Real-time performance of a multi-model ensemble-based extended range forecast system in predicting the 2014 monsoon season based on NCEP-CFSv2

A. K. Sahai*, R. Chattopadhyay, S. Joseph, R. Mandal, A. Dey, S. Abhilash, R. P. M. Krishna and N. Borah

Indian Institute of Tropical Meteorology, Pune 411 008, India

The real-time validation of any strategy to forecast the Indian summer monsoon rainfall requires comprehensive assessment of performance of the model on subseasonal scale. The multi-model ensemble (MME) approach based on the NCEP-CFS version 2 models, as developed and reported earlier, has been employed to forecast the 2014 monsoon season on the extended range scale with 3–4 pentad lead time (where a pentad corresponds to five-day average). The present study reports the broad performance of the MME employed on experimental basis to forecast the salient features of the real-time evolution of the 2014 monsoon season during June to September. The MME is successful in predicting both these features well in advance (3–4 pentad or 15–20 days lead time). The assessment of the model performance at pentad scale lead time shows that the weak monsoon conditions that are evident in precipitation and lower level wind anomalies are well captured as a whole up to four pentad advance lead time. The subseasonal propagation during onset and withdrawal is also evident in the forecast. Finally, the region-wise performance shows that the spatial extent of the skillful forecast encompasses central India as well as the monsoon zone for the 2014 monsoon season. Considering the natural variation in the forecast skill of extended range forecast itself as reported in earlier studies, the 2014 monsoon forecast seems to be skillful for operational purposes. For other regions (e.g. North East India), the forecast could be skillful at times, but it still requires further research on how to improve the same.

Keywords: Monsoon forecast, multi-model ensemble, pentad, lead time.

REAL-TIME extended range predictions of Indian summer monsoon intraseasonal oscillations (MISO) with 2–3 weeks lead time provide valuable outlook to the end-users and have high demand from agricultural stakeholders and dam managers dealing with flood forecasts. The development of extended prediction tool for intraseasonal oscillations in the tropics (MISOs and Madden–Julian oscillations) is a challenging problem and also a continuous process that is highly dependent on model bias. Several studies have adopted a number of strategies to achieve this target through inter-comparison of dynamical models based on diagnostics of operational forecasts of tropical intraseasonal oscillations[1–8].

Evaluation of the seasonal performance of monsoon forecasts as well as its failure as a whole is important to understand several peculiarities within a season, especially the catastrophic incidents[9,10]. Modelling and forecasting the monsoon variability with improved methods will ensure the development of an advanced strategy of adaptation for the stakeholders within a season[11]. The aim of this article is to highlight the performance of the experimental real-time extended range prediction of Indian summer monsoon for the 2014 monsoon season based on the strategy described in Abhilash *et al.*[12]. The 2014 monsoon season was a marginal drought year with rainfall recorded at 88% of long-period averages (refer later in the text). The scheme so developed employs a state-of-the-art multi-model and multi-ensemble framework of the CFSv2, the ocean–atmosphere coupled model from National Centre for Environmental Prediction (NCEP), USA and adopted at the Indian Institute of Tropical Meteorology, (IITM) Pune, as a part of the National Monsoon Mission Project of the Ministry of Earth Sciences, Government of India (http://www.tropmet.res.in/monsoon/).

The statistical performance of this multi-model and multi-ensemble version framework of NCEP–CFSv2 in the hindcast mode has been reported recently[12]. The phrase 'multi-model' here refers to (i) the same model run with two different spatial resolutions (T126 and T382), and (ii) the GFS model (essentially the atmospheric version of the CFSv2) run with sea surface temperature (SST) generated from the CFSv2 forecast runs with some bias correction[13]. The pros and cons of the individual components of the CFSv2 as referred above are described in several

*For correspondence. (e-mail: sahai@tropmet.res.in)

recent studies[12–14]. The phrase 'muti-ensemble' here refers to the ensemble of runs generated from the perturbed initial conditions (ICs) for each component model as described in the last few sentences. The method of generating the perturbed initial conditions has been discussed in earlier studies[7,13,15] and also briefly later in the text. This article deals with the extended range forecast (i.e. forecast with 2–3 weeks lead time in advance, hereafter referred to as the forecast) generated by this multi-model and multi-ensemble framework of CFSv2, and will be referred to as 'MME' (multi-model ensemble) runs of CFSv2.

The results of Abhilash *et al.*[12], clearly depict that it is useful to employ the multi-model and multi-ensemble version of the CFSv2 model in predicting the intraseasonal variability with a lead time of 2–3 weeks during the June–September monsoon season. This, however, is needed to be verified for case studies as the intrinsic interannual variation in forecast skill[16,17] could come across the operational implementation. The 2014 monsoon season shows several interesting features and in order to implement the real-time version of the MME to be suitable for operational agencies, it is required that it should be tested thoroughly. With this objective, the present article assesses the performance of the extended range prediction system of CFSv2-MME for the year 2014 in great detail. Especially, the focus will be on the critical assessment of the performance of CFSv2 forecast by highlighting different months. We will also describe onset, withdrawal and a few important spells.

The 2014 summer monsoon season

According to the 'end of season report' of India Meteorological Department (IMD), the official forecasting agency of India, the 2014 monsoon season may be characterized by the following seasonal as well as subseasonal features: (a) The seasonal quantum of rainfall is 777.5 mm, which is 12% below normal or long period average (LPA). (b) There was a hiatus in the progression of onset phase. (c) The 2014 monsoon season was a slow starter and the monthly rainfall over the country as a whole was 57% of LPA in June, impacting the sowing of rice over the northern rice belt. The rainfall recovered later and was 90% of LPA each in July and August. The relatively greater rainfall during these two monsoon months provided some relief to the agrarian community. Rainfall was 108% of LPA in September, which was important for planning urban hydrological roadmap until the next monsoon season. (d) The season was characterized by one cyclonic storm 'Nanauk', two monsoon depressions and 10 low pressures. (e) The intraseasonal (MISO) variability of rainfall was prominent during end of July to beginning of August associated with low pressure regions along the monsoon trough. Since the monsoon has recovered from early hiatus during 2014, it is necessary to perform the skill analysis of early months when the monsoon is weak to later months when monsoon is strong. In this article, the performance of MME based on CFSv2 will be organized in order to highlight this recovery in the latter months.

Data, method and forecast strategy

Data

The study uses the NCEP–NCAR reanalysis[18] for the year 2014 for comparing the dynamical fields (wind and temperature) and the IMD–TRMM merged data[19] for comparison of rainfall with model forecast. The long-term mean (climatology) is calculated based on 30 years (1981–2010) of NCEP data and 14 years (1998–2011) of rainfall data.

