The evaluation of weekly extended range river basin rainfall forecasts and a new bias correction mechanism for flood management in India

Pulak Guhathakurta, Ashwini Kumar Prasad, Rajib Chattyopadhyay, Neha Sangwan, Nilesh Wagh, D.R. Pattanayak, D.S. Pai, M. Mohapatra

India Meteorological Department, Pune, India

Abstract

Operational extended range forecasts are being disseminated once every week by the India Meteorological Department (IMD) for several sectorial applications. These forecasts show a reduction in amplitude and variance as a function of lead-time. Such reductions in variance can be due to several physical factors: inherent forecast model bias, a problem relating to initial conditions, lead-dependent statistical biases, etc. A week-by-week analysis shows that such biases are not systematic. Rainfall forecasts are underestimated in some regions, while others overestimate rainfall amplitude. To correct the bias in the extended range weekly averaged forecast, a statistical post-processing method (normal ratio correction) is proposed to make the outlook more valuable at a longer lead-time. The correction method is based on the World Meteorological Organization (WMO) technical guidance on rainfall estimation and is also shown to be useful for rainfall forecasts. In this analysis, we evaluate the extended range forecast skill at the river sub-basin-scale and show that there are several river sub-basins over the central Indian region where the correction has improved the model forecast in the one to two-week range. Although this analysis was tailored toward making the river basins and sub-basins of India more readily realizable for flood forecasters, it can be used for any administrative boundaries such as block, district, or state-level requirements.

Keywords
Extended range prediction, rainfall bias correction, flood management.

Submitted 22 July 2021, revised 15 February 2022, accepted 17 February 2022
DOI: 10.26491/mhwm/146785

1. Introduction

One of the major purposes of the extended range forecast is to provide a high-resolution spatial-temporal forecast on a weekly scale, with up to 2-3 weeks lead time. The weekly outlooks can provide important input to the decision-making process of various stakeholders (Pattanaik, Das 2015; Chattopadhyay et al. 2018; Pattanaik et al. 2019; Sahai et al. 2019a-b). The extended range forecasts of the India Meteorological Department (IMD) are being generated at a spatial resolution of 1°X1° Grid. Forecasts of rainfall in different river basins have several important hydrometeorological applications, especially in flood forecasting based on rainfall variables (Ming et al. 2020; Webster, Hoyos 2004; Webster et al. 2010; Gilewski, Nawalany 2018; Sayama et al. 2020; Gilewski 2021). In this regard, the most critical application is the forecast of heavy rainfall (daily rainfall of 7 cm or more) in the river basins. This can lead to flooding and inundations, and a considerable loss of life and property. Hence, for flood forecasting and other hydrometeorological requirements, the precipitation forecast must be as quantitative as possible. However, it is occasionally found that there is a significant bias in amplitude and variance in the forecasted rainfall. This often leads to a severe underestimation of rainfall outlook, thereby adding additional input errors to hydrological models that use these quantitative rainfall forecasts.
The Ministry of Earth Sciences, the Government of India, and the IMD have the mandate to provide rainfall forecast and long-range outlooks in the S2S (seasonal to sub-seasonal) scale\(^1\). Such predictions are required to be as accurate as possible. Bias in rainfall forecasts is a common problem in raw model forecast data, which can be problematic for quantitative precipitation forecast. Such amplitude biases in rainfall in longer lead times arise due to inefficient representation of model physics and dynamics, or due to systematic errors in the large-scale forcing. To make the forecast more useful, these biases should be reduced as much as possible. One crucial error in rainfall forecasts is the underrepresentation of rainfall amplitude after a forecast lead-time of a few days. The forecast often shows that the variance is severely underrepresented in the forecasted rainfall, as lead-time increases. Several statistical post-processing methods, using complex to simple approaches to correct the rainfall bias, exist to improve the rainfall forecast under such circumstances (Boé et al. 2007; Leander, Buishand 2007; Ghimire et al. 2019). These bias corrections are shown to improve hydrological forecasts (Teutschbein, Seibert 2012). The results show that a bias in rainfall arising due to improper amplitude attenuations as a function of lead-time, could be corrected under many circumstances – provided climatological or observed rainfall amplitude is known for any lead day. This shows promise for correcting amplitude bias arising in operational dynamical models (Singh et al. 2017; Jabbari, Bae 2020).

The presented study aimed to provide better basin-wise weekly rainfall forecasts for the river sub-basins of India, using a novel method to correct the forecast bias in the extended range weekly forecast. The forecast ability of extended range weekly rainfall forecasts, as well as bias-corrected extended range weekly rainfall forecasts, were satisfactory in both 1-week and 2-weeks lead time. These forecasts can thus be used as model inputs for flood forecasting. Although this study focused on basin-wise rainfall forecast, the method is general and can be applied to the average rainfall of any administrative boundaries or geographical locations like districts or states.

