarly for the ensemble forecast data; (iii) a date is chosen from the sample dataset of the observations. Subsequently, the ranks of the observation sample are identified; (iv) from the ensemble set of the forecasts, a forecast having the same rank as that of the selected observations is chosen, and (v) the steps 3–4 are repeated for all ensemble members.

3.4. Verification measures

In this study, we use various measures to verify the efficacy of the post-processing method. The measures used are Bias Percentage

(Bias %), Continuous Rank Probability Score (CRPS), Root Mean Square Error (RMSE), Nash-Sutcliffe efficiency (NSE), Relative Operating Characteristic (ROC) and the spread-skill analysis.

3.4.1. Bias percentage (Bias %)

Bias calculates the difference between the mean of the ensemble precipitation forecasts and the mean of the observations. Bias (%) is defined as the percentage deviation of the mean of calibrated forecasts from the observations, as shown in Eq. (3).

$$BIAS\% = \frac{\sum_1^t F_{x_{fcst}}(t) - \sum_1^t F_{y_{obs}}(t)}{\sum_1^t F_{y_{obs}}(t)} * 100 \quad (3)$$

Here x_{fcst} could either be raw $x_{rawfcst}$ or post-processed forecasts x_{fcst} , and y_{obs} represents observation.

3.4.2. Continuous rank probability score (CRPS)

The CRPS compares the cumulative distribution function (CDF) of the forecast ($F_{x_{fcst}}$) with the corresponding CDF of the observation ($F_{y_{obs}}$).

$$CRPS = \int_{-\infty}^{\infty} (F_{x_{fcst}}(t) - F_{y_{obs}}(t))^2 dt \quad (4)$$

CRPS reduces to mean absolute error for deterministic QPFs and allows efficient comparison between ensemble and deterministic forecasts. Smaller values of CRPS are favorable (closer to zero), meaning that the forecasts are accurate and vice versa.

3.4.3. Root mean square error (RMSE)

RMSE gives the standard deviation of the model prediction error and a smaller RMSE value indicates better performance of the presented model.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_{x_{fcst}}(t) - F_{y_{obs}}(t))^2} \quad (5)$$

Here x_{fcst} could either be raw $x_{rawfcst}$ or post-processed forecasts x_{fcst} .

3.4.4. Nash-Sutcliffe efficiency (NSE)

The NSE is a normalized statistics that determines the relative magnitude of the residual variance (noise) compared to the measured variance.

$$NSE = 1 - \left[\frac{\sum_1^t (F_{y_{obs}}(t) - F_{x_{fcst}}(t))}{(F_{x_{fcst}}(t) - F_{y_{obs}}(t))} \right] \quad (6)$$

\hat{y}_{obs} is the true mean of observation. NSE ranges from -Inf to 1 and values equal to 1 indicates a perfect match of the forecast to the observed data. $NSE = 0$, indicates that the model predictions are as accurate as the mean of the observed data, $-Inf < NSE < 0$, indicates that the observed mean is better predictor than the model.

3.4.5. Relative operating characteristic (ROC)

ROC curves are used to evaluate the ability of a forecasting method to distinguish between events and non-events, and plot the hit rate against the false alarm rate. The hit rate refers to the probability of a forecast to detect events that exceed a predefined precipitation threshold, while the false alarm rate refers to the probability of a forecast predicting erroneous events (Robbins and Titley, 2018). For a skillful calibrated forecast, hit rates are expected to exceed the false-alarm rates and the ROC curves are closer to the top-left corner of the plot. On the other hand, the detecting ability of the forecast is considered low when the ROC curves are closer to the diagonal. The Area Under Curve (AUC) is calculated for each ROC curve to distinguish the measure of separability. The AUC is expected to be higher than 0.6 for better distinguishing separate events.

3.4.6. Spread-skill analysis

The aim of the spread-skill analysis is to compare the spread in forecast ensembles with the forecast error (Nester et al., 2012). The forecast error is defined as the difference between the observations and the ensemble mean. In this study, we perform the spread-skill analysis between the available 11 forecast members of the perturbed NCMRWF and the 300 generated members of ACQ and SCQ, arranging them in increasing order, followed by their grouping into ten bins. Consequently, the average ensemble spread and forecast error from each bin is plotted.

3.5. Implementation of RPP

In this study, we apply RPP to obtain the calibrated QPFs over the Indian River Basins at two spatial scales, i.e., at the centroids of each sub-basin and at each individual grids. To this end, we analyse the performance of the RPP to post-process monsoonal precipitation using two experiments at each of the two spatial scales. First, using the annual data producing ACQ and second with seasonal

data producing SCQ for both the gridded and sub-basin averaged precipitation. The annual run used sub-basin averaged forecast and observation precipitation at each sub-basins for a period of 39 months from July 2018 to September 2021, whereas for seasonal run of RPP, monsoon season (June, July, August, and September) data of year 2019–2021 is used. For validation, a leave-one-month-out cross-validation strategy is applied as the data record used is of smaller time period. Specifically, one month data is left for validation while data from other months are used for parameter estimation and optimization. Once the parameters are obtained, the left one-month data is used to produce calibrated forecasts with 300 ensemble members. This is repeated for the five lead times of the NCMRWF forecast data at each sub-basins of the study area. It is noteworthy that the annual experiment uses annual data producing ACQ, however for better comparison with SCQ we extract and present the results for monsoon season only. The RPP calibrated QPFs at the sub-basin scale are evaluated using the measures of bias percentage and CRPS. The discriminating ability of the calibrated QPFs are investigated in terms of hit rate and false alarm rate using a precipitation threshold of 5 mm. Finally, the spread skill of the calibrated ensemble forecasts is analyzed with respect to the available perturbed ensemble forecast from NCMRWF. Each of the mentioned evaluations is undertaken for both ACQ and SCQ at the lead times of 1-day, 3-day and 5 days. The RPP calibrated QPFs at the grid scale are evaluated using bias percentage and CRPS.

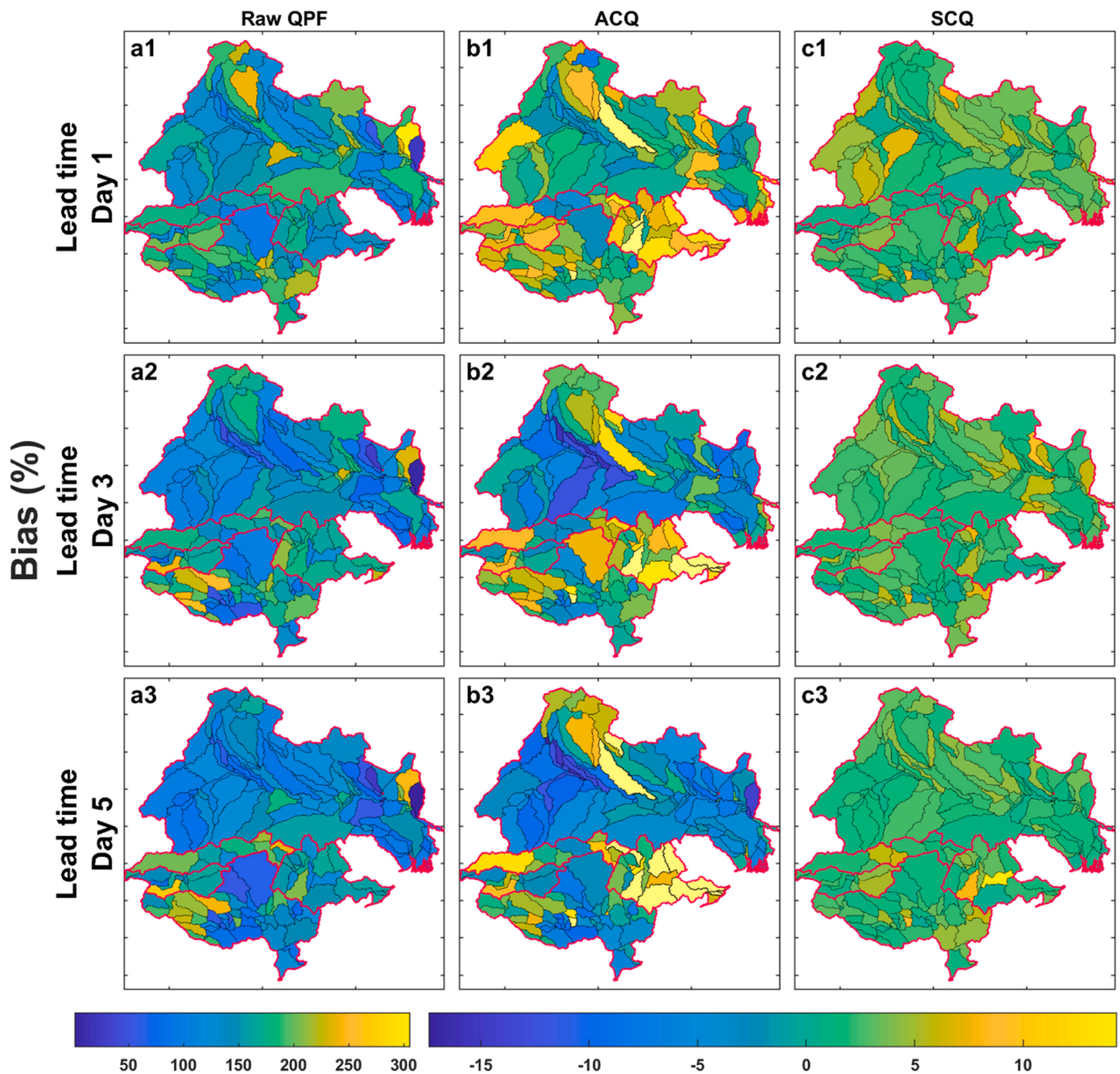


Fig. 2. Spatial plots of the sub-basin averaged bias percentage of raw QPFs (a1 to a3), Post-processed QPFs through SBA_RPP (b1 to c3) of daily precipitation including lead time day-1, day-3, and day-5 for the five river basins of Ganga, Mahanadi, Godavari, Narmada, and Tapi.