Model and forecast strategy

We have used the latest version of the ocean–atmosphere coupled model of NCEP, i.e. CFSv2. This has an atmospheric component (GFS) that is coupled to an ocean model, sea-ice model and land surface model. The ocean model is the GFDL Modular Ocean Model version 4p0d (MOM4)[21]. The initial conditions for the 2014 monsoon season are prepared from a coupled data assimilation system (CDAS) with T574L64 resolution atmospheric assimilation and MOM4-based oceanic assimilation, which is a real-time extension of the CFSR (ref. 22).

As discussed in Abhilash *et al.*[12], we have used a combination of three different versions of CFS for the forecast. The first version is developed by running the standalone GFS at T126 resolution, which has slightly different physics options (than the GFS that is coupled to CFS). The standalone GFS is forced with daily bias corrected and forecasted SST from the CFSv2 T126 run. Bias correction involves subtracting climatology of bias as a function of calendar day and lead time, with optimum interpolation sea surface temperature (OISST)[23] observations as the reference. We denote this two-tier forecast as GFSbc, with 'bc' indicating bias-corrected boundary conditions. For more details readers can refer to Abhilash *et al.*[13].

The other two versions of CFS run independently are the CFS T126 (~110 km) and the CFS T382 (~38 km) runs. They are so designed that the effect of high resolution (e.g. orography) and the large-scale processes are well represented in them. Further model and experimental details and skills of GFSbc, CFST126 and CFST382 are available in the literature[13–15].

Based on performance experience and aiming to maximize the operational skill depending on computer resources, we choose to pool these variants: 11 members of CFST126, 11 members of CFST382 and 21 members

of GFSbc. The model integrations to generate the ensemble members are initiated starting from 16 May and continued until 28 September at every five-day interval during the summer monsoon season. The forecast consensus is given by averaging the 43 ensemble members.

The performance of the model for the 2014 monsoon season is evaluated by averaging the forecast on a pentad (five-day mean) scale. The pentad-scale averaging is shown to give better statistics in the extended range as reported in earlier studies[7]. To elucidate if the forecast starts from 16 May IC, pentad 1 (P1) lead-time forecast corresponds to the forecast averaged over the period 17–21 May lead time; pentad 2 (P2) lead-time forecast corresponds to the forecast averaged over the period 22–26 May, and so on.

The forecasts are disseminated through deterministic as well as probabilistic statistics. Probabilistic forecasts from ensemble members at different lead times are validated with respect to predefined categories (quantiles) based on observations. The categories are defined by classifying the observed rainfall into above normal (AN), below normal (BN) and near normal (NN) using the tercile classification method. These categories correspond to the canonical 'active', 'break' and 'normal' spells over the Indian region. Tercile are the three intervals, i.e. the lower, middle and upper thirds of the climatologically distributed values of rainfall, where each category has an equal climatological probability of 33.33%. To determine the tercile ranges, the observed rainfall data were ranked in descending order. The three categories are defined as AN, NN or BN separated by the values 1/3 and 2/3 of the way down the ranked list.

Verification of forecast

In the present study we use a skill score method based on the definition of fractional skill score[24–26] to verify the performance of the model on different spatial scales at various lead times.

Let I_F (ix, iy, im, iyr, ip) be the binary value depending on yes-forecast (1) or no-forecast (0) of an event at a particular grid point (ix, iy) where $ix = 1, \ldots, Nx$ (total number of grids in longitude) and $iy = 1, \ldots, Ny$ (total number of grid points in latitude) for a member $im = 1, \ldots, Nm$ (total number of members in an ensemble forecast) for a year $iyr = 1, \ldots, Nyr$ (total number of years) and for a particular pentad $ip = 1, \ldots, Np$ (total number of forecasted pentads). In the present implementation $Nx = 26$ (65°–90°E), $Ny = 21$ (10°–30°N), $Nm = 43$, $Nyr = 10$ (for hindcast) and 1 (for forecast), and $Np = 24$ (all pentads during JJAS). Similarly, let I_O (ix, iy, iyr, ip) be the binary value of the observed occurrence yes (1) or no (0). Then the forecast fraction at a point is (ix, iy) of a year, iyr and a pentad ip for a given spatial neighborhood window (n) is given as

$$F_n(ix, iy, iyr, ip) = \frac{1}{n * n * nmem} \sum_{jj=iy-N}^{iy+N} \sum_{ii=ix-N}^{ix+N} \sum_{im=1}^{Nm} I_F(ii, jj, im, iyr, ip), \quad (1)$$

where $n = 2*N + 1$, ($N = 0, 1, 2, \ldots, 10$ in the present study). Similarly, observed fraction at that point is defined as

$$O_n(ix, iy, iyr, ip) = \frac{1}{n * n * nmem} \sum_{jj=iy-N}^{iy+N} \sum_{ii=ix-N}^{ix+N} I_O(ii, jj, iyr, ip). \quad (2)$$

Then the Fractions Brier Score (FBS) is defined as

$$\mathrm{FBS}(n) = \frac{1}{N_x * N_y * N_{yr} * N_p} \sum_{ix=1}^{N_x} \sum_{iy=1}^{N_y} \sum_{iyr=1}^{N_{yr}} \sum_{ip=1}^{N_p} (F_n(F_n(ix, iy, iyr, ip) - O_n(ix, iy, iyr, ip))^2. \quad (3)$$

$\mathrm{FBS}_{\mathrm{ref}}$ is defined as

$$\mathrm{FBS}_{\mathrm{ref}}(n) = \frac{1}{N_x * N_y * N_{yr} * N_p} \sum_{ix=1}^{N_x} \sum_{iy=1}^{N_y} \sum_{iyr=1}^{N_{yr}} \sum_{ip=1}^{N_p} ((F_n(ix, iy, iyr, ip))^2 + (O_n(ix, iy, iyr, ip))^2. \quad (4)$$

Finally FSS for a particular grid scale n is defined as

$$\mathrm{FSS}(n) = 1 - \mathrm{FBS}(n)/\mathrm{FBS}_{\mathrm{ref}}(n). \quad (5)$$

When the value of FSS is plotted against n, we can get the smallest scale in the value of n for which the FSS curve crosses the $\mathrm{FSS}_{\mathrm{target}}$ line. $\mathrm{FSS}_{\mathrm{target}}$ is defined as $\mathrm{FSS}_{\mathrm{target}} = 0.5 + f_0/2$, where f_0 is the base value (= 0.33 for tercile categories). $\mathrm{FSS}_{\mathrm{random}}$ is defined as f_0.