2. Data and methodology

2.1. Data

The current study uses the daily observed rainfall data from the IMD 0.25 Deg × 0.25 Deg gridded data (Pai et al. 2014) and daily 1Deg × 1Deg rainfall forecast data received from the IMD_IITM extended range forecast system \([\text{ERF}]\). Daily observed gridded rainfall data of 0.25 Deg × 0.25 Deg has been generated by the IMD from the quality controlled daily rainfall data of rain gauge stations\(^2\). The dataset covers a geographical domain of 6.5\(^\circ\)-38.5\(^\circ\)N and 66.5\(^\circ\)-100.0\(^\circ\)E and contains only values from land regions. The extended range forecast models generate precipitation forecast data up to four weeks in advance, based on the conditions observed at any given time (Chattopadhyay et al. 2019; Pattanaik et al. 2019, Sahai et al. 2019b). Currently, an operationally extended range forecast is disseminated once every week. For every week’s operational forecast, there is corresponding “on the fly” hindcasts for the same set of recorded

\(^1\) https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1706073
\(^2\) https://imdpune.gov.in/Clim_Pred_LRF_New/Grided_Data_Download.html
conditions since 2003. The data is generated for the global domain. In this analysis we have used the data for the river basins of India. The Central Water Commission divides the country into 25 major river basins and 101 river sub-basins (Fig. 1 and Table 1). River sub-basins shapefiles were obtained from the Central Water Commission. Using the shapefile, the gridded data (both observation and forecast data) was masked and basin averaged data was prepared for each of the 101 river basins.

![Map of 101 River Sub-basins of India As Per Central Water Commission (CWC)](image)