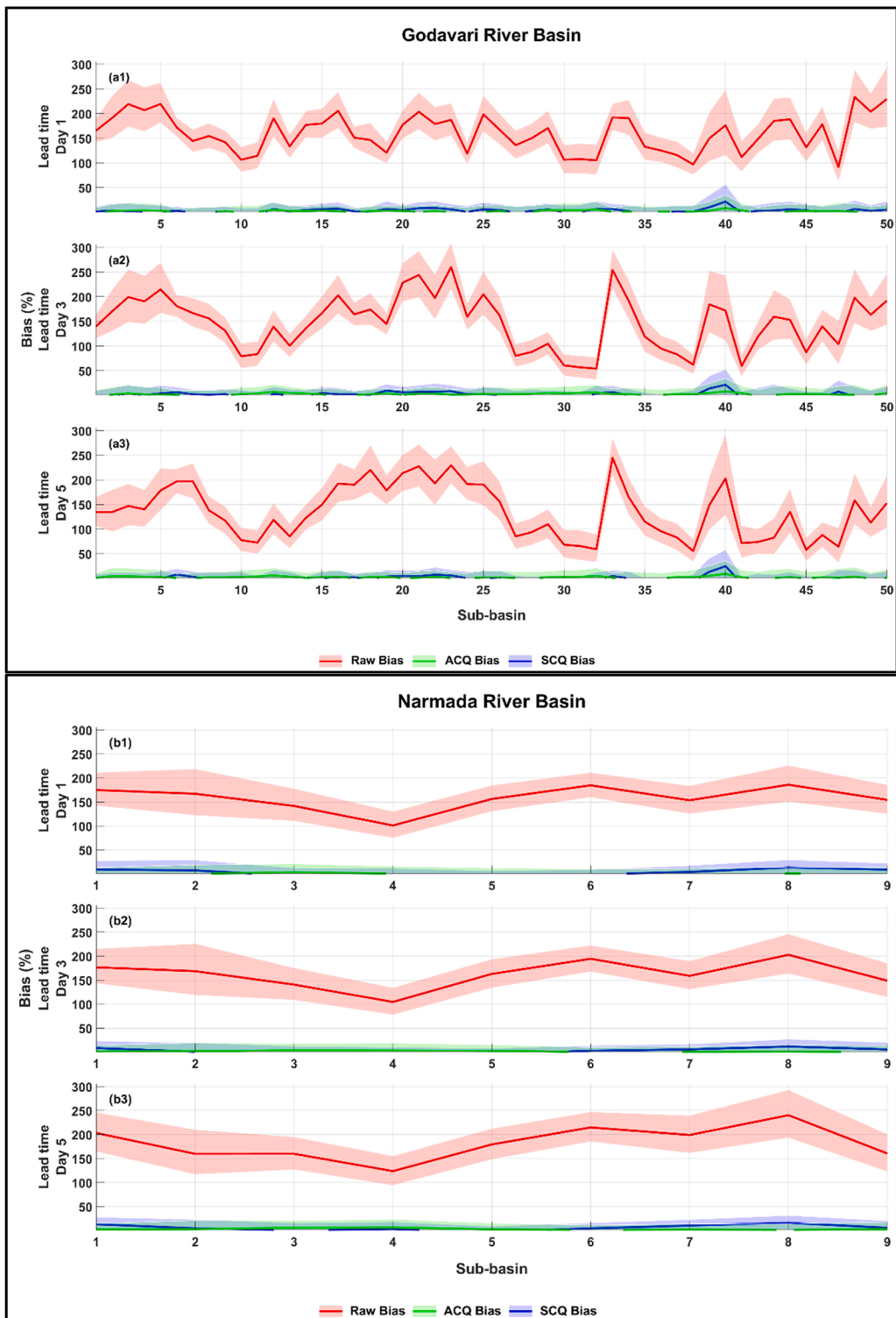


Fig. 3. Sub-basin averaged bias (%) with 5 % and 95 % confidence interval (shaded region) for raw QPFs (red), Post-processed QPFs for ACQ (green), and SCQ (violet) through gridded RPP to SBA and SBA_RPP for lead times day-1, day-3, and day-5 for the river basins Godavari and Narmada.

4. Results

4.1. Evaluation of the calibrated QPFs at the sub-basin scale

Here, the performance of the calibrated forecasts in each sub-basin (ACQ and SCQ) is evaluated using the measures of bias percent and CRPS at lead times of 1-day, 3-day and 5 days. Please note that the evaluation presented is for the monsoon season only in both cases of ACQ and SCQ.

4.1.1. Evaluation using percentage bias

Fig. 2 shows the spatial plot of sub-basin averaged bias percent of the raw QPF (a1-a3), ACQ (b1-b3) and SCQ (c1-c3) at lead times of 1-day, 3-day and 5-days for all the 177 sub-basins. We observe high spatial variability in the bias% of the raw deterministic forecast as depicted in Fig. 2 (a1-a3). The range of bias is also large with it ranging between -95 to 300 . Moreover, we find greater variability among the sub-basins of the Ganga River, western Tapti, and western Godavari basins as opposed to the sub-basins of other areas in all three lead times. Large bias over these sub-basins could be attributed to the presence of mountainous regions, orographic influence, complex topography and frequent events of extreme precipitation. Further, results suggest that the ACQ obtained after post-processing

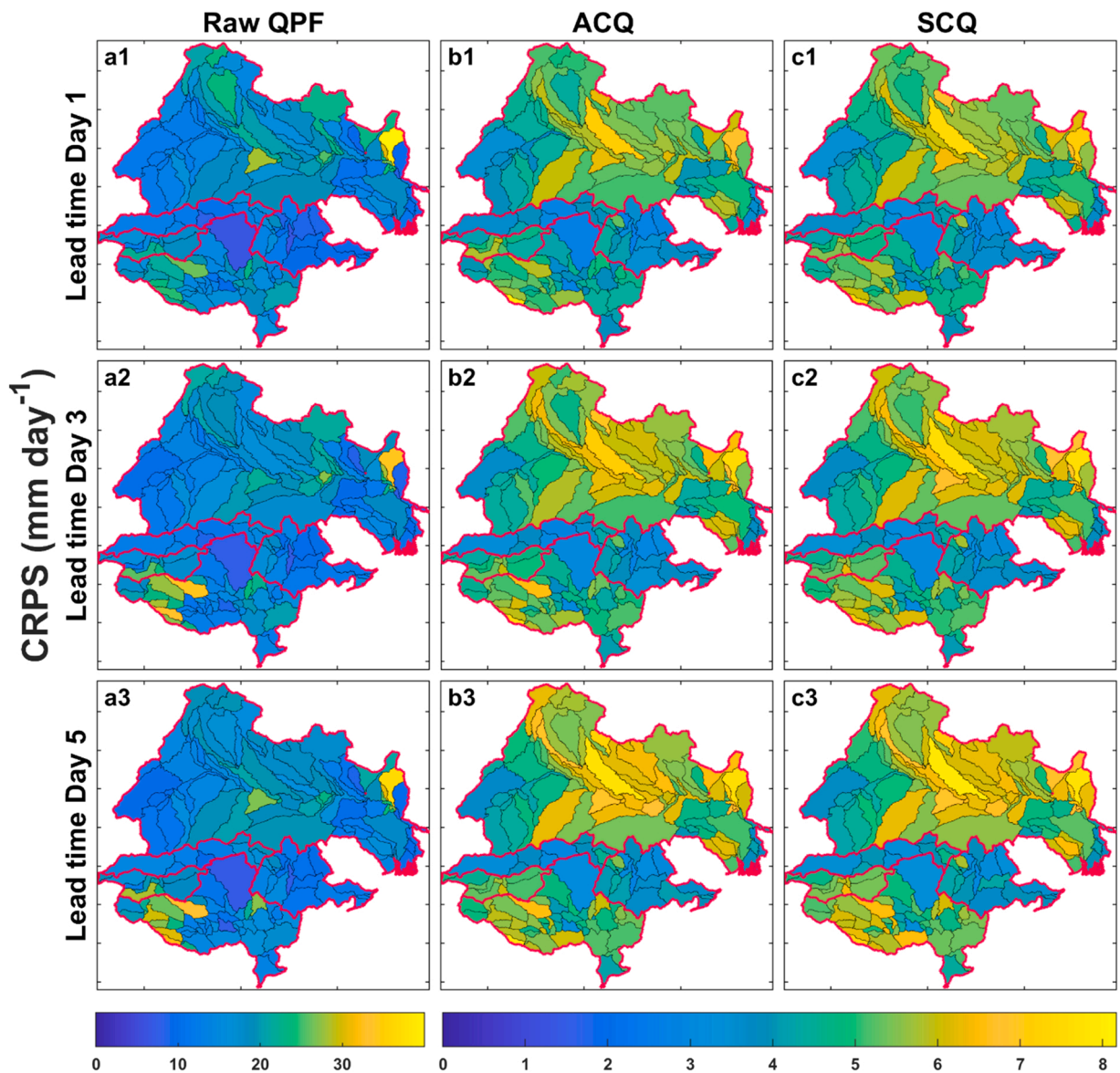


Fig. 4. Spatial plots of the sub-basin averaged CRPS (mm day^{-1}) of raw QPFs, Post-processed QPFs through RPP of ACQ and SCQ including lead time day-1, day-3, and day-5 for five river basins of Ganga, Mahanadi, Godavari, Narmada, and Tapti.

the raw annual QPF shows improvement in the bias percent as depicted in Fig. 2 (b1-b3). For the majority of the sub-basins, the bias values have decreased (−20 to 20) as compared to the bias of raw QPF (−95–300). In the case of ACQ, the bias % throughout the sub-basins shows overestimation of precipitation values mostly in Ganga and Mahanadi basins with bias 5–10% and under estimation in all