Annual cycle

Figure 1 shows the seasonal cycle of rainfall for the 2014 monsoon season. It is clear that the monsoon season

shows two strong active spells during mid-July and the beginning of September (bars in Figure 1). The forecasted seasonal cycles are also shown in the figure. The forecasted seasonal cycles at P1, P2, P3, P4 in advance lead time are shown in Figure 1 as lines. It is clear that the seasonal cycles, including the active spells are captured by the MME forecast. However, the amplitudes are underestimated by the MME. In observations, the JJAS mean rainfall (mm day^{-1}) is 7.27, whereas it is 6.41, 6.74, 6.9 and 6.65 respectively, for P1, P2, P3 and P4 forecasts. The variability however is less in P3 and P4 forecasts compared to P1 and P2 forecasts. This is due to the fact that at larger lead time, the spread among the members of MME increases leading to larger inter-member spread than at lower lead time. Larger spread is due to the larger growth of error at large lead times. The computed ensemble mean forecast (43 member mean) at these larger lead times is thus affected by the cancellation of the larger amplitude values (signal plus large error) of individual members (smoothing of random errors) for any forecast day leading to the clustering towards central (mean) value. All the forecasted days have equal probability of large error value at larger lead times. The daily ensemble mean forecast for all the days during the monsoon season at larger lead times thus has less (absolute) deviation about the season mean (expected) value compared to smaller lead times, which lead to smaller standard deviations.

Another important aspect of the annual cycle for the 2014 season is the abrupt increase in rainfall in July and decrease in rainfall in mid-August. The MME could predict these features in almost all pentad leads. This is reflected in the correlation of the seasonal cycle from observation and model forecasts at various lead times. The correlations between the observation and the forecasts at P1, P2, P3, P4 lead times are 0.72, 0.73, 0.70 and 0.71 respectively, indicating that the seasonal cycle of rainfall is captured with good fidelity. It is also noticed that there is some shift in the seasonal cycle. We believe that the problem in this case arises due to systematic bias in the model, where the MME gets more rainfall over the foothills (wet bias) and less rainfall (dry bias) over central India (Figure 2 of Abhilash *et al.*[12]). This can arise if the intraseasonal oscillations have different northward-propagating phase speed in model and observation, so that there is relative phase shift between the two.

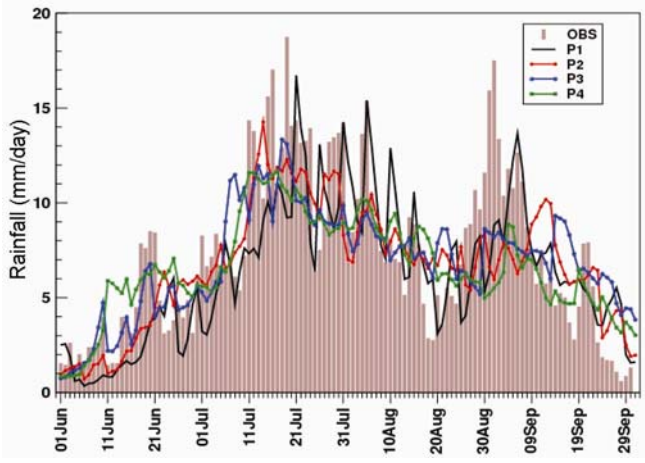

Figure 1. The seasonal cycle of rainfall (mm day^{-1}) for the 2014 monsoon season. The bars denote daily rainfall from IMD observation. The forecasted seasonal cycles at P1, P2, P3 and P4 lead times are shown as black, red, blue and green curves respectively. The correlation coefficients between the observations and the forecasts at P1, P2, P3, P4 lead times are 0.72, 0.73, 0.70, and 0.71 respectively.

Pentadwise seasonal performance

Figure 2 shows the spatial pattern of seasonal mean rainfall forecasted at P1, P2, P3 and P4 lead times. The JJAS (i.e. June to September) mean is shown in the left panels and the JJAS anomaly pattern is shown in the right panels. The anomaly as well as the mean patterns forecasted at different lead times indicate that over the Indian land mass, the MME performs with better fidelity in capturing the overall dryness (negative rainfall anomalies) in a consistent manner. Over the oceanic region, the negative

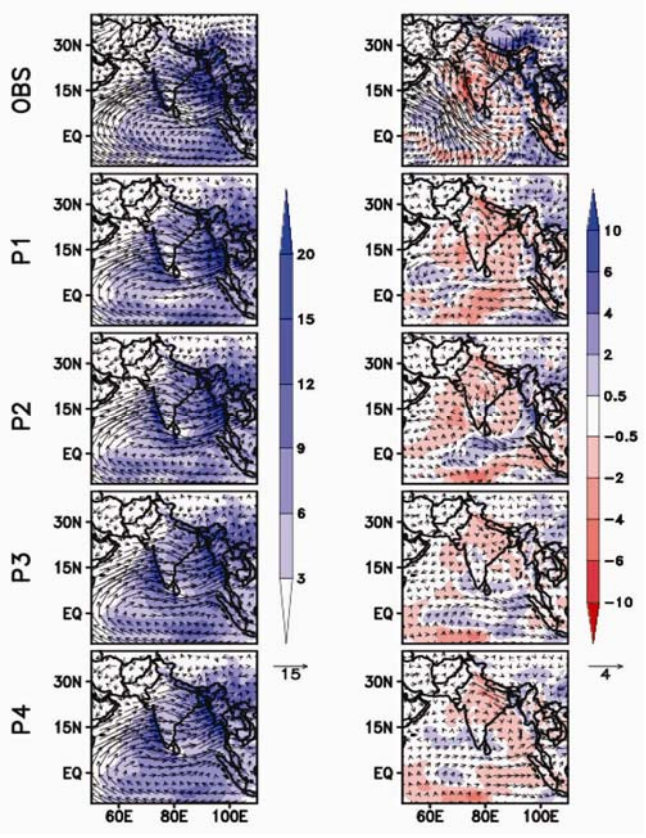

Figure 2. The June–September (JJAS) averaged rainfall and 850 hPa wind. (Top left panel) the actual observation. (Top right panel) Anomaly from JJAS climatology. The forecasts at P1, P2, P3 and P4 pentads are shown in subsequent panels. The left (right) column represents actual (anomalous) spatial patterns corresponding to P1–P4 pentad leads. (Units: rainfall, mm day^{-1}; winds, m s^{-1}.)

anomalies over this head Bay indicate dry bias prevailing at all lead-time forecasts. However, the positive rainfall anomalies associated with the Myanmar coast are seen to be well captured by the MME at all lead times. It could also predict the persistent low-level (850 hPa) northerlies over the central North Indian plain. Thus, both the rainfall and wind pattern at the lower level are well captured by the MME in terms of magnitude and seasonal mean spatial distribution.

Monthly performance of pentad-scale forecast

Figure 3 *a* shows the June rainfall anomaly from the reanalysis of 850 hPa wind and rainfall. The observed data show that June is characterized by a strong anticylonic anomalous wind pattern and is associated with strong negative rainfall anomalies all over the Indian region (top left panel) that contribute to –43% of the long-term mean rainfall during the month. It is interesting to note that the MME clearly predicts the large-scale dryness and anticyclonic persistent wind patterns at P1–P4 lead times.