**Fig. 1. Locations of the 101 river sub-basins of India.**

**Table 1. List of the 101 river sub-basins of India and their codes.**

| Sub Basin No. | SB CODE | Major Basin | SUB_BASIN |
|---------------|---------|-------------|-----------|
| 1             | ARA     | Ganga Basin | Above Ramganga Confluence |
| 2             | BAI     | Brahmani and Baitarni Basin | Baitarni |
| 3             | BAM     | Indus (Up to border) Basin | Barmer |
| 4             | BAN     | Ganga Basin | Banas |
| 5             | BAR     | Barak and others Basin | Barak |
| 6             | BEA     | Indus (Up to border) Basin | Beas |
| 7             | BHA     | West flowing rivers of Kutch and Saurashtra including Luni Basin | Bhadar and other west flowing rivers |
| 8             | BHG     | Ganga Basin | Bhagirathi and others (Ganga Lower) |
| 9             | BHL     | Krishna Basin | Bhima Lower |
| 10            | BHT     | West flowing rivers South of Tapi Basin | Bhatsool and others |
| 11            | BHU     | Krishna Basin | Bhima Upper |
| 12            | BRA     | Brahmani and Baitarni Basin | Brahmani |
| 13            | BRL     | Brahmaputra Basin | Brahmaputra Lower |
|   |   |   |
|---|---|---|
| 14 | BRU | Brahmaputra Basin | Brahmaputra Upper |
| 15 | CAL | Cauvery Basin | Cauvery Lower |
| 16 | CAM | Cauvery Basin | Cauvery Middle |
| 17 | CAU | Cauvery Basin | Cauvery Upper |
| 18 | CHA | Indus (Up to border) Basin | Chautang and others |
| 19 | CHE | Indus (Up to border) Basin | Chenab |
| 20 | CHL | Ganga Basin | Chambal Lower |
| 21 | CHR | Indus (Up to border) Basin | Churu |
| 22 | CHU | Ganga Basin | Chambal Upper |
| 23 | DAL | Drainage Area of Lakshadweep Islands Basin | Drainage Area of Lakshadweep Islands |
| 24 | DAM | Ganga Basin | Damodar |
| 25 | DAN | Drainage Area of Andaman and Nicobar Islands Basin | Drainage Area of Andaman and Nicobar Islands |
| 26 | GAN | Ganga Basin | Gandak and others |
| 27 | GDL | Godavari Basin | Godavari Lower |
| 28 | GDM | Godavari Basin | Godavari Middle |
| 29 | GDU | Godavari Basin | Godavari Upper |
| 30 | GHA | Ganga Basin | Ghaghara |
| 31 | GHG | Ganga Basin | Ghaghara Confluence to Gomti confluence |
| 32 | GHO | Indus (Up to border) Basin | Ghaghari and others |
| 33 | GIL | Indus (Up to border) Basin | Gilgit |
| 34 | GOM | Ganga Basin | Gomti |
| 35 | GTK | East flowing rivers between Godavari and Krishna Basin | East flowing rivers between Godavari and Krishna |
| 36 | IMP | Minor rivers draining into Myanmar Basin | Imphal and others |
| 37 | IND | Godavari Basin | Indravati |
| 38 | JHE | Indus (Up to border) Basin | Jhelum |
| 39 | KAS | Ganga Basin | Kali Sindh and others up to Confluence with Parbati |
| 40 | KOS | Ganga Basin | Kosi |
| 41 | KPO | Minor rivers draining into Bangladesh Basin | Karnaphuli and others |
| 42 | KRL | Krishna Basin | Krishna Lower |
| 43 | KRM | Krishna Basin | Krishna Middle |
| 44 | KRU | Krishna Basin | Krishna Upper |
| 45 | KTP | East flowing rivers between Krishna and Pennar Basin | East flowing rivers between Krishna and Pennar |
| 46 | KYN | Barak and others Basin | Kynchiang and other south flowing rivers |
| 47 | LIN | Indus (Up to border) Basin | Lower Indus |
| 48 | LUL | West flowing rivers of Kutch and Saurashtra including Luni Basin | Luni Lower |
| 49 | LUU | West flowing rivers of Kutch and Saurashtra including Luni Basin | Luni Upper |
| 50 | MAJ | Godavari Basin | Manjira |
| 51 | MAL | Mahanadi Basin | Mahanadi Lower |
| 52 | MAM | Mahanadi Basin | Mahanadi Middle |
| 53 | MAN | Minor rivers draining into Myanmar Basin | Mangpu Lui and others |
| 54 | MAU | Mahanadi Basin | Mahanadi Upper |
| 55 | MHO | Minor rivers draining into Bangladesh Basin | Muhury and others |
|   |   |   |
|---|---|---|
| 56 | MHU | Mahi Basin | Mahi Upper |
| 57 | NAG | East flowing rivers between Mahanadi and Godavari Basin | Nagvati and others |
| 58 | NAM | Narmada Basin | Narmada Middle |
| 59 | NAO | Barak and others Basin | Naoch chara and others |
| 60 | NAU | Narmada Basin | Narmada Upper |
| 61 | NET | West flowing rivers South of Tapi Basin | Nethravati and others |
| 62 | PAL | East flowing rivers between Pennar and Cauvery Basin | Palar and other |
| 63 | PAM | East flowing rivers between Pennar and Cauvery Basin | Pamba and others |
| 64 | PAR | West flowing rivers South of Tapi Basin | Periyar and others |
| 65 | PEL | Pennar Basin | Pennar Lower |
| 66 | PEU | Pennar Basin | Pennar Upper |
| 67 | PON | East flowing rivers between Pennar and Cauvery Basin | Ponnaiyar and other |
| 68 | PRA | Godavari Basin | Pranathi and others |
| 69 | RAM | Ganga Basin | Ramganga |
| 70 | RAN | West flowing rivers of Kutch and Saurashtra including Luni Basin | Drainage of Rann |
| 71 | RAV | Indus (Up to border) Basin | Ravi |
| 72 | SAL | Sabarmati Basin | Sabarmati Lower |
| 73 | SAR | West flowing rivers of Kutch and Saurashtra including Luni Basin | Saraswati |
| 74 | SAU | Sabarmati Basin | Sabarmati Upper |
| 75 | SHE | West flowing rivers of Kutch and Saurashtra including Luni Basin | Shetranjuli and other east flowing rivers |
| 76 | SHK | Area of North Ladakha not draining into Indus Basin | Shaksgam |
| 77 | SHY | Indus (Up to border) Basin | Shyok |
| 78 | SON | Ganga Basin | Sone |
| 79 | SUB | Subernarekha Basin | Subernarekha |
| 80 | SUL | Indus (Up to border) Basin | Sutlaj Lower |
| 81 | SUM | Area of North Ladakha not draining into Indus Basin | Sulmar |
| 82 | SUU | Indus (Up to border) Basin | Sutlaj Upper |
| 83 | TAM | Tapi Basin | Tapi Middle |
| 84 | TAU | Tapi Basin | Tapi Upper |
| 85 | TON | Ganga Basin | Tons |
| 86 | TUL | Krishna Basin | Tungabhadra Lower |
| 87 | TUU | Krishna Basin | Tungabhadra Upper |
| 88 | UGO | Ganga Basin | Upstream of Gorriti confluence to Muzaffarnagar |
| 89 | UIN | Indus (Up to border) Basin | Upper Indus |
| 90 | VAI | East flowing rivers South of Cauvery Basin | Vaippar and others |
| 91 | VAM | East flowing rivers between Mahanadi and Godavari Basin | Vamsadhara and other |
| 92 | VAR | West flowing rivers South of Tapi Basin | Varrar and others |
| 93 | VAS | West flowing rivers South of Tapi Basin | Vasishthi and others |
| 94 | WAR | Godavari Basin | Wardha |
| 95 | WEI | Godavari Basin | Weinganga |
| 96 | YAL | Ganga Basin | Yamuna Lower |
2.2. Methodology

Weekly grid point cumulative rainfall data for the years 2003-2019 were used. Average cumulative rainfall (in mm) in each sub-basin for every week was calculated using the Raster Statistics Method in the QGIS Software. More details on the operational extended-range forecast can be seen in a study by Sahai et al. (2019b). A diagram illustrating how the operational forecast is generated currently, is shown in Figure 2. The operational extended-range forecast is an ensembled mean of four dynamical models. Two of them are high resolution (denoted by suffix T382 or ~38 km), and two are low resolution models (indicated by suffix T126 or ~110 km). Two of the models have coupled models (CFS), and two are atmospheric models (GFS). Each model shares the same dynamic core but slightly different physics and resolutions. Each model has 4-member ensemble runs. Thus, we have a total of $4 \times 4 = 16$ ensemble members from the CFST126, CFST382, GFST126, and GFST382 models for runs from each set of condition. Atmospheric and oceanic initial conditions were generated by NCMRWF and INCOIS, respectively. The sea surface temperature boundary conditions for the GFS were derived from the CFS runs. Since the CFS sea surface temperature has a bias, a simple bias correction using observed climatology was applied to generate the final input boundary conditions for the GFS.