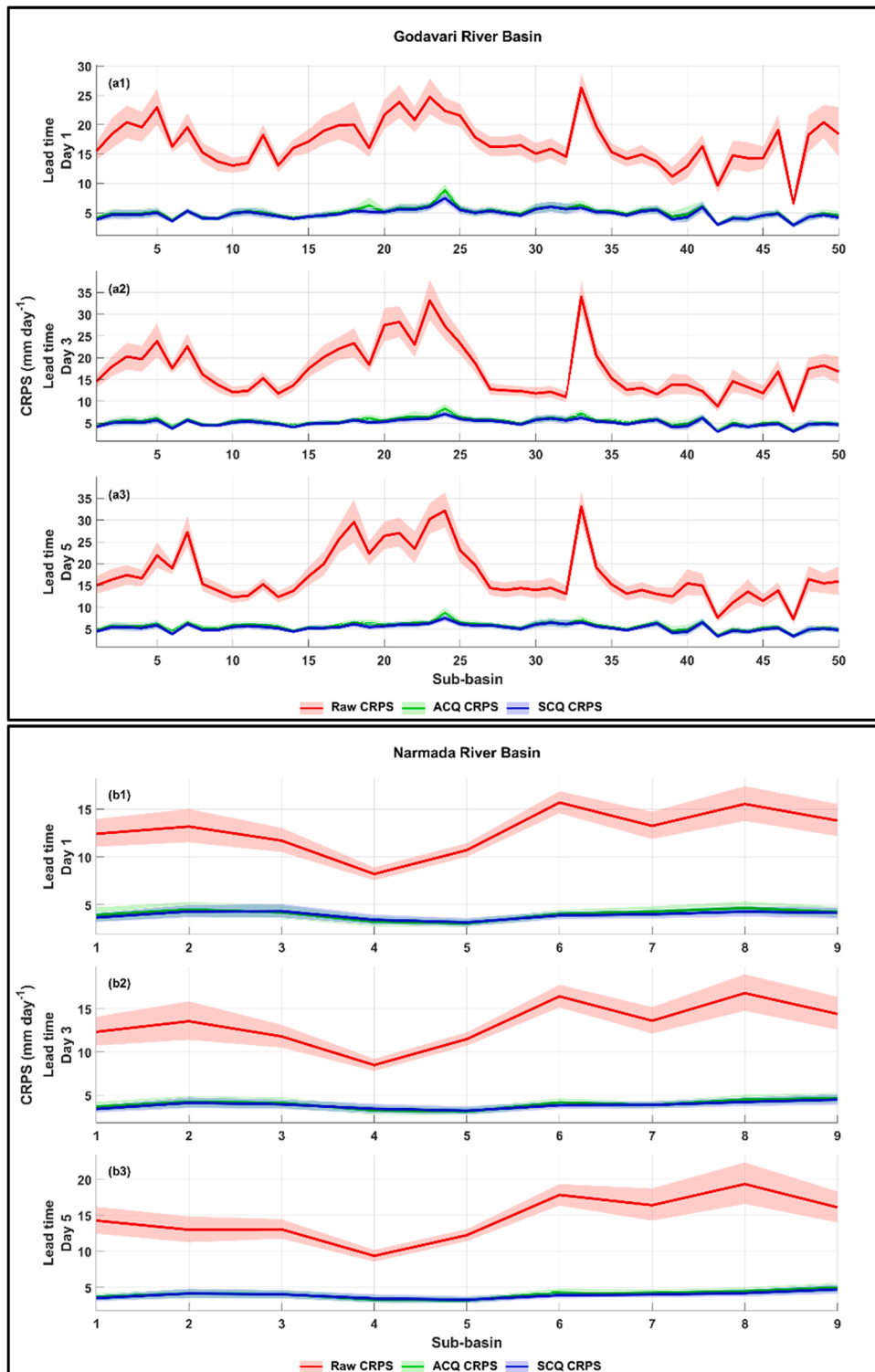


Fig. 5. Sub-basin averaged CRPS (mm day^{-1}) with 5% and 95% confidence interval (shaded region) for raw QPFs (red), post-processed QPFs from ACQ (green) and SCQ (violet) for the lead times day-1, day-3, and day-5 for the river basins Godavari and Narmada.

other sub-basins with bias less than -3% . Furthermore, in the case of SCQ, we observe a significant reduction in the bias percent as shown in Fig. 2 (c1-c3). The SCQs show better performance in all three lead times; however, we find very few sub-basins showing overestimation of the values at lead times 3-day and 5-day. Furthermore, the majority of the sub-basins show bias in the narrow range of -5 – 5 which means that calibration of the raw forecast using seasonal data performs better than the calibration using annual data. Moreover, this also suggests that the quality of the calibrated QPF depends upon the length of data and not necessarily on the location and size of the area under study.

Fig. 3 shows the bias plots of the raw deterministic QPF, ACQ and SCQ at the lead times of 1-day, 3-day and 5-day for the Godavari (larger basin) and Narmada (smaller basin) at 5% and 95% confidence intervals. In both the basins, we observe that the post-processing approach is well able to reduce the bias in the raw QPF. Specifically, in the case of Godavari (Fig. 3; a1-a3), we see that the bias in raw QPF has a large range between the positive and negative values in all three lead times. From the calibrated results, we observe that the performance of both ACQ and SCQ improves the quality of QPFs. SCQ shows a relatively straight line originating near the zero bias mark and running as a straight line parallel to the x-axis, denoting stable and skillful generation of forecast across all the fifty sub-basins. Similar observations can be made for all three lead times. In contrast, the ACQ performs better than the raw QPF; however, the results are slightly underestimated compared to the SCQ. The 5% and 95% confidence intervals around both ACQ and SCQ show a decrease when compared to that of the raw QPF (hence a narrow-shaded area). Furthermore, in the case of Narmada (Fig. 3; b1-b3), we observe nearly similar results to those observed for the Godavari basin. The raw QPF varies within a large range of positive and negative values which slightly increases with the increase in lead times. Compared to the Godavari basin, the ACQ of the Narmada performs better (less biased, more stable); however, it is still underestimated when compared to the ACQ. Overall, results from both the river basins suggest that the bias becomes least in SCQ which is followed by ACQ.

4.1.2. Evaluation using CRPS

The CRPSs are estimated based on the 300 ensemble members generated using the post-processing approach for both annual and seasonal calibrated QPFs. Fig. 4 shows the spatial plot of sub-basin averaged CRPS of the raw QPF (a1-a3), ACQ (b1-b3) and SCQ (c1-c3) at lead times of 1-day, 3-day and 5 days for all the 177 sub-basins. Results show that, in the case of raw QPF, the CRPSs vary within 0–40 mm for all three lead times. Higher CRPS are observed in the Godavari River basin (bottom half of the sub-basins), while in the majority of the other sub-basins, it ranges between 5 and 20 mm only. The finding remains the same in all three lead times with small variations. Further, the results show that the raw QPFs have significantly improved upon calibration in both the cases of ACQ and SCQ. The range of CRPS has reduced to 0–8 mm (after calibration) as opposed to 0–40 mm (before calibration), suggesting reasonable performance of the post-processing approach. The majority of the sub-basins are within a narrow range of 0–5 mm, while some sub-basins in the upper Ganga and western Godavari show a higher range of CRPS (5–8 mm). Moreover, the Fig. 4 shows that the spatial

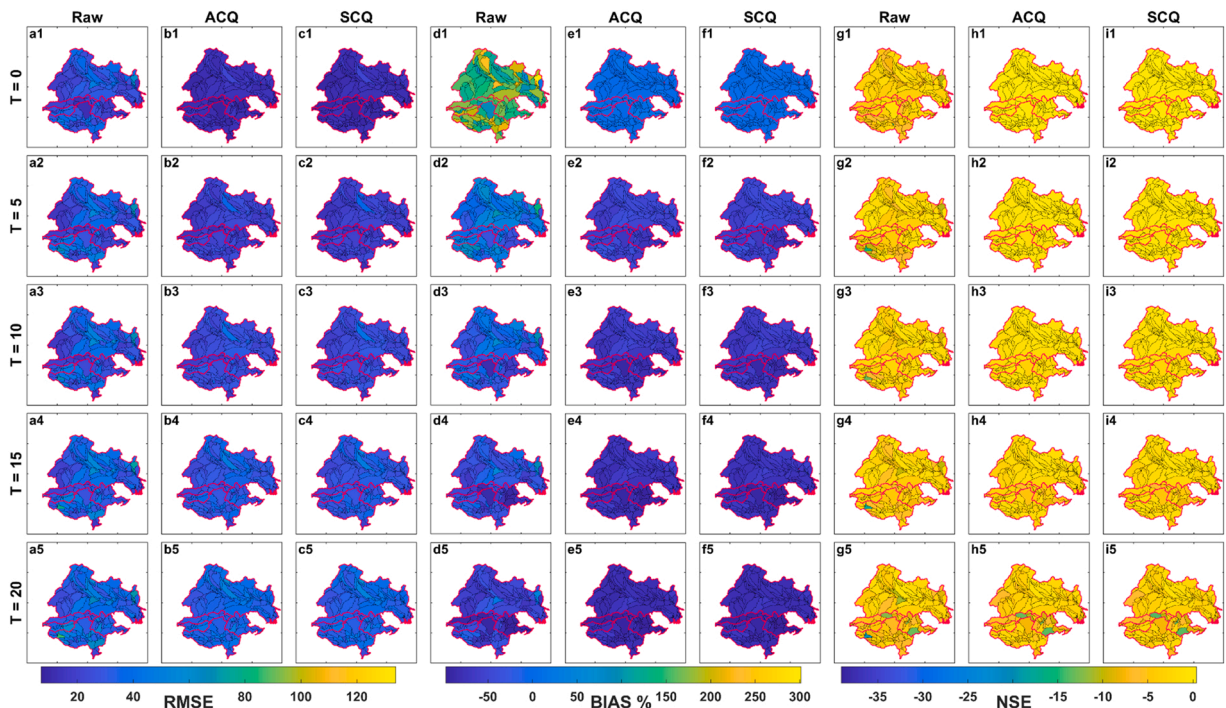


Fig. 6. Error statistics (RMSE (a1 to c5), Bias % (d1 to f5), and NSE (g1 to i5)) showing effect of post-processing on precipitation of different intensity (threshold 0, 5, 10, 15, and 20 mm per day) presented for lead time day-1 with raw QPF (1st column), ACQ (2nd column) and SCQ (3rd column).

variability of CRPS of ACQ and SCQ are almost similar which is contrary to the results observed using bias percentage.