Figure 3 *b* shows the observed July rainfall anomaly and the forecast performance at different lead times for the year 2014. As discussed earlier, the performance of monsoon has increased considerably in July, even though the monthly anomaly pattern shows drying over most parts of the country (top right panel). The head Bay is known for the formation of monsoon lows and depressions. The increased rainfall over the head Bay region was predicted by the MME to some extent, at almost all lead times. Although the overall dryness is evident in all the lead times, it is clear that the region-wise performances at advance lead times show bias. For example, the southern part of India as well as the west coast are shown to have positive rainfall biases at P3 and P4 lead times.

It is clear from Figure 3 that the large-scale drying observed in June 2014 had reduced to some extent in July due to the increased rainfall activity during the month (Figure 1). From agriculture point of view, it is important that this revival of monsoon from dry spell be pre-informed correctly. Figure 4 shows latitude–time evolution of rainfall averaged over 65°–95°E from 1 July in observations as well as the MME. It is clear from the figure that the first week of July witnessed an active spell after the onset of monsoon (top right panel). The forecasts for the same days with P1–P4 lead times are shown in subsequent panels (top to bottom). As expected, P1 shows greater skill in predicting the anomaly amplitudes as well as phase. However, the northward-propagating spell is well captured from P2–P4 lead-time forecasts also.

It has been shown by Abhilash *et al.*[12] that the probabilistic forecasts from the MME are superior to individual models participating in the MME. Therefore, we have analysed the skill of the IITM–MME in predicting the

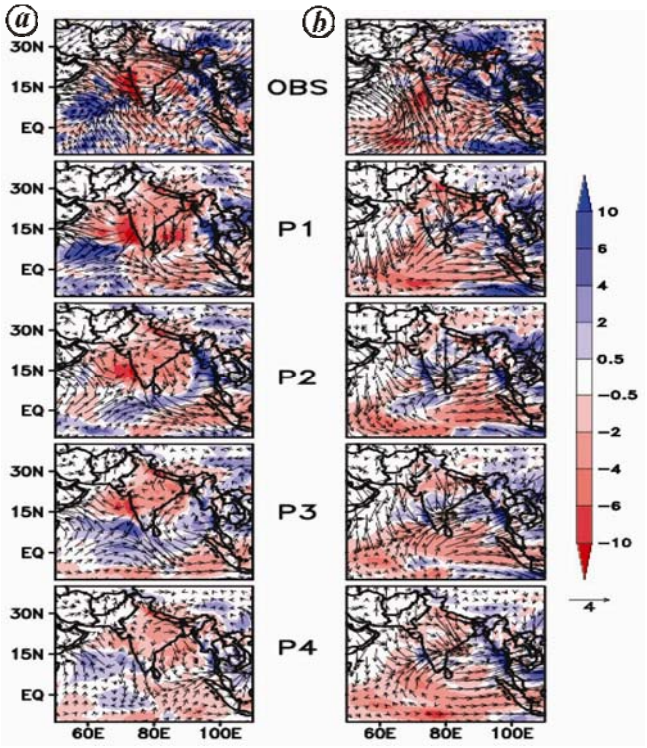

Figure 3. Same as Figure 2, but for the spatial patterns of monthly anomalies from observation (top panels) and forecasts at P1, P2, P3 and P4 lead times (subsequent panels) averaged during (***a***) June and (***b***) July. (Units: rainfall, mm day^{-1}; wind, m s^{-1}.)

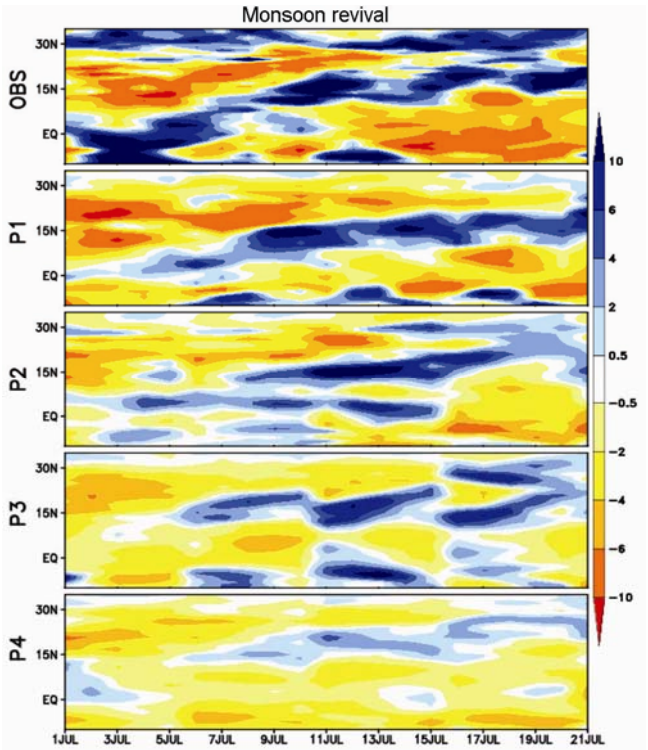

Figure 4. Plot showing the northward (Hovmöller) propagation of rainfall (mm day^{-1}) during the revival phase in July. (Top panel) observation. (Bottom panel) Model forecasts at P1–P4 lead times.

revival phase in July 2014. However, to develop confidence on the fidelity of the IITM–MME, it is also important to compare the performance of the same with the operational forecasts of other centres. The northward propagation of rain bands from IITM–MME forecasts during the revival phase has been compared with NCEP–CFS and Japan Meteorological Agency (JMA) operational forecasts[27,28]. The NCEP–CFS and JMA had some success in predicting the 2013 Uttarakahand extreme event[29]. Figure 5 shows the Hovmöller diagram of precipitation averaged over 65°–85°E. From the figure it is clear that the northward propagation features and its intensity are better captured in the IITM–MME system, although the general features are comparable with the NCEP–CFS forecast as expected. However, the precipitation bias over the equatorial Indian Ocean and Indian land mass is slightly reduced in the IITM–CFS forecasts compared to NCEP–CFS and JMA forecasts initiated around the same date (i.e. 25 June 2014). The northward extend of the precipitation and its general structure are well captured in the IITM–CFS compared to the other two forecasts. This shows the clear gain of the IITM multi-model CGMME system, and obviously the seminal role of the GFSbc model in improving the skill of the MME along with lower and higher resolution CFSv2.

The August forecast analysis shows that the rainfall in the west coast as well as in North East India is seen to be forecasted well even at the fourth pentad (P4) lead time (figure not shown). Also, the anomalous anticyclones over Central India are forecasted in a consistent manner (figure not shown).