There are various metrics for evaluating rainfall forecasts (e.g., Barnston 1992; Huang, Zhao 2022). Several papers have used root mean square errors and correlation coefficients as a first order measure to evaluate the deterministic forecast ability of rainfall in the extended range (e.g., Joseph et al. 2019). Similarly, there are several methods to evaluate hydrological forecasts (e.g., Hoshin et al. 2009; Gilewski, Nawalany 2018). In this study, we computed and compared the "normalized root mean square error" (NRMSE) and "correlation coefficient" of raw extended range weekly forecast data (hereafter ERF) and bias-corrected forecast data (hereafter BERF). For basin averaged extended range forecast, it would be shown that the bias-corrected forecast improves the model performance.
Fig. 2. Schematics of the end-to-end forecast and dissemination system implemented for Extended Range Prediction in India. Abbreviations: CFSv2: Coupled forecasting system version2; GFSv2: Global forecasting system; T382 and T126 suffixes specify the CFS and the GFS model horizontal spatial resolutions and different spectral truncations. T382 implies models CFS or GFS model run at ~38 km resolution. In comparison, T126 indicates the model to run at ~110 km spatial resolution. The suffix "bc" suggests that the GFS runs with bias-corrected sea surface temperature (SST) as boundary conditions. "FCST" indicates forecast runs, and ICs indicate initial conditions from which the model is run. INCOIS (Indian National Center for Ocean Information services) and NCMRWF (National Centre for Medium-Range Weather Forecasting) generate the ICs (i.e., initial conditions) for the operational forecasts. Models are an adapted version of models developed at the National Centre for Environmental Predictions (NCEP), USA, which also generates the ICs. MISO or MJO is the intra-seasonal monsoon oscillations and the Madden Julian Oscillations.

2.3. Bias correction with Normal Ratio Method

For the estimation of missing or unknown rainfall values, a normal ratio method is suggested by WMO (2018). We have adopted a similar approach to perform bias correction of the raw extended rainfall forecast (ERF) by multiplying the Bias correction ratio with the raw ERF rainfall. The normal ratio method is generally used for rainfall estimation, whereas difference correction is advised for temperature and other parameters.

According to the normal ratio method, the missing precipitation is given as:

$$ P_x = \frac{1}{n} \sum_{i=1}^{n} \frac{N_x}{N_i} P_i $$  \hspace{1cm} (1)

$$ P_x $$ is the missing precipitation for any storm at the interpolation station 'x', $$ P_i $$ is the precipitation for the same period for the same storm at the "i"th station of a group of index stations, $$ N_x $$ is the normal...
In statistical modeling, the precipitation value for the 'x' station and \( N_i \) the normal precipitation value for 'i' station. In our bias correction method, \( P_i \) is the precipitation from raw \( ERF \), \( N_i \) is the climatology of \( P_i \), \( N_i \) is the observed climatology, and \( P_x \) is the bias corrected \( ERF \).

Figure 3 shows the climatological differences between raw \( ERF \) rainfall and realized rainfall of 101 sub-basins of India for each of the 18 weeks of southwest monsoon. The first week of this period was from 30th May to 5th June and the last week was from 26th Sept to 2nd Oct (as 18th week). It can be seen that \( ERF \) has no systematic bias, as it is overestimating in some areas and underestimating in others. These differences also changed as the monsoon progresses. During the initial onset phase of the monsoon in June, the \( ERF \) climatology was higher than the observed climatology in most sub-basins. Still, during the peak monsoon period from July to August, \( ERF \) underestimated the rainfall for most sub-basins. Particularly during week number 8 (18th Jul to 25th July), \( ERF \) climatology was less significant for all the sub-basins of India – except one sub-basin in the extreme eastern parts of India. Another important finding in \( ERF \) was overestimation throughout the season, except one or two weeks for the sub-basins over Bihar, east UP, and adjacent areas. Thus, bias correction based on the normal ratio method has to be applied for all the weeks separately. This overcomes both the underestimating and overestimating of the raw \( ERF \) rainfall forecast and makes the prediction closer to the realized one. Thus, the bias correction ratio was different for each basin, as well as for each week during the monsoon onset, progress, and retreat phases.

The bias correction ratio for each of the 101 sub-basins and all the 18 weeks during the southwest monsoon season was estimated by the ratio of Actual Rainfall Climatology (for the same week in the period 2003-2019) and \( ERF \) Climatology (for the same week in the period 2003-2019). Here we have used the normal ratio in equation (1) for the estimation of missing rainfall, as the bias correction ratio in our bias correction method. This assists in improving the forecast value by giving weight to observed climatology.

Therefore, to improve the accuracy of the sub-basin rainfall forecast, we have adopted a new bias correction method given as follows:

The Bias Corrected Rainfall Forecast for each week and each basin = \( ERF \) (Rainfall) for that week X Bias correction ratio for the corresponding week of the same basin.

The correlation coefficient is one of the possible choices for forecast verification (Barnston 1992) and is given as:

\[
r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}
\]  

(2)

where: \( r \) – correlation coefficient; \( x_i \) – values of the x-variable in a sample; \( \bar{x} \) – mean of the values of the x-variable; \( y_i \) – values of the y-variable in a sample; \( \bar{y} \) – mean of the values of the y-variable.