Fig. 5 shows the CRPS plots of the raw deterministic QPF along with ACQ and SCQ at the lead times of 1-day, 3-day and 5-days for the Godavari and Narmada River basins at 5 % and 95 % confidence intervals. Visual inspection suggests that the post-processing approach is successful in reducing the CRPS in both the basins for both ACQ and SCQ. Specifically, in the Godavari basin (Fig. 5; a1-a3), we find that the CRPS of raw QPF fluctuates between 7 and 35 mm among the sub-basins. The results show that the CRPS has significantly reduced, in the case of both ACQ and SCQ, upon post-processing of the raw QPF. Moreover, the plots of ACQ and SCQ are less differentiated and are located within close proximity; however, we find that SCQ performs slightly better than ACQ in all three lead times. Furthermore, in the case of Narmada (Fig. 5; b1-b3), we observe similar results as the calibrated QPFs perform reasonably well as compared to the raw QPF. The CRPS of the raw QPF is in higher range (8 –15 mm) while the calibrated QPFs (ACQ and SCQ) are observed in a narrow range of 2–5 mm. Similar to the observation in the case of the Godavari basin, the post-processing approach shows comparable CRPS between ACQ and SCQ; however, the performance of SCQ is slightly better for all the sub-basins of the Narmada River.

4.2. Evaluation of QPFs at varying intensities

In this section, the performance of the calibrated forecasts (ACQ and SCQ) is evaluated at varying intensities at each sub-basin using the multiple error statistics such as RMSE, Bias %, and NSE at lead times of 1-day, 3-day and 5 days. The thresholds (T) of varying intensities are taken by carefully studying the daily sub-basin averaged observation data for the selected time period. To keep around 10 %, 15 % and 25 % data values above the higher thresholds, T = 10 mm, T = 15 mm, and T = 20 mm per day are selected respectively. According to that, the precipitation values are equally divided by taking five threshold values of T = 0 mm, T = 5 mm, T = 10 mm, T = 15 mm, and T = 20 mm per day. Fig. 6 presents the spatial plots of error statistics calculated for sub-basin averaged raw forecast from NCMRWF and QPFs from ACQ and SCQ in comparison with observation (IMERG) at each sub-basin using the T-values. For convenience, the results are presented for precipitation of lead time day 1 at the five river basins of Ganga, Godavari, Mahanadi, Narmada and Tapti. In Fig. 9 (a1 to a5, b1 to b5, and c1 to c5) presents RMSE calculated with observation for raw forecasts, ACQ, and SCQ at T = 0 mm, T = 5 mm, T = 10 mm, T = 15 mm, and T = 20 mm per day respectively. Similarly, results for Bias % and NSE are presented in the next six columns (d1 to i5) of the figure. The collective understanding of Fig. 9 shows that the values from each error statistics increases with increasing intensity of rainfall. Fig. 6 (a1 to a5) shows increasing RMSE values with increasing intensity of the precipitation and varies between 20 and 70 mm/day. The RMSE presented for ACQ (Fig. 6 b1 to b5) and SCQ (Fig. 6 c1 to c5) show lower RMSE values ranging from 0 to 40 mm/day, increasing with higher intensities. Furthermore, the bias % values of raw forecast show increased negative bias with increasing intensities of rainfall (Fig. 6 d1 to d5). The calibrated QPFs from both ACQ and SCQ show underestimation of the precipitation values at higher intensities (Fig. 6 e1 to f5). Similarly, NSE also supports that the ACQ and SCQ calibrated QPFs are underestimated at higher intensities of precipitation.

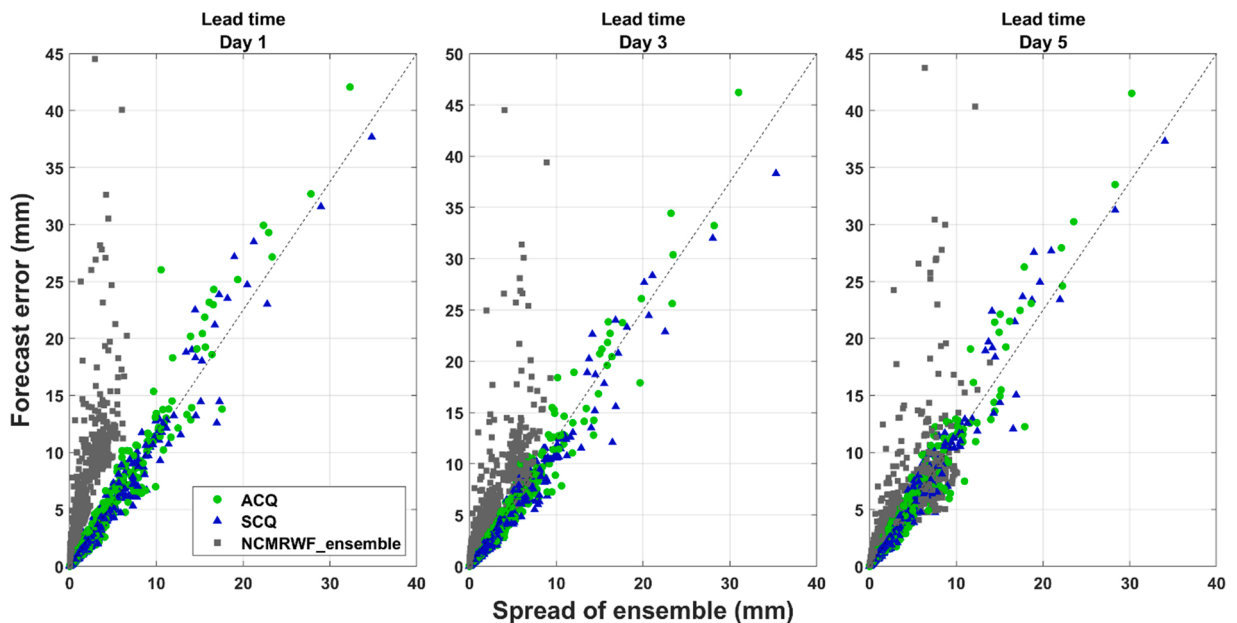


Fig. 9. The spread-skill analysis plot for monsoon season data for the NCMRWF ensemble forecast, ACQ, and SCQ run of RPP QPFs for lead time day 1, day-3, and day-5 for all the sub-basin.

4.3. Evaluation of the discriminating ability of the calibrated QPFs

In this section, we use the ROC plots to evaluate the ability of the calibrated QPFs to discriminate the precipitation events among the different sizes of the sub-basin. The sizes of the sub-basins depend upon the smallest, moderate, and largest area in the respective river basins of Godavari, Mahanadi, Narmada, Tapti and Ganga. Precipitation of more than 80th percentile (for each sub-basin averaged daily precipitation) is selected as the threshold for both annual and seasonal calibration at the lead times of 1, 3 and 5 days.

Fig. 7 and Fig. 8 show the ROC curves for the calibrated QPFs (for events of precipitation greater than 80th percentile for each sub-basin) to produce ACQs and SCQs, respectively. The figures show that the post-processed forecasts can discriminate between the events and non-events in most sub-basins. In both the cases of ACQ (Fig. 7) and SCQ (Fig. 8), we observe that the ROC curves are closer to the top-left corner of the plot for the lead time 1, thus indicating a good ability of forecasts to discriminate between precipitation events. However, as the lead time increases, the curves begin to move towards the diagonal, thus representing a lower discriminating ability of lead time 3-day and 5-day when the daily precipitation is more than 80th percentile. For instance, in Fig. 7 (b1), we find that the curve of lead time 1 is closer to the top-left corner of the plot (denoting a higher hit rate than false alarm rate) while the curves of the other two lead times are closer to the diagonal. In terms of AUC too, the value for the first lead time is found to be highest (0.795) followed by the second (AUC = 0.727) and third lead time (AUC = 0.657). Similar inference can be drawn for other basins, for instance, in Fig. 7 (e3), where the curve for the lead time 1-day in large-sized sub-basin of Ganga signifies a larger probability of hit rate while the other two lead times have a lower probability (larger probability of false alarm). Fig. 8 presents the ROC curves and AUC values from SCQ which depicts results similar to the ACQ. Furthermore, we observe that the temporal length of data used to set up the model has no significant influence over the discriminating ability of the forecast lead times since it remains largely the same for both ACQ and SCQ.

4.4. Spread-skill analysis of the ensemble QPFs

In this section, we compare the spread-skill of the model generated 300 ensemble members of ACQ and SCQ with the available perturbed forecasts from NCMRWF having 11 ensemble members. Fig. 9 shows the forecast error versus spread of the ensembles for ACQ and SCQ along with the NCMRWF perturbed forecast for the lead times of 1-day, 3-day and 5-days. From the figure, we observe that the NCMRWF perturbed ensemble contain high forecast error with minimal spread. This is evident since the scatter points are located close to the Y-axis in all three lead times. On the other hand, the ACQ and SCQ are located along the diagonal (1:1 line) suggesting good agreement between the magnitude of the error and spread of ensemble. Closer inspection of the plot suggests that the spread skill of SCQ is comparatively better than that of the ACQ, across all three lead times. The majority of the scatter points of ACQ are situated to the left side of the diagonal while those of the SCQ is well coinciding with the diagonal. Overall, the results show that, compared to the perturbed ensemble forecasts from NCMRWF, the calibrated QPFs (ACQ and SCQ) have a significantly better spread of