Analysis of the forecast for the month of September indicates that the Kashmir region indicates positive rainfall anomalies in response to the heavy rainfall event that occurred in the first week of the month (figure not shown). However, the MME failed to predict the Kashmir flood event during early September 2014. The overall performance at P1–P4 lead times in forecasting the positive rainfall anomalies over the Central Indian region is satisfactory. However, the exact region of occurrence of maximum rainfall within Central India varies at different lead times (figure not shown).

Average performance over the monsoon zone and other homogeneous regions

The performance of the MME over the monsoon zone region as a whole and other homogenous regions as defined by IMD is an important metric of verification as the agricultural output, river-water availability for managing irrigation as well as flood management in this zone critically depend on monsoon rainfall over these regions.

Figure 6 shows the deterministic as well as probabilistic performance of the MME for the monsoon zone region[25].

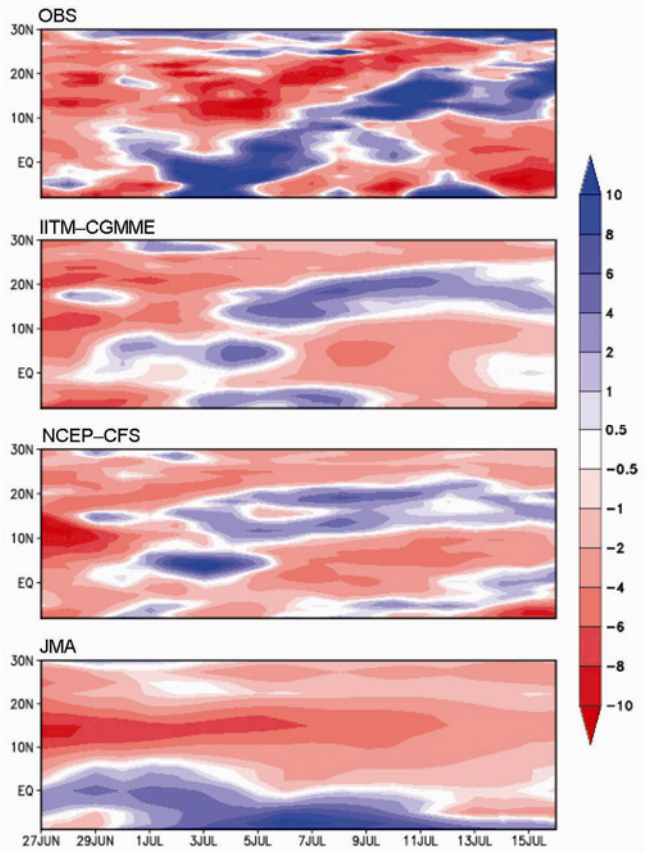

Figure 5. Comparison of the revival phase for different operational forecasts (IITM–CGMME, NCEP–CFS and JMA). (Top panel) Redrawn from Figure 4. The shading represents the rainfall anomalies (mm day^{-1}).

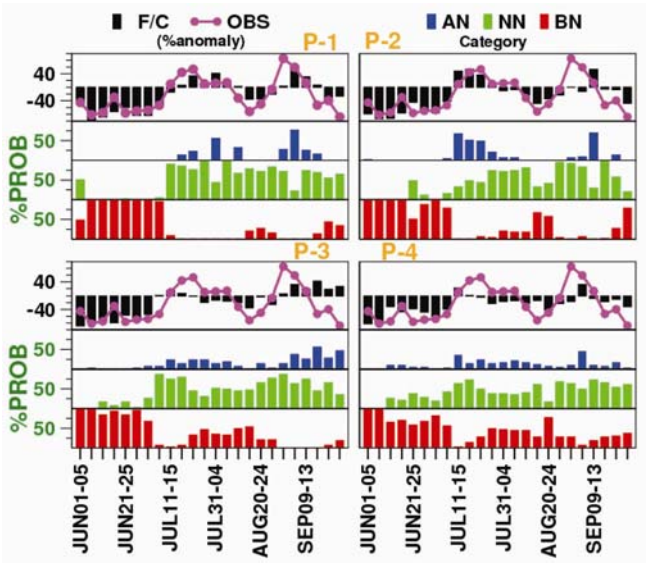

Figure 6. Area-averaged pentads mean rainfall percentage departure over monsoon zone (MZI, refer text) during JJAS 2014. The pink curve is for observation and the bars denote the forecasts. The lower three subpanels in each panels of P1–P4 are for above normal (AN), near normal (NN) and below normal (BN) forecasted categories.

The pink curve shows the observed rainfall variation area averaged over the monsoon zone[30], while the black bars in different panels show the forecast at different lead times (P1–P4) mentioned at the top of each panel. The noteworthy feature is that the extended dry spell during June is well predicted by the model even at the larger (P3 and P4) lead times. June is the crucial month for agricultural sowing activity over the Indian region as a whole. Thus, the skillful ~20 day (P4 lead time) forecast could be useful for agricultural planning as well as water management in dams and rivers.

The probabilistic forecasts of different rainfall categories such as above normal (AN, blue bars), below normal (BN, red bars) and near normal (NN, green bars), are shown in the subpanels below each panel of Figure 6 for each lead time. It is evident from the figure that the probabilistic forecasts are consistent with the deterministic forecast. The use of probabilistic forecast is worth mentioning for an active spell that occurred after 24 August (Figure 6). Although the deterministic forecast captured the spell for one or two pentads, the probabilistic forecast shows consistent in increase in blue bars up to third pentad lead time. Such systematic increase in the probability of concurrency of AN category would add quantitative value to the skill of deterministic forecasts. Table 1 shows the correlation skill (CC) and the root mean square error (RMSE, %) for the deterministic forecast at each lead-time. The forecast skill is lower in the fourth pentad (P4 lead time) as expected. However, the correlation and RMSE indicate that the skill is operationally useful for the stakeholders.

Table 1 also provides the skills for other homogenous regions for the year 2014. It may be seen that the CC and RMSE show useful predictability skill up to the fourth pentad in advance for the monsoon zone (MZI), northwest India (NWI) and Central India (CEI). The south peninsula India (SPI) and North East India (NEI) show less CC skill compared to other regions.

It is worth exploring whether the skill of the model is only for larger regions or it is in general for the grid scale also. We make an analysis of correlation skill from grid scale to gradually increasing to larger grid scale by averaging the smaller grid scale (Figure 7). For each grid resolution (1×1, 3×3, 5×5 and 9×9 degrees), the CC value obtained by comparing the observed and predicted rainfall for the whole season (24 pentads) has been plotted in Figure 7. It is evident from the figure that over most of the Indian region, the skill remains high from 1×1 to 3×3 to 5×5 to 9×9 grid scale, thus providing the operational confidence to the forecasters in real time. This skill is also evident in the hindcast statistics. This would be quantified more when we describe the fractional skill score later in the text.