In statistical modeling, another way of measuring the quality of the fit of the model, is the \( RMSE \) (also called Root Mean Square Deviation) (Barnston 1992) given by:
\[ RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \]  

where \( y_i \) is the \( i \)th observation of \( y \) and \( \hat{y}_i \) the predicted \( y \) value given the model. If the predicted responses are very close to the correct responses, the RMSE will be small. If the predicted and true responses differ substantially – at least for some observations – the RMSE will be large.

To compare RMSE of rainfall forecast of the different river basins with different mean rainfall patterns, we have used Normalized Root Mean Square Error (NRMSE) as:

\[ NRMSE = \frac{RMSE}{\text{Mean (observed values)}} \]  

In the next sections, the NRMSE and the correlation coefficient will be used as the standard skill score measures to evaluate the improvement in the rainfall forecast.
3. Results and discussion

3.1. Performances of forecast

From a hydrological forecast perspective, the monsoon onset phase is perhaps the most important phenomenon. Every year, the onset over Kerala, and its subsequent propagation over the Indian Landmass, is monitored for agrometeorological predictions. The onset phase is often associated with a northward propagating rainfall pulse, providing rain over large regions of India and several river basins.

Rainfall during the onset phase of the monsoon is crucial for agricultural planning. Additionally, most flood events occur during July and August, when the monsoon is active. Week-by-week performances of the week 1 extended-range forecast, as well as the bias-corrected forecast, are shown for June (Fig. 4a), July (Fig. 4b), August (Fig. 4c), and September (Fig. 4d) of 2003-2019.

For all the weeks, the Normalized RMSE of bias-corrected ERF was less than 1 in most cases and for most sub-basins. In week 1 (Fig. 4a), due to bias correction, NRMSE of ERF has been reduced from 2.4 to 0.5 for the Drainage Area of Andaman and Nicobar Islands sub-basin, from 1.6 to 0.4 for the Drainage Area of Lakshadweep Islands sub-basin, from 1.7 to 1.4 for the Sulmar sub-basin, 1.0 to 0.7 for the Kynchiang sub-basin, and other south-flowing rivers of Barak basin during onset phases of the SW monsoon. For all four weeks of June (Fig. 4a), NRMSE of these sub-basins were high (more than 1.5) for raw ERF, whereas due to bias correction, NRMSE has come down by around 0.5. Furthermore, for all four weeks of June, NRMSE of bias-corrected ERF was less than the NRMSE of raw ERF. This was within 0.2 to 0.9 for all the sub-basins, except a few sub-basins in the first week and one sub-basin in the second and third weeks.

In the first week of July (27th Jun to 3rd July) (Fig. 4b), the bias corrections of several sub-basins have helped to improve the NRMSE by keeping it less than 0.8. In the following two weeks, though the NRMSE of bias-corrected ERF was less than the raw ERF for all the sub-basins, there was no significant improvement. However, in the last two weeks of July, significant improvement of the ability of bias-corrected ERF was seen for most of the sub-basins.
Fig. 4a. Normalized RMSE of Bias corrected ERF (BERF) and raw ERF of Week 1 rainfall during June.

Fig. 4b. Normalized RMSE (NRMSE) of Bias corrected ERF (BERF) and raw ERF of Week 1 rainfall during July.
There is a remarkable improvement in the skill of bias-corrected ERF for the first two weeks of August (Fig. 4c), as NRMSE of bias-corrected ERF was between 0.2 to 0.6 in most of the sub-basins. Since most of the floods in India occur during July and August, bias correction can help improve flood forecasts and better flood management.

Even during all the weeks of September (Fig. 4d), bias correction reduced the NRMSE value to well below 1.0 of the NRMSE value and greater than 1 of raw ERF.

Fig. 4c. Normalized RMSE of Bias corrected ERF (BERF) and raw ERF of Week 1 rainfall during August.
3.2. The spatial pattern of Extended Range onset forecast skill for the period 2003-2019

To demonstrate the skill of the extended range forecast for the 1-week and 2-weeks lead-time during the monsoon season, we have computed the correlation coefficient and the normalized root mean square error (NRMSE) map between the ERF and observed rainfall for the years 2003-2019. The samples consisted of 18 weeks and 17 years (18×17 = 306 samples) for each of the 101 sub-basins for the monsoon season. Figures 5a-b show the basin-wise map of the correlation coefficients and normalized root mean square error for the raw ERF (left panels), respectively. The plot indicates relatively high correlations in the central and northern Indian basins and relatively low correlations in the southern peninsular basins. Furthermore, there were low correlations and higher NRMSE in the Jammu, Kashmir, and Ladakh regions. The root mean square error in Figure 5b shows that the model had the lowest error in central and northern India.

Similarly, Figures 5c-d show the same skill metrics for the bias-corrected forecast. The bias-corrected forecast shows some improvement in correlation skills in the Maharashtra sub-basins and some basins of peninsular India. There was also a significant decrease in RMSE over the basins of central to southern peninsular India.