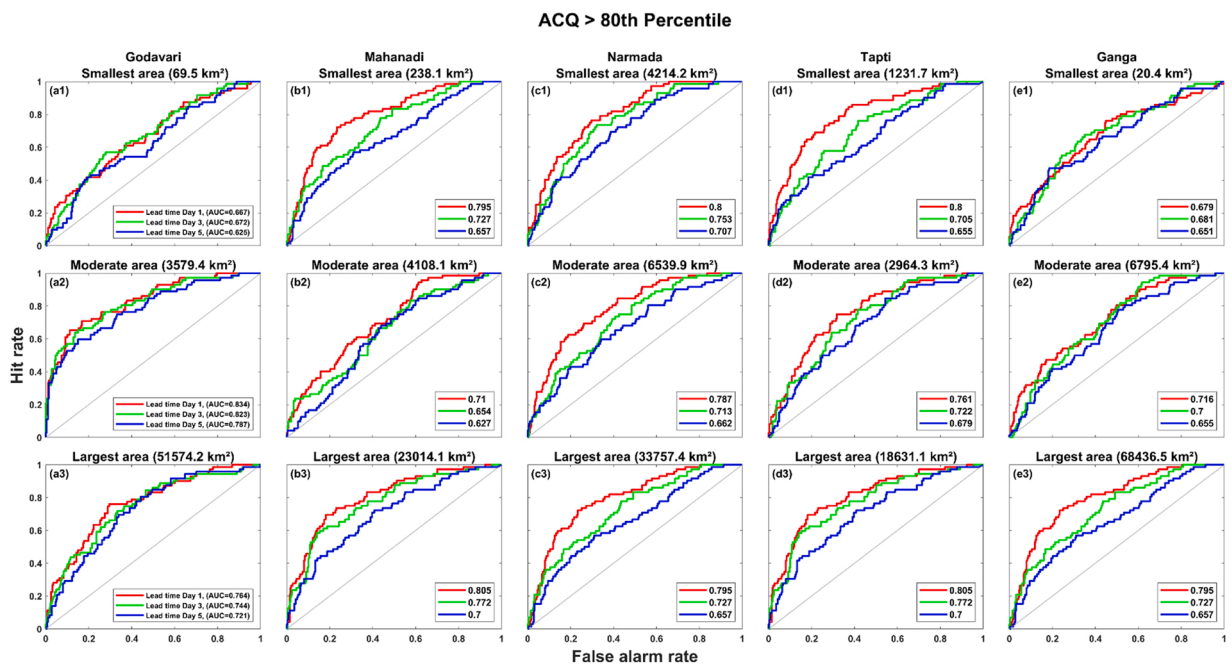


Fig. 7. Relative operating characteristic (ROC) curves presented for sub-basin averaged post-processed QPFs for events of precipitation greater than 80th percentile for day-1 for lead time 1 (red), 3 (green), and 5 (blue) days for river basins Godavari (a1-a3), Mahanadi (b1-b3), Narmada (c1-c3), Tapti (d1-d3), and Ganga (e1-e3) using annual dataset of July 2018 to Sep 2021. The results presented here correspond to the output of RPP run under ACQ.

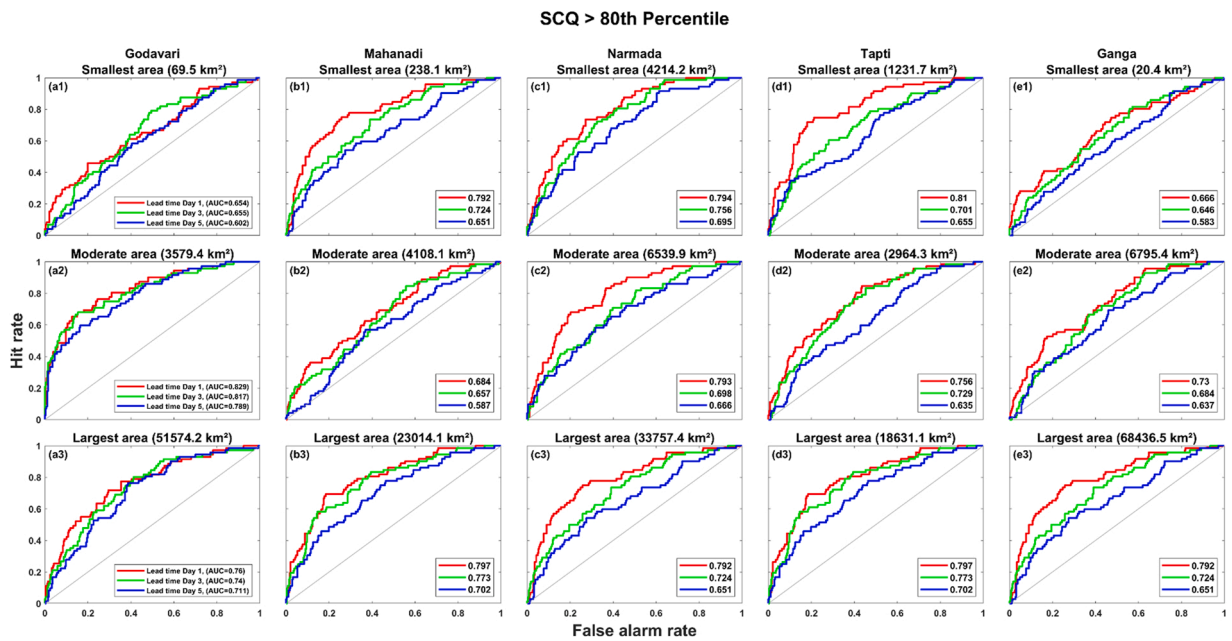


Fig. 8. Relative operating characteristic (ROC) curves obtained from SCQ results. ROCs are presented for sub-basin averaged post-processed QPFs for events of precipitation greater than 80th percentile for day-1 for lead time 1 (red), 3 (green), and 5 (blue) days for river basins Godavari a1-a3, Mahanadi (b1-b3), Narmada (c1-c3), Tapi (d1-d3), and Ganga (e1-e3) using monsoon (JJAS) data of 2019–2021.

ensemble along with low forecast error. Moreover, between ACQ and SCQ, the spread-skill of SCQ is higher.

4.5. Evaluation of the calibrated QPFs at the grid-scale

Here, the performance of the gridded calibrated forecasts (both ACQ and SCQ) in each basin is evaluated using the spatial maps, bias percent and CRPS at lead times of 1-day, 3-day and 5 days. Please note that the evaluation is presented only for the monsoon season in both the cases of ACQ and SCQ.

4.5.1. Evaluation using spatial maps

Fig. 10 presents the results of the calibrated QPFs in form of spatial maps at the gridded scale showing the gridded NCMRWF raw forecast (a1-a3), IMERG observation (b1-b3), and RPP calibrated mean QPFs from ACQ (c1 to c3) and SCQ (d1 to d3). The primary aim of applying the post-processing method is to bring the forecast closer to the observation. Visual inspection of Fig. 10 shows that RPP is capable of bringing the QPFs closer to the observation at all the basins throughout the lead times. The ACQ shows higher values at upper part of the Ganga basin which contains mountainous regions, whereas the SCQ calibrated forecasts shows a spatially matching pattern with the IMERG. The pattern of higher values of precipitation observed in the north-eastern part of Ganga basin has been maintained by the SCQ calibrated QPFs. The western part of the Ganga basin attributes to lower precipitation as its farther away from the Bay of Bengal. The similar pattern of lower precipitation is maintained in the ACQ and SCQ, which explains that RPP can manage to produce calibrated forecasts even using forecasts containing excessive zero values in the time series. The raw forecasts of a1 to a3 show overestimated values of up to 80 mm/day at some places, while IMERG ranges from 0 to 22 mm/day. The application of RPP on each grid through ACQ and SCQ is capable of bringing the calibrated forecasts to the same range as observation i.e., 0–22 mm/day.

4.5.2. Evaluation using percentage bias

Fig. 11 shows the spatial plot of gridded bias percent of the raw QPF (a1-a3), ACQ (b1-b3) and SCQ (c1-c3) at lead times of 1-day, 3-day and 5 days for all the five river basins. We observe high spatial variability in the bias % of the raw deterministic forecasts as seen in Fig. 11 (a1-a3) with bias values ranging from – 50–400 %. The upper part of the Ganga basin is showing high spatial variability if Bias % in comparison to the other basins. Large bias over the Northern Ganga basin could be attributed to the presence of mountainous regions, orographic influence, complex topography and frequent events of extreme precipitation. Further, results suggest that the ACQ obtained after post-processing the raw annual QPF shows improvement in the bias percent as depicted in Fig. 11 (b1-b3). For majority of the grids, the bias values decrease to (–40 to 40) as compared to the bias of raw QPF (–50–400). Furthermore, in the case of SCQ, we observe a significant reduction in the bias percent as shown in Fig. 11 (c1-c3). The bias for the majority of grids is in a narrow range of – 10–40 which means that calibration of the raw forecast using seasonal data for gridded precipitation performs better than the calibration using annual data. Moreover, this also suggests that the quality of the calibrated QPFs on gridded data with higher resolution depends upon the length of data and not necessarily on the location and size of the area under study. Both ACQ and SCQ