Table 1. The root mean square error (RMSE, %) and correlation (CC) of pentad rainfall averaged over the different homogenous regions of India for the 2014 JJAS season

RMSE	MZI	NEI	NWI	CEI	SPI
P1	31.38	37.75	47.17	35.35	46.85
P2	34.09	40.05	41.20	37.66	55.45
P3	46.20	36.72	63.17	49.91	52.68
P4	37.83	29.91	46.59	41.30	55.55
CC	MZI	NEI	NWI	CEI	SPI
P1	0.82	0.31	0.69	0.81	0.65
P2	0.79	0.22	0.76	0.78	0.48
P3	0.57	0.21	0.43	0.56	0.52
P4	0.72	0.52	0.68	0.72	0.44

Spell-wise forecast performance

We now discuss the performance of the MME in getting the distinct phases of monsoon season of 2014.

Onset: The onset spell of monsoon bears important 'psychological' confidence-building information to the stakeholders. The delay in the onset spell has spiral effect

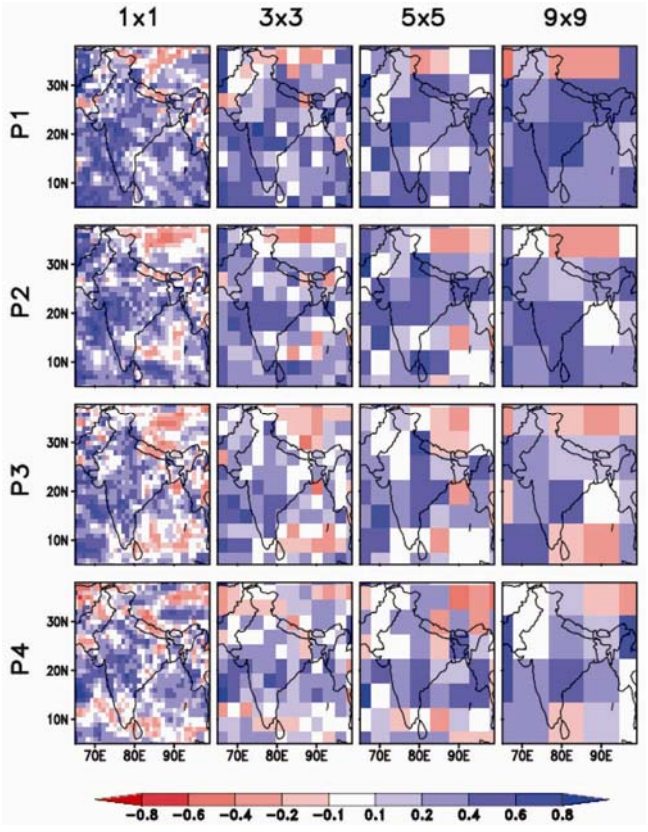

Figure 7. Spatial distribution of temporal correlation of rainfall on pentad scale computed for each grid during the 2014 monsoon season for different averaging of grid resolution. Each grid gives the correlation value for 24 pentads of model forecast and observation. The grid resolution sizes computed from the same data are shown at the top of each panel. The lead times (P1–P4) for the computations of correlation are shown on the left of each row.

in deciding the agricultural activity in the following season. Joseph *et al.* has defined a method to forecast the onset date from the MME using three indices defined from the large-scale wind circulation as well as rainfall, and the same methodology has been followed in the present study to forecast the onset. The rainfall index, ROK, is defined as the rainfall averaged over 74°–78°E; 8°–12°N; the wind index, UARAB, is defined as the zonal wind at 850 hPa averaged over 55°–75°E; 5°–12°N, while U_{depth} is defined as the zonal wind at 600 hPa over the same region for the UARAB index. Thirty-day mean values of forecasted ROK and UARAB (hereafter termed as ROKM and UARABM respectively) starting from 17 May are computed for each ensemble member and for each year. MOK for each member is then defined individually as the date on which both ROK and UARAB exceed 50% of their mean (ROKM and UARABM respectively), and one of them surmounts 70% of its mean, for five consecutive days, provided the value of Udepth during the period exceeds zero. The ensemble mean MOK date (of all 43 members) is treated as the final predicted MOK date. Figure 8 *a* depicts the evolution of ROK and UARAB indices from 16 May initial conditions. The predicted onset date (denoted by black circle in the figure), i.e. 5 June, is in good agreement with the actual onset date declared by IMD (6 June).

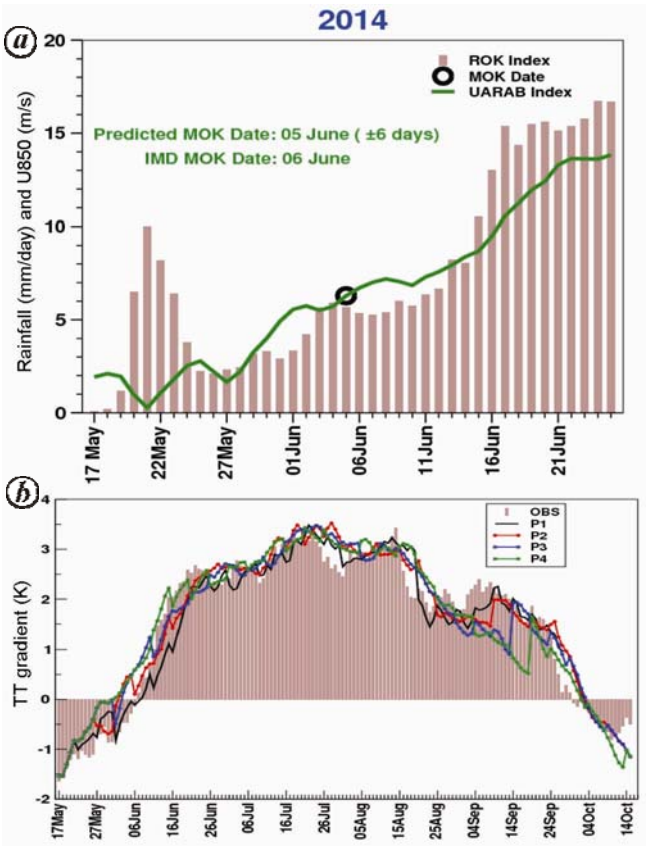

Figure 8. ***a***, Evolution of ROK and UARAB indices (see text for definition) during the monsoon onset phase. ***b***, The seasonal cycle of TTG index from observation (bars) and the forecasts at P1–P4 lead times. The correlations between the observation and the forecasted TT gradient during the 2014 monsoon season at P1, P2, P3, P4 lead times are 0.96, 0.96, 0.95 and 0.94 respectively.