Figure 6 shows the same skill plots but for the 2-week forecast.
Fig. 5. (a) Correlation map of 1-week actual forecast, (b) RMSE of the actual forecast, (c) same as (a) but showing the cc map for bias-corrected forecast, (d) same as (b) but showing RMSE of bias-corrected forecast.
3.4. Floods in Maharashtra and Bihar during 2019 and the evaluation of skill forecast for the year 2019

In 2019, several parts of the country had experienced severe floods affecting lakhs of people (Shagun 2019; Kambli 2020). During July and August 2019, heavy flooding occurred in Maharashtra due to intense rainfall. The Sangli and Kolhapur district in the Krishna sub-basin experienced severe floods of long durations. Substantial losses of life, property and crops were reported. At the beginning of the flood period, i.e., from 27th Jul to 3rd Aug, heavy rainfall events were localized in the northern part of the Konkan and adjoining North Madhya Maharashtra. Many stations in the Pune and Nasik districts recorded rainfall of more than 150 mm/day from 3rd to 5th August. Towards the latter part of the week, the rainfall belt shifted towards southern Madhya Maharashtra. Mahabaleshwar recorded the highest rainfall of 380 mm on 5th Aug 2019. It is also observed that the Kolhapur district continuously experienced heavy rain throughout this period, with the highest rainfall amounts on 6th Aug 2019. Gaganbawda recorded its highest rainfall of 340 mm on 6th Aug 2019. It is also seen that, though heavy rainfall occurred in the western part of the districts in Madhya Maharashtra, their eastern parts were devoid of rainfall. Furthermore, during the heavy rain spell of Aug 2019, many stations in the Kolhapur district and western parts of the Satara district have surpassed their previous record of 7 days rainfall. Compared to 2018, rainfall over the region was widespread and remained very intense for an extended period from 27th Jul to 13th Aug 2019.
(Government of Maharashtra 2020). The expert committee of the Government of Maharashtra recommended that IMD 1-week and 2-week river sub-basin rainfall forecasts should be used in flood forecasting to improve the accuracy of the forecast. Another major affected state was Bihar, where around 306 lives were lost due to floods and heavy rain.

We have analyzed the 1-week forecasted rainfall of raw ERF compared to the actual rainfall for all the sub-basins of these two states, and showed how the bias-corrected forecast could have helped the flood management. The losses could have been minimized by using the bias-corrected forecast for these regions.

Figure 7 shows the realized, bias-corrected ERF and ERF rainfall for 18 weeks of SW Monsoon season of 2019. This includes the sub-basins viz. Godavari Upper, Godavari Middle, Wardha, Wainganga, Tapi Middle, Bhima Upper, Krishna Upper, Bhatso and others, and Vasishti and other Flood-affected Maharashtra states. In the 9th and 10th weeks (25th Jul to 31st Jul and 1st Aug to 7th Aug), all nine of these sub-basins reported a significant increase in rainfall compared to previous weeks, which raw ERF was not able to predict in most of the cases. The sub-basins Weinganga, Vasishti, and others also reported increased rainfall activity in 11-weeks. The raw ERF underestimated the rainfall for all these basins. Applying the bias correction forecast to rainfall from these basins was almost comparable to that of the realized rainfall, indicating the usefulness of the bias-corrected Week 1 rainfall in improving flood management.
Fig. 7. Realized, Bias corrected $ERF$ ($BERF$) and $ERF$ rainfall for 18 weeks of the 2019 SW Monsoon season in 9 sub basins of the flood affected Maharashtra state.
For the Bihar flood, we have selected four sub-basins viz. Ghaghara, Ghaghara Confluence to Gomti confluence, Gandak and others, and Koshi. Figure 8 shows the realized, bias-corrected ERF and ERF rainfall for 18 weeks of the 2019 SW Monsoon season for sub-basins of Flood-affected Bihar. In the 6th and 7th weeks, all four sub-basins have reported increased rainfall activities causing devastating flooding over this region. The week 1 raw ERF rainfall has been overestimated in all these cases. The bias correction could help to minimize the differences between observed rainfall and forecast rainfall.

Fig. 8. Realized, Bias corrected ERF (BERF) and ERF rainfall for 18 weeks of SW Monsoon season 2019 of sub basins of flood-affected districts of the Bihar state.

To see the performance of raw ERF and bias-corrected ERF for the year 2019, the correlation coefficient between observed and forecast rainfall and normalized RMSE was calculated using 18 samples (all eighteen weeks of SW monsoon 2019) for both the 1-week and 2-week lead forecasts. Figure 9 shows the (a) correlation and (b) RMSE of the raw extended range forecast, calculated using the weekly data for the year 2019. (c) same as (a) but after using bias correction. (d) same as (b) but after using bias correction for the 1-week lead forecast.

The left column shows the raw extended range forecast, and the right column shows the corresponding bias-corrected forecast. There was a significant improvement in the correlation coefficient for most sub-basins, mainly over the northern and central parts of India. The normalized root means square error shows that there was a considerable improvement in the bias-corrected forecast, especially in the east and central parts of India, as normalized RMSE has been reduced to less than 0.3 due to bias correction over these parts. Additionally, in the western parts of Maharashtra, NRMSE has been reduced from near 1 in raw ERF to less than 0.5.
Fig. 9. (a) Correlation and (b) RMSE of the raw extended range forecast calculated using the weekly data for the year 2019 of 1-week forecast. (c) same as (a) but after using bias correction. (d) same as (b) but after using bias correction.

Figure 10 shows the (a) correlation and (b) RMSE of raw extended range forecast calculated using the weekly data for the year 2019. (c) same as (a) but after using bias correction. (d) same as (b) but after using bias correction for the 2-week lead forecast.