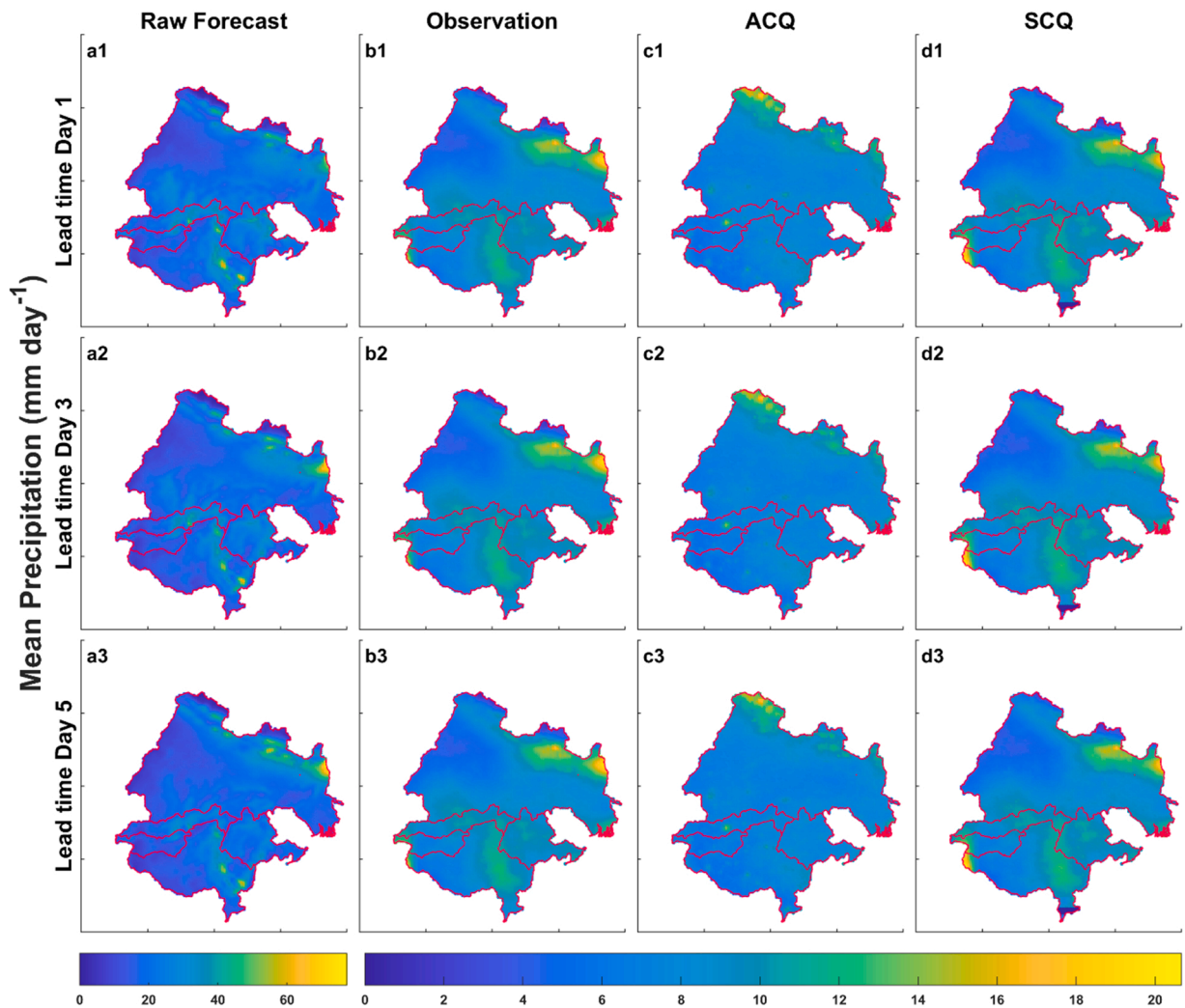


Fig. 10. Spatial plots of the gridded raw QPFs, mean post-processed QPFs of ACQ and SCQ through RPP of daily precipitation including lead time day-1, day-3, and day-5 for the five river basins of Ganga, Mahanadi, Godavari, Narmada, and Tapi.

performs similarly throughout all three lead times 1-day, 3-day, and 5-day suggesting ability of the RPP-model to produce calibrated QPFs with less bias even at higher lead times.

4.5.3. Evaluation using CRPS

The CRPSs are estimated based on the 300 ensemble members generated using the post-processing approach for both the annual and seasonal calibrated QPFs. Fig. 12 shows the spatial plot of mean CRPS of the gridded raw QPF (a1-a3), ACQ (b1-b3) and SCQ (c1-c3) at lead times of 1-day, 3-day and 5 days for all the basins. For the raw NCMRWF gridded QPFs, the mean CRPS varies from 0 to 65 mm for all three lead times. The higher raw mean CRPS values are observed in the lower Godavari basin and north eastern Ganga basin. Further, the results show that the raw gridded QPFs have significantly improved upon applying RPP for both ACQ and SCQ. The range of CRPS has reduced to 0–20 mm (after calibration) as opposed to 0–65 mm (before calibration), suggesting reasonable performance of the post-processing approach. Fig. 12 (b1 to b3) show increase in CRPS values with increasing lead times while SCQ (Fig. 12 (c1 to c3)) is showing similar variability in CRPS throughout all the three lead times of 1-day, 3-day and 5-day. Moreover, the Fig. 12 depicts that the SCQ calibrated gridded forecasts is performing better than the ACQ calibrated gridded forecasts.

5. Discussion

The rainfall post-processing approach applied in this study shows promising performance in improving the quality of the raw QPFs over the five river basins located across diverse landscapes in India. Results show that the approach is successfully able to reduce the bias in both the cases of ACQ and SCQ at both gridded and sub-basin level. The spatial plots and the line plots of the sub-basin averaged results suggest that the bias has significantly reduced in SCQ (close to zero) compared to ACQ. The reason for the better performance of

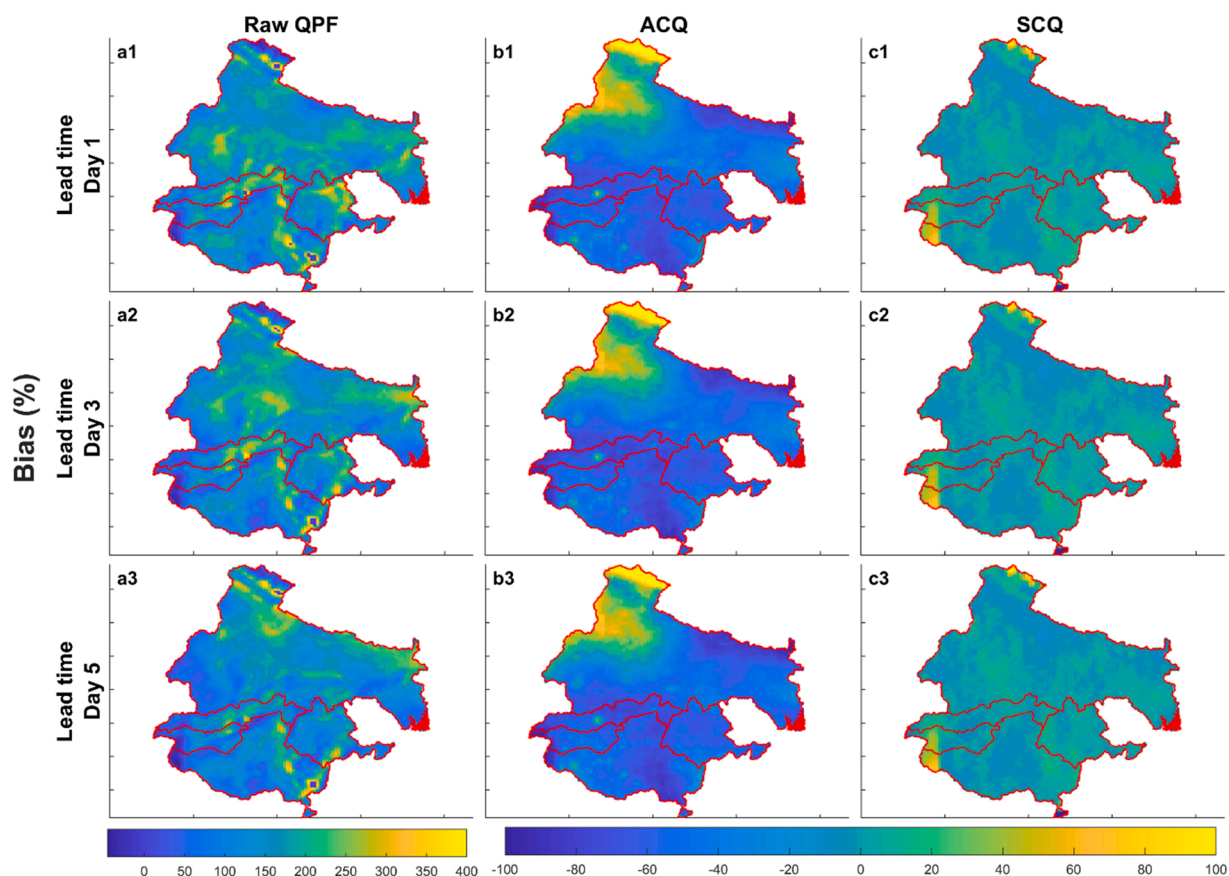


Fig. 11. Spatial plots of bias percentage of gridded raw QPFs, Post-processed QPFs through RPP of daily precipitation including lead time day-1, day-3, and day-5 for the five river basins of Ganga, Mahanadi, Godavari, Narmada, and Tapi.

the seasonal calibrated QPF could be that NCMRWF produces better forecasts for monsoon precipitation. In India, the distribution of both forecast and observed precipitation data in monsoon period is very different than the non-monsoon period. Representing both monsoon and non-monsoon data in a single run (ACQ) is inadequate. Moreover, the calibrated QPFs also show significantly reduced CRPS for both ACQ and SCQ. However, the difference between the CRPSs of ACQ and SCQ is much less as opposed to the results from bias percentage (SCQ showed considerably less bias compared to ACQ). The CRPSs of both ACQ and SCQ are comparable when observed using spatial plots. However, when observed using the line plots, we find that the SCQ performs slightly better than the ACQ. This reiterates the fact that the post-processor is more skillful and produces reliable ensemble forecasts with seasonal rainfall data (having less zero rainfall values).

Further, the ROC plots used to evaluate the ability of the calibrated QPFs to discriminate between the events (hit rate probability) and non-events (false alarm probability) of precipitation show that the shorter lead times (day-1) have greater ability to discriminate between events with respect to the longer lead times (day-3 and day-5). Moreover, we observe that the discrimination performance of the calibrated forecasts increases with the increase in the size of the sub-basin, which indicates that the size of the basin influences the performance of the calibrated QPFs. Furthermore, we observe that the discriminating ability of ACQ and SCQ are largely similar across the lead times and the various sub-basin size. This inference suggests that the temporal length of the data has no significant influence over the discriminating ability of the forecast lead times.

Further, we analyze the spread-skill of the calibrated ensembles of both ACQ and SCQ by comparing them with the perturbed ensemble forecasts from NCMRWF and find that the ACQ and SCQ have considerably better ability to represent the forecast errors than the NCMRWF perturbed forecasts. Moreover, between ACQ and SCQ, the spread-skill of SCQ is found to be better.