The quantitative indicator of onset as well as the overall strength of the monsoon is the tropospheric temperature gradient (TTG) index as defined in Xavier *et al.*[31]. Figure 8 *b* shows the seasonal evolution of the TTG index and the forecasted evolution. Based on TTG, the onset is defined as the date when the TTG index (black curve) changes from negative to positive[26]. In 2014, this happened around 6 June (zero crossing date for the bars). It is clear from the plot that for the forecasted onset date, the P4 forecast shows earlier onset around 2 June, which is 3–4 days earlier compared to observations. As we come closer (P3, P2, P1), the onset is predicted closer to the observed value. Although a 3–4 day error is less than the standard deviation of onset forecast in observation (6–7 days), the skill is required to be enhanced further. The TT gradient cycle shows high correlation at all lead times when the seasonal cycle is compared with observation. The correlations between the observation and the forecasted TT gradient during the 2014 monsoon season at P1, P2, P3, P4 lead times are 0.96, 0.96, 0.95 and 0.94 respectively.

The apparent skill in the model forecast is further evident when we plot the depth of westerlies over the region 55°–80°E; equator 10°N from observation and from forecast (Figure 9). The increase in the depth of westerlies up to 600 hPa, the hallmark of monsoon circulation is forecasted between 6 and 11 June at all lead times. Although the strength of the westerlies is underestimated in the model forecast at all the lead times, the temporal accuracy (with a pattern CC of 0.93, 0.95, 0.93 and 0.94 for P1, P2, P3 and P4 respectively) would help in this case to accurately forecast the establishment of the monsoon onset over southern India and its subsequent progression.

The northward propagation of rainfall (averaged over 65°–95°E) during onset phase is shown in Figure 10, for observation as well as P1–P4 forecasts. The figure clearly delineates the persistent negative rainfall anomalies associated with the subdued monsoon activity over Central India during June. It is worthwhile to note that the MME could predict this dry onset phase with great fidelity in almost all pentad leads, which was very much beneficial in the planning of agricultural activities.

Withdrawal: In this section we discuss the forecast of several features associated with withdrawals. The important index for withdrawal is the TTG index as discussed earlier. Withdrawal is defined as the day when the TTG index changes its sign from positive to negative. As shown in Figure 8 *b*, it is clear that the forecast of TTG index was made with good fidelity at all pentad lead times. Thus, the indications of withdrawals are captured well in the MME framework.

Figure 11 shows the propagation of rainfall anomalies averaged over Indian longitudes (65°–95°E). In observation (top panel), consistent drying above 15°N is noticed from 9 September onwards. The equatorial convection persisted around this time. The same features are shown for the MME forecasts (P1–P4) in subsequent panels. It may be seen that the model forecast shows drying after 10 September in P1, P2 and P4 forecasts. Curiously, the third pentad forecast shows positive rainfall anomalies above 15°N at the same time. Thus, the withdrawal forecasts are not consistent at the third pentad lead time.

The skill of the 2014 forecast

The skill of the 2014 forecast is defined based on fractional skill score (FSS) metric as defined earlier. The skill metric is defined in order to get an objective idea on the performance of the model on various spatial scales and at various lead times (P1–P4). Figure 12 shows a plot of fractional skill score vs number of grids for which the forecast is skillful. The left panels shows the FSS skills based on 10 years (2001–2010) of hindcast during monsoon season and the right panels show the FSS skill of 2014 monsoon season. The hindcast skill gives an idea of climatolofical or expected skill score as a function of lead time and grid scale. This plot gives an indication of the minimum spatial scale (in terms of grid scales in which the forecast is compared with observation) for which the forecast is valid at each lead time and for each category of forecast BN, NN and AN). Thus, the probabilistic forecast is quantified category-wise for this year for operational purposes.

The random forecast skill is indicated by a 'random' line in each panel and the target skill for which the forecast at each lead time is valid is shown by the 'target' line. From Figure 12, it is clear that in the 10-year-average

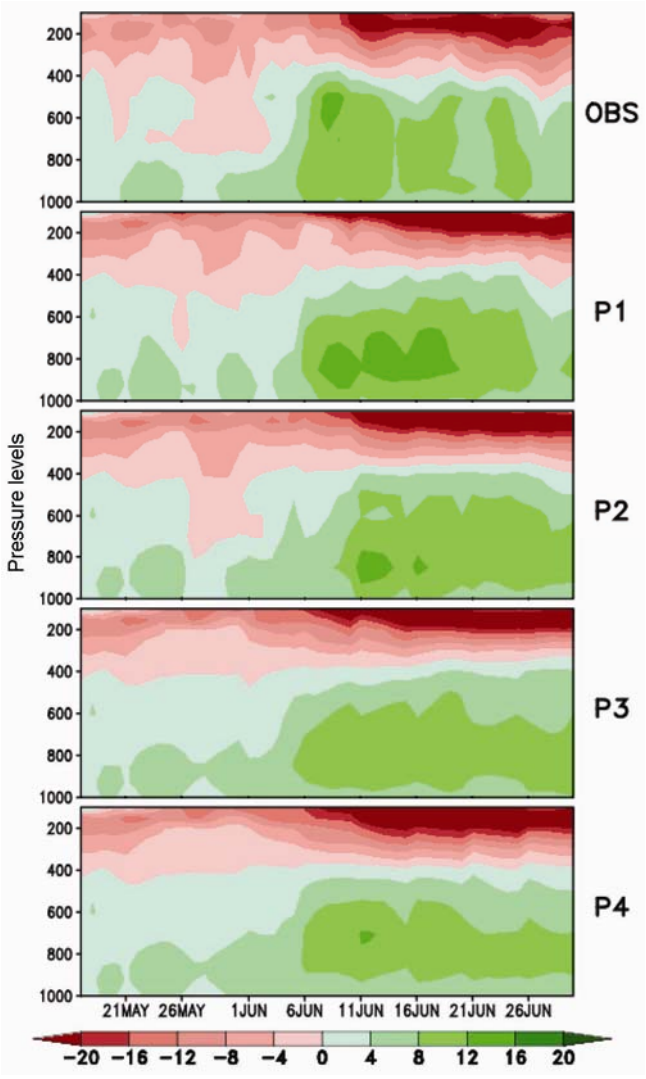

Figure 9. Time–height evolution of zonal wind (m s^{-1}) area-averaged over the region 55°–80°E; equator 10°N during onset phase. (Ordinate represents vertical pressure level in hPa). The pattern correlations between observations and pentads P1, P2, P3, P4 are 0.93, 0.95, 0.93 and 0.94 respectively.