The left column shows the actual extended range forecast, and the right column shows the corresponding bias-corrected forecast. In the 2-week forecast, the correlation coefficient for the sub-basins of Maharashtra has been increased from around 0.7-0.8 in raw ERF to 0.93-0.97. The correlation coefficient is between 0.7-0.8 in most of the sub-basins of central India in the bias-corrected forecast. Normalized RMSE is also less than 0.5 in the bias-corrected forecast for most of the sub-basins of India, with central India being less than 0.3.
4. Conclusions

For efficient flood and disaster management, an accurate rainfall forecast is essential to provide a quantitative prediction of precipitation during the June to September (monsoon) season over river basins of the Indian subcontinent. The weekly averaged extended range rainfall forecast of up to 2-weeks lead-time is important, as it provides a valuable input for generating flood forecast models in a time-scale that is crucial for water and dam management. A proper rainfall forecast with a longer lead time is always desirable to manage floods and their impact on disaster risk reduction. India's present operational flood forecasting models are primarily dependent on 1-3 days quantitative rainfall forecast and a forecast of up to 5 days generated by India Meteorological Department. In the extended range (i.e., 2-weeks lead time) the rainfall forecast is often not accurate, owing to the decrease in rainfall amplitude. In the current study, we have provided a comprehensive basin averaged rainfall skill analysis over different sub-basins of India, using the extended range retrospective forecast and proposing a bias correction method to improve the rainfall forecast in the extended range. We have found that the extended forecast has an unsystematic bias (i.e., overestimation and underestimation) for weekly averaged rainfall. The bias in precipitation is not systematic, and different sub-basins show the bias of different amplitude. Such amplitude biases would likely impact
Our bias corrected forecast has shown significant skill in predicting sub basin rainfall of 1-week as well as 2-weeks lead time.

We hypothesized that a part of the amplitude bias might be associated with systematic forecast model bias. Due to rainfall forecast error associated with model physics, dynamics, and several other factors, such biases can arise. Using an amplitude correction method based on the "Normal Ratio" correction method from the WMO manual, we devised an approach to see if the normal ratio correction would improve the first-order skill scores (root mean square error and correlation) for weekly extended range forecast over the Indian land region. The results show an encouraging improvement in statistical skill scores for several river basins over India. The long-term (2003-2019) skill analysis shows enough improvement in the weekly mean forecast. Similarly, case studies over the Maharashtra and Bihar river basins for 2019 show significant improvement in the weekly mean rainfall estimates. We also verified the week-by-week forecast from the onset to the withdrawal phase. The onset phase rainfall forecast over different sub-basins shows sufficient improvement. We propose that the analysis could be used as a background for operational forecast bias correction using the normal ratio method. This can be implemented for products based on extended range forecast and all forecast products in the sub-seasonal to seasonal (s2s) time-scale. Thus, these extended range basin rainfall forecasts of 1-week and 2-week lead times have shown good skill during the 2003-2019 period. In addition to existing flood forecasting systems of the central water commission of India, these findings can be used for generating flood forecasts with longer lead times to reduce disaster impacts.

Acknowledgment
Authors are grateful to the National Water Informatics Centre Ministry of Jalshakti for providing the river basin and sub-basin shapefiles that help initiate the weekly river subbasin wide rainfall forecast. We also acknowledge the India Meteorological department for making grid point rainfall data accessible to for observation and forecasting.

References
Barnston A.G., 1992, Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score, Weather and Forecasting, 7 (4), 699-709, DOI: 10.1175/1520-0434(1992)007<0699:CATCRA>2.0.CO;2.
Boé J., Terray L., Habets F., Martin E., 2007, Statistical and dynamical downscaling of the Seine basin climate for hydrometeorological studies, International Journal of Climatology, 27 (12), 1643-1655, DOI: 10.1002/joc.1602.
Chattopadhyay R., Krishna R.P.M., Joseph S., Dey A., Mandal R., Sahai A.K., 2018, A comparison of extended-range prediction of monsoon in the IITM-CFSv2 with ECMWF S2S forecast system, IITM Research Report No. RR-139, available online https://www.tropmet.res.in/~lip/Publication/RR-pdf/RR-139.pdf (data access 28.02.2022).
Chattopadhyay R., Susmitha J., Abhilash S., Mandal R., Dey A., Phani R., Saranya G., Kaur M., Pattanaik D.R., Sahai A.K., 2019, Understanding the intraseasonal variability over Indian region and development of an operational extended range prediction system, Mausam, 70 (1), 31-36, DOI: 10.54302/mausam.v70i1.166.
Ghimire U., Srinivasan G., Agarwal A., 2019, Assessment of rainfall bias correction techniques for improved hydrological simulation, International Journal of Climatology, 39 (4), 2386-2399, DOI: 10.1002/joc.5959.
Gilewski P., 2021, Impact of the grid resolution and deterministic interpolation of precipitation on rainfall-runoff modeling in a sparsely gauged mountainous catchment, Water, 13 (2), DOI: 10.3390/w13020230.

Gilewski P., Nawalany M., 2018, Inter-comparison of rain-gauge, radar, and satellite (IMERG GPM) precipitation estimates performance for rainfall-runoff modeling in a mountainous catchment in Poland, Water, 10 (11), DOI: 10.3390/w10111665.