In addition to applying the RPP at the sub-basin scale, we also implemented the same at the individual grids of each basin. Upon analysis, we find that similar to the observations made at the sub-basin scale, RPP is well able to produce calibrated QPFs at the gridded scale. This suggests that the spatial scale involved in the post-processing approach have minimal role in the changes of results. However, we observe that the change in seasonal length of data does have a significant role to play at the gridded scale. For instance, the calibrated QPFs performed relatively better when only the seasonal data is provided to the model as opposed to the annual data. Consequently, at the gridded scale, SCQ is found to be closer to the observation compared to ACQ. One of the main reasons behind such an observation could be the presence of larger number of zero values in annual data compared to that of the seasonal data. Overall, results show that the post-processing approach is successful in preserving the spatial variation in calibrated QPFs from both annual and

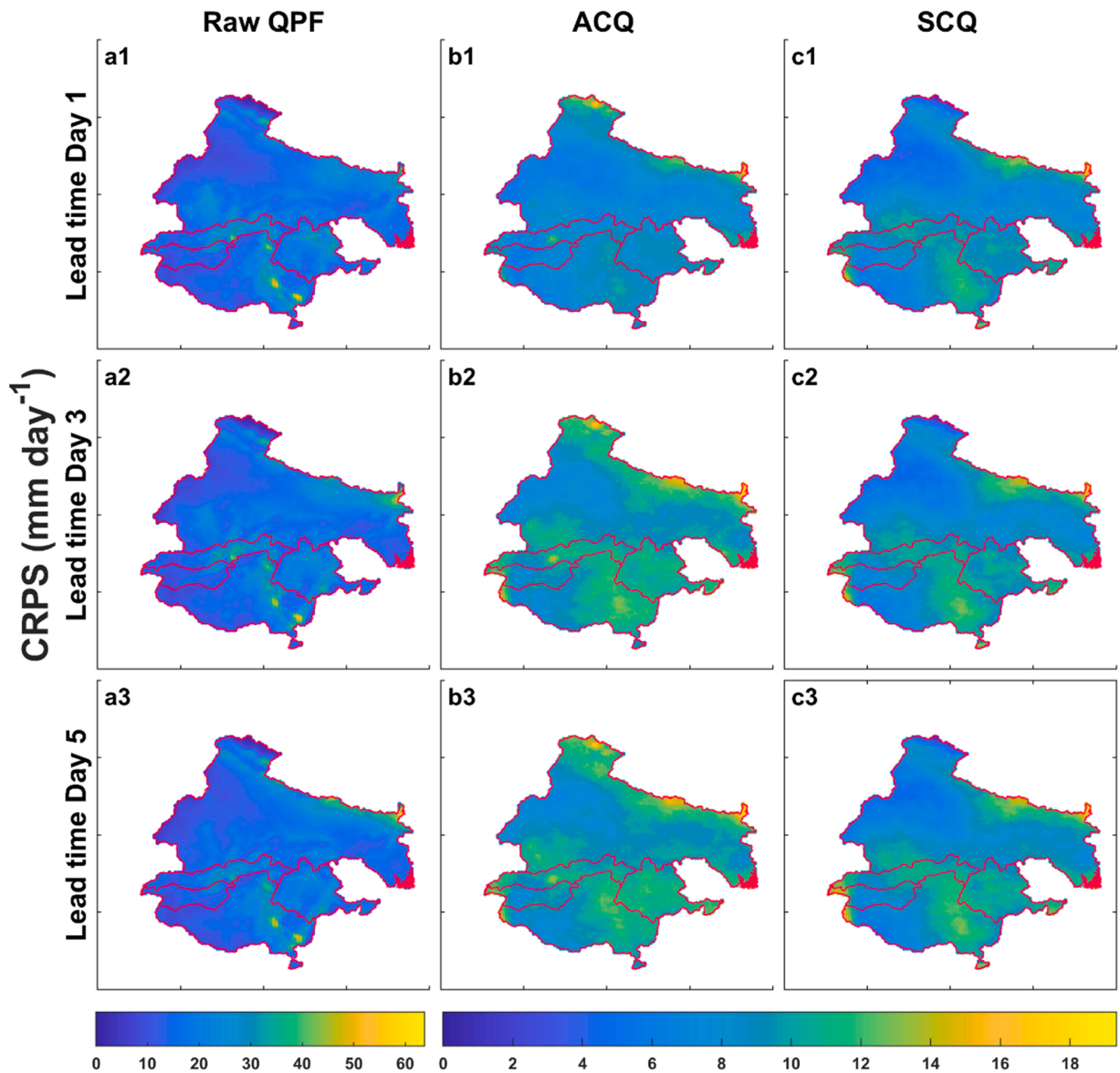


Fig. 12. Spatial plots of CRPS (mm day^{-1}) of gridded raw QPFs, Post-processed QPFs through RPP of daily precipitation including lead time day-1, day-3, and day-5 for the five river basins of Ganga, Mahanadi, Godavari, Narmada, and Tapi.

seasonal runs located across varied topography in India at both the sub-basin as well as the gridded scale. In this study, we also evaluate the results of the sub-basin with those at the gridded scale to understand the influence of spatial resolution, data length and lead times on the performance of the RPP. We observe that the performance of RPP does not depend upon the spatial scale of the data; however, its performance shows limitation when larger number of days with zero precipitation (dry days) are present in the data.

6. Conclusions

Availability of skillful and bias-free precipitation forecasts are essential at the sub-basin level for water management and decision-making. In this study, we attempt to statistically post-process deterministic QPFs obtained from the NCMRWF model and produce calibrated QPFs. To this end, a Bayesian joint probability model is applied to the gridded and sub-basin averaged precipitation over five river basins and lead times. The model is run using the raw QPFs at two temporal lengths (annual and seasonal) to produce 300 reliable ensemble members of both ACQ and SCQ. The spatio-temporal correlations in the ensemble members of both ACQ and SCQ are further adjusted corresponding to the correlations observed in the historical observation data using the Schaake shuffle. The calibration of the raw QPFs is carried out for all the grids and the 177 sub-basins of various sizes across five major river basins of India having diverse climatic and hydrological characteristics. Results show that the calibrated forecasts are less biased, more skillful, and more reliable

than the deterministic forecast. Specifically, the performance of seasonal calibrated QPF is better than that of the annual calibrated QPF. Moreover, the calibrated QPFs can discriminate the extreme daily precipitation events (and non-events) regardless of the size of the sub-basins. Further, in terms of the spread-skill (forecast error vs ensemble spread) of forecasts, we find that the skill of calibrated ACQ and SCQ is significantly better when compared to the available NCMRWF perturbed forecast with 11 ensemble members. Upon analysis of the results at the grid-scale, we find that the post-processing approach works well using the data at the gridded scale. This shows that the approach does not significantly depend upon the spatial scale. However, we observe that the presence of larger zero values in the annual data influences the results. At the grid-scale, results from SCQ are comparatively better compared to those from ACQ. In terms of lead times, results at both the spatial scales show deterioration in performance with increase in the lead times. The study is significant since it generates daily calibrated QPFs at the sub-basin scale which can be helpful to policy makers for streamflow forecasting, reservoir operations and issuing warnings regarding flood. The study area selected in the study is significantly dense in population where majority of the livelihood depends upon agriculture. The post-processing approach can be used to provide timely and reliable agro-met advisory services to minimize crop loss.

The post-processing approach can be further applied to a wide range of hydrological research. Future work will investigate the capability of the approach for grid-based precipitation at fine spatial resolution. Moreover, the efficacy of the approach can also be evaluated for producing continuous precipitation data with efficient forecasting skills at rain gauges.

CRedit authorship contribution statement

Nibedita Samal: Conceptualization, Methodology, Formal analysis, Writing – original draft, Investigation, Validation, Visualization, Data curation. **R Ashwin:** Data curation, Formal analysis, Visualization. **Akshay Singhal:** Writing and editing manuscript, Investigation, Visualization. **Sanjeev Kumar Jha:** Conceptualization, Methodology, Editing manuscript, Investigation, Supervision, Funding acquisition. **David E. Robertson:** Software, Conceptualization, Editing manuscript, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ejrh.2022.101284](https://doi.org/10.1016/j.ejrh.2022.101284).

References

- Ahmed, S., Coulibaly, P., Tsanis, I., 2014. Improved spring peak-flow forecasting using ensemble meteorological predictions. *J. Hydrol. Eng.* 20, 04014044. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001014](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001014).
- Bowler, N.E., Arribas, A., Mylne, K.R., Robertson, K.B., Beare, S.E., 2008. The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.* 134, 703–722. <https://doi.org/10.1002/QJ.234>.
- Buizza, R., 2018. Ensemble forecasting and the need for calibration. *stat. postprocessing ensemble. Forecast* 15–48. <https://doi.org/10.1016/B978-0-12-812372-0.00002-9>.
- Cai, C., Wang, J., Li, Z., 2019. Assessment and modelling of uncertainty in precipitation forecasts from TIGGE using fuzzy probability and Bayesian theory. *J. Hydrol.* 577, 123995. <https://doi.org/10.1016/j.jhydrol.2019.123995>.
- Clark, M.P., Hay, L.E., 2004. Use of medium-range numerical weather prediction model output to produce forecasts of streamflow. *J. Hydrometeorol.* 5, 15–32. [https://doi.org/10.1175/1525-7541\(2004\)005<0015:UOMNWP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0015:UOMNWP>2.0.CO;2).
- Das, S., Jain, M.K., Gupta, V., 2022. A step towards mapping rainfall erosivity for India using high-resolution GPM satellite rainfall products. *CATENA* 212, 106067. <https://doi.org/10.1016/J.CATENA.2022.106067>.
- Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J. Hydrol.* 158, 265–284. [https://doi.org/10.1016/0022-1694\(94\)90057-4](https://doi.org/10.1016/0022-1694(94)90057-4).
- Froude, L.S.R., 2010. TIGGE: comparison of the prediction of northern hemisphere extratropical cyclones by different ensemble prediction systems. *Weather Forecast* 25, 819–836. <https://doi.org/10.1175/2010WAF2222326.1>.
- Ghosh, S., Das, D., Kao, S.C., Ganguly, A.R., 2012. Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes. *Nat. Clim. Chang.* 2, 86–91. <https://doi.org/10.1038/nclimate1327>.
- Gneiting, T., Raftery, A.E., Westveld, A.H., Goldman, T., 2005. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* 133, 1098–1118. <https://doi.org/10.1175/MWR2904.1>.
- Guhathakurta, P., Revadekar, J., 2017. Observed variability and long-term trends of rainfall Over India. *Springer Geol.* 1–15. https://doi.org/10.1007/978-981-10-2531-0_1.
- Hamill, T.M., Whitaker, J.S., 2006. Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. *Mon. Weather Rev.* 134, 3209–3229. <https://doi.org/10.1175/MWR3237.1>.
- Hamill, T.M., Whitaker, J.S., Mullen, S.L., 2006. Reforecasts: an important dataset for improving weather predictions. *Bull. Am. Meteorol. Soc.* 87, 33–46. <https://doi.org/10.1175/BAMS-87-1-33>.