Government of Maharashtra, 2020, A report on flood 2019 (Krishna Sub-basin), available online https://wrd.maharashtra.gov.in/Upload/PDF/Vol%2020%20Main%20Report.pdf (data access 28.02.2022).

Gupta H.V, Kling H., Yilmaz K.K., Martinez G.F., 2009, Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377 (1-2), 80-91, DOI: 10.1016/j.jhydrol.2009.08.003.

Huang Z., Zhao T., 2022, Predictive performance of ensemble hydroclimatic forecasts: Verification metrics, diagnostic plots and forecast attributes, WIREs Water, e1580, DOI: 10.1002/wat2.1580.

Jabbari A., Bae D.-H., 2020, Improving ensemble forecasting using total least squares and lead-time dependent bias correction, Atmosphere, 11 (3), DOI: 10.3390/atmos11030300.

Joseph S., Sahai A.K., Phani R., Mandal R., Dey A., Chattopadhyay R., Abhilash S., 2019, Skill evaluation of extended-range forecasts of rainfall and temperature over the meteorological subdivisions of India, Weather and Forecasting, 34 (1), 81-101, DOI: 10.1175/WAF-D-18-0055.1.

Kamblé K., 2020, Top 5: Biggest floods to affect India in 2019, The Weather Channel, available online https://weather.com/en-IN/india/news/news/2020-01-08-top-5-biggest-floods-affect-india-2019 (data access 28.02.2022).

Leander R., Buishand T.A., 2007, Resampling of regional climate model output for the simulation of extreme river flows, Journal of Hydrology, 332 (3-4), 487-496, DOI: 10.1016/j.jhydrol.2006.08.006.

Ming X., Liang Q., Xia X., Li D., Fowler H.J., 2020, Real-time flood forecasting based on a high-performance 2-D hydrodynamic model and numerical weather predictions, Water Resources Research, 56 (7), DOI: 10.1029/2019WR025583.

Pai D.S., Rajeevan M., Sreejith O.P., Mukhopadhyay B., Satbhai N.S., 2014, Development of a new high spatial resolution (0.25°×0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region, MAUSAM, 65 (1), DOI: 10.54302/mausam.v65i1.851.

Pattanaik D.R., Das A.K., 2015, Prospect of application of extended range forecast in water resource management: a case study over the Mahanadi River basin, Natural Hazards, 77, 575-595, DOI: 10.1007/s11069-015-1610-4.

Pattanaik D.R., Sahai A.K., Mandal R., Muralkrishna R.P., Dey A., Chattopadhyay R., Joseph S., Tiwari A.D., Mishra V., 2019, Evolution of operational extended range forecast system of IMD: Prospects of its applications in different sectors, MAUSAM, 70 (2), DOI: 10.54302/mausam.v70i2.170.

Sahai A.K., Chattopadhyay R., Joseph S., 2019a, Extended range forecast, Geography and You, 19, 16-21.

Sahai A.K., Chattopadhyay R., Joseph S., Krishna P.M., Pattanaik D.R., Abhilash S., 2019b, Chapter 20 – Seamless prediction of monsoon onset and active/break phases, [in:] Sub-Seasonal to Seasonal Prediction, A.W. Robertson, F. Vitart (eds.), Elsevier, 421-438, DOI: 10.1016/B978-0-12-811714-9.00020-6.

Sayama T., Yamada M., Sugawara Y., Yamazaki D., 2020, Ensemble flash flood predictions using a high-resolution nationwide distributed rainfall-runoff model: case study of the heavy rain event of July 2018 and Typhoon Hagibis in 2019, Progress in Earth Planetary Science, 7, 75, DOI: 10.1186/s40645-020-00391-7.

Shagun K., 2019, Indian rivers crossed highest flood level 25 times in August 2019, Down to Earth, available online https://www.downtoearth.org.in/news/natural-disasters/indian-rivers-crossed-highest-flood-level-25-times-in-august-2019-66818 (data access 28.02.2022).

Singh A., Sahoo R.K., Nair A., Mohanty U.C., Rai R.K., 2017, Assessing the performance of bias correction approaches for correcting monthly precipitation over India through coupled models, Meteorological Applications, 24 (3), 326-337, DOI: 10.1002/met.1627.
Teutschbein C., Seibert J., 2012, Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology, 456-457, 12-29, DOI: 10.1016/j.jhydrol.2012.05.052.

Webster P.J., Hoyos C., 2004, Prediction of monsoon rainfall and river discharge on 15-30-day time scales, Bulletin of the American Meteorological Society, 85 (11), 1745-1765, DOI: 10.1175/BAMS-85-11-1745.

Webster P.J., Jian J., Hopson T.M., Hoyos C.D., Agudelo P.A., Chang H., Curry J.A., Grossman R.L., Palmer T.N., Subbiah A.R., 2010, Extended-range probabilistic forecasts of Ganges and Brahmaputra floods in Bangladesh, Bulletin of the American Meteorological Society, 91 (11), 1493-1514, DOI: 10.1175/2010BAMS2911.1.

WMO, 2018, Guide to Climatological Practice, WMO-No. 100, World Meteorological Organization, Geneva, 139 pp.