- Imhoff, R.O., Brauer, C.C., Overeem, A., Weerts, A.H., Uijlenhoet, R., 2020. Spatial and temporal evaluation of radar rainfall nowcasting techniques on 1,533 events. *Water Resour. Res.* 56. <https://doi.org/10.1029/2019WR026723>.
- Jha, S.K., Shrestha, D.L., Stadnyk, T.A., Coulibaly, P., 2018. Evaluation of ensemble precipitation forecasts generated through post-processing in a Canadian catchment. *Hydrol. Earth Syst. Sci.* 22, 1957–1969. <https://doi.org/10.5194/HESS-22-1957-2018>.
- Kumar Singh, A., Tripathi, J.N., Singh, K.K., Singh, V., Sateesh, M., 2019. Comparison of different satellite-derived rainfall products with IMD gridded data over Indian meteorological subdivisions during Indian Summer Monsoon (ISM) 2016 at weekly temporal resolution. *J. Hydrol.* 575, 1371–1379. <https://doi.org/10.1016/j.jhydrol.2019.02.016>.
- Lakshmi, D.D., Satyanarayana, A.N.V., Chakraborty, A., 2019. Assessment of heavy precipitation events associated with floods due to strong moisture transport during summer monsoon over India. *J. Atmos. Sol.-Terr. Phys.* 189, 123–140. <https://doi.org/10.1016/J.JASTP.2019.04.013>.
- Li, M., Jin, H., Shao, Q., 2021. Improvements in subseasonal forecasts of rainfall extremes by statistical postprocessing methods. *Weather Clim. Extrem.* 34, 100384. <https://doi.org/10.1016/J.WACE.2021.100384>.
- Li, R., Wang, K., Qi, D., 2018. Validating the integrated multisatellite retrievals for global precipitation measurement in terms of diurnal variability with hourly gauge observations collected at 50,000 stations in China, 10,423–10,442. *J. Geophys. Res. Atmos.* 123. <https://doi.org/10.1029/2018JD028991>.
- Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., Di, Z., 2017. A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdiscip. Rev. Water* 4. <https://doi.org/10.1002/wat2.1246>.
- Li, W., Duan, Q., Ye, A., Miao, C., 2019. An improved meta-Gaussian distribution model for post-processing of precipitation forecasts by censored maximum likelihood estimation. *J. Hydrol.* 574, 801–810. <https://doi.org/10.1016/j.jhydrol.2019.04.073>.
- Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, R.E., Hertig, E., Wibig, J., Huth, R., Wilcke, R.A.I., 2015. VALUE: A framework to validate downscaling approaches for climate change studies. *Earth's Futur* 3, 1–14. <https://doi.org/10.1002/2014EF000259>.
- Medina, H., Tian, D., Marin, F.R., Chirico, G.B., 2019. Comparing GEFs, ECMWF, and postprocessing methods for ensemble precipitation forecasts over Brazil. *J. Hydrometeorol.* 20, 773–790. <https://doi.org/10.1175/JHM-D-18-0125.1>.
- Nester, T., Komma, J., Viglione, A., Blöschl, G., 2012. Flood forecast errors and ensemble spread—a case study. *Water Resour. Res.* 48, 10502. <https://doi.org/10.1029/2011WR011649>.
- Nicotina, L., Alessi Celegon, E., Rinaldo, A., Marani, M., 2008. On the impact of rainfall patterns on the hydrologic response. *Water Resour. Res.* 44. <https://doi.org/10.1029/2007WR006654>.
- Nipen, T., Stull, R., 2011. Calibrating probabilistic forecasts from an NWP ensemble. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.* 63, 858–875. <https://doi.org/10.1111/J.1600-0870.2011.00535.X>.
- Prakash, S., Mitra, A.K., AghaKouchak, A., Liu, Z., Norouzi, H., Pai, D.S., 2018. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* 556, 865–876. <https://doi.org/10.1016/j.jhydrol.2016.01.029>.
- Ramos, M.H., Van Andel, S.J., Pappenberger, F., 2013. Do probabilistic forecasts lead to better decisions? *Hydrol. Earth Syst. Sci.* 17, 2219–2232. <https://doi.org/10.5194/HESS-17-2219-2013>.
- Ray, K., Giri, R.K., Ray, S.S., Dimri, A.P., Rajeevan, M., 2021. An assessment of long-term changes in mortalities due to extreme weather events in India: A study of 50 years' data. 1970–2019. *Weather Clim. Extrem.* 32, 100315. <https://doi.org/10.1016/J.WACE.2021.100315>.
- Robbins, J.C., Tittle, H.A., 2018. Evaluating high-impact precipitation forecasts from the met office global hazard map (GHM) using a global impact database. *Meteorol. Appl.* 25, 548–560. <https://doi.org/10.1002/MET.1720>.
- Robertson, D.E., 2013. Post processing rainfall forecasts from numerical weather prediction models for short term streamflow forecasting. *Hydrol. Earth Syst. Sci. Discuss.* 10, 6765. <https://doi.org/10.5194/hessd-10-6765-2013>.
- Roulin, E., Vannitsem, S., 2012. Postprocessing of ensemble precipitation predictions with extended logistic regression based on hindcasts. *Mon. Weather Rev.* 140, 874–888. <https://doi.org/10.1175/MWR-D-11-00062.1>.
- Roxy, M.K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., Terray, P., Rajeevan, M., 2017. A threefold rise in widespread extreme rain events over central India. *Nat. Commun.* 2017 81 (8), 1–11. <https://doi.org/10.1038/s41467-017-00744-9>.
- Saminathan, S., Medina, H., Mitra, S., Tian, D., 2021. Improving short to medium range GEFs precipitation forecast in India. *J. Hydrol.* 598, 126431. <https://doi.org/10.1016/j.jhydrol.2021.126431>.
- Schaake, J., Demargne, J., Hartman, R., Mullusky, M., Welles, E., Wu, L., Herr, H., Fan, X., Seo, D.J., 2007. Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrol. Earth Syst. Sci. Discuss.* 4, 655–717. <https://doi.org/10.5194/hessd-4-655-2007>.
- Scheppen, A., Zhao, T., Wang, Q.J., Robertson, D.E., 2018. A Bayesian modelling method for post-processing daily sub-seasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments. *Hydrol. Earth Syst. Sci.* 22, 1615–1628. <https://doi.org/10.5194/hess-22-1615-2018>.
- Shrestha, D.L., Robertson, D.E., Bennett, J.C., Wang, Q.J., 2015. Improving precipitation forecasts by generating ensembles through postprocessing. *Mon. Weather Rev.* 143, 3642–3663. <https://doi.org/10.1175/MWR-D-14-00329.1>.
- Shrestha, D.L., Robertson, D.E., Bennett, J.C., Wang, Q.J., Perraud, J., 2016. Preliminary Analysis of Rainfall Forecast Post-Processor Model Parameters. Singh, A., Jha, S.K., 2021. Identification of sensitive parameters in daily and monthly hydrological simulations in small to large catchments in Central India. *J. Hydrol.* 601, 126632. <https://doi.org/10.1016/J.JHYDROL.2021.126632>.
- Singh, A., Tiwari, S., Jha, S.K., 2021. Evaluation of quantitative precipitation forecast in five Indian river basins. <https://doi.org/10.1080/02626667.2021.1982138>.
- Singhal, A., Jha, S.K., 2021a. An application of Multiple-point statistics downscaling approach over North-West Himalayas in avalanche-prone areas. *Int. J. Clim.* 1–20. <https://doi.org/10.1002/joc.7342>.
- Singhal, A., Jha, S.K., 2021b. Can the approach of vulnerability assessment facilitate identification of suitable adaptation models for risk reduction? *Int. J. Disaster Risk Reduct.* 63, 102469. <https://doi.org/10.1016/j.ijdr.2021.102469>.
- Singhal, A., Cheriamparambil, A., Jha, S.K., 2022a. Spatial extrapolation of statistically downscaled weather data over the Northwest Himalayas at major glacier sites. *Environ. Model. Softw.* 149, 105317. <https://doi.org/10.1016/j.envsoft.2022.105317>.
- Singhal, A., Raman, A., Jha, S.K., 2022b. Potential use of extreme rainfall forecast and socio-economic data for impact-based forecasting at the district level in Northern India. *Front. Earth Sci.* 0, 761. <https://doi.org/10.3389/FEART.2022.846113>.
- Tian, J., Liu, J., Yan, D., Ding, L., Li, C., 2019. Ensemble flood forecasting based on a coupled atmospheric-hydrological modeling system with data assimilation. *Atmos. Res.* 224, 127–137. <https://doi.org/10.1016/J.ATMOSRES.2019.03.029>.
- Verkade, J.S., Brown, J.D., Reggiani, P., Weerts, A.H., 2013. Post-processing ECMWF precipitation and temperature ensemble forecasts for operational hydrologic forecasting at various spatial scales. *J. Hydrol.* 501, 73–91. <https://doi.org/10.1016/J.JHYDROL.2013.07.039>.
- Wang, Q.J., Robertson, D.E., Chiew, F.H.S., 2009. A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resour. Res.* 45, 1–18. <https://doi.org/10.1029/2008WR007355>.
- Wang, Q.J., Zhao, T., Yang, Q., Robertson, D., 2019. A seasonally coherent calibration (SCC) model for postprocessing numerical weather predictions. *Mon. Weather Rev.* 147, 3633–3647. <https://doi.org/10.1175/MWR-D-19-0108.1>.
- Yagli, G.M., Yang, D., Srinivasan, D., 2020. Ensemble solar forecasting using data-driven models with probabilistic post-processing through GAMLSS. *Sol. Energy* 208, 612–622. <https://doi.org/10.1016/J.SOLENER.2020.07.040>.