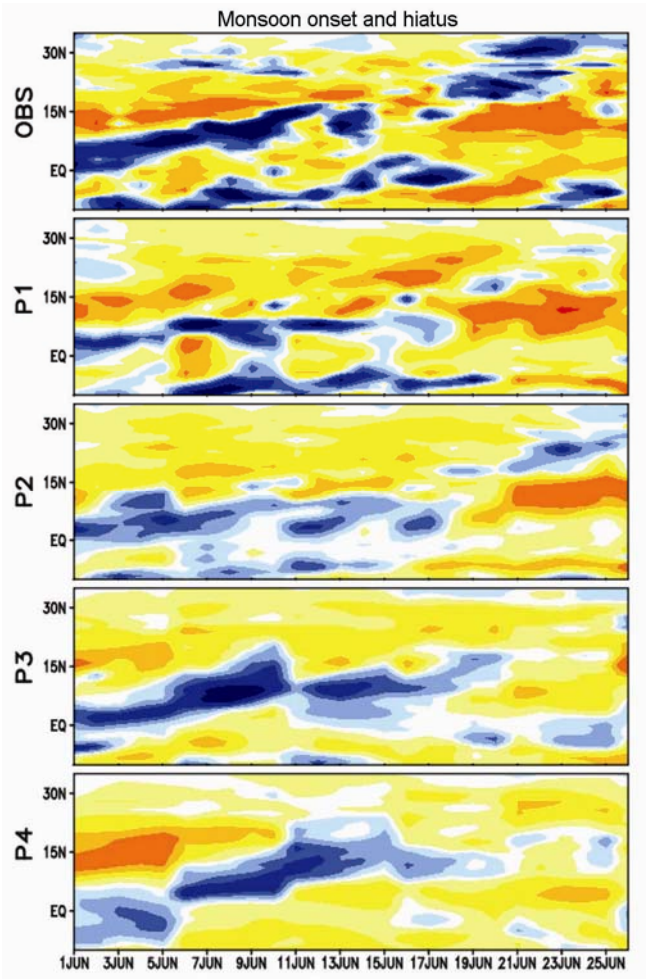

Figure 10. The Hovmöller plot of rainfall progression during onset phase of the 2014 monsoon season. The top panel is for observation and the subsequent panels are for the forecast at P1–P4 lead times respectively.

case (left panels), for BN and AN category, as we increase the lead time the spatial scale for which the forecast is skillful also increases. For example, for both the AN and BN categories, skillful forecast in high resolution (1 × 1 degree grid box) can be given with confidence only for P1 lead time, while the fourth pentad (P4) forecast can be delivered skilfully only in a 4 × 4 degree grid box. This is intuitively expected that as the uncertainty window increases, the lead time increases from the perspective of Lorenz theory of error growth. For the NN category, all scales converge to the same 2 × 2 degree grid box. Such scale separation in forecast skill indicates that the extended range forecasts are equally skillful when the MISO-induced phase is going towards AN and BN phases related to intraseasonal oscillations. However, since a priori it is impossible to know which IC run is going to be in the correct AN/BN phase, such situations simply indicate that there are days during the monsoon season when the forecasts are more skilful than the others.

For the year 2014, the situation is different with forecast at each lead time being equally skilful compared to the other. Thus, it indicates that for individual years, the skill of MME forecast could be different from the long-term expectation value. In 2014, a forecast of high skill is possible even at large lead time (P4). Thus, extended range forecast is reasonably skilful within a grid box of 2 × 2, 3 × 3 and 4 × 4 scales for BN, NN and AN categories. The interesting point here is that the *AN* spells could be equally well predictable for large lead time under certain circumstances using the MME.

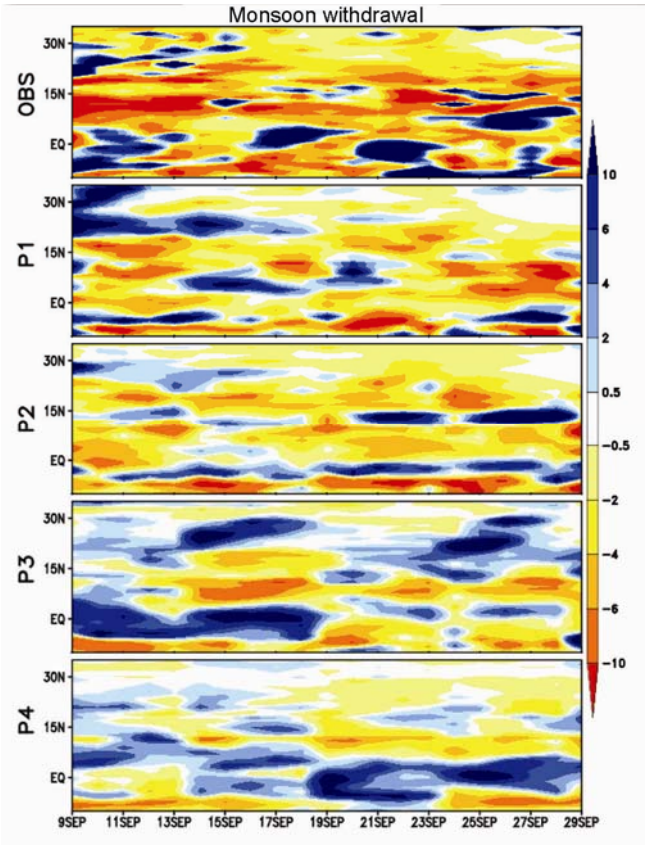

Figure 11. Same as Figure 10, but for monsoon withdrawal phase.

Discussion and conclusion

The overall performance of CFSv2 MME for the 2014 monsoon season is discussed in this article. The forecasts based on MME version could be useful to the operational forecasters for real-time extended range prediction. It is clear that the MME approach as proposed earlier is quite successful in terms of capturing the overall spatial pattern of rainfall during the 2014 monsoon season. The onset phase in June was weak and the monsoon revived during July. The MME technique was found to be useful in predicting these features with large lead time, than the individual set of model runs. The monthly analysis of forecasts made have clearly shows that even though regional details are required to be improved, the overall forecasting efficiency of large-scale patterns is encouraging with spells predicted well in advance up to the fourth pentad lead time. The onset, progression and withdrawal show consistent skill in forecast. Although it is noticed that the forecast at all lead times are not consistent among themselves, the overall performance of the MME for the 2014 monsoon season is encouraging. The basic improvement in the skill lies in predicting the large-scale northward propagating spells of rainfall associated with intraseasonal oscillations.

Also, we verify our forecast skill for category-wise forecasts using a new metric to verify the extended range forecast, e.g. fractional skill score. This score applied on the MME technique shows that practical skill may be achievable in extended range on the 4 × 4 degree grid scale (~400 km) at P4 pentad. Also, the AN spells could be predicted with equal fidelity using the MME for the year 2014.

The IITM CFSv2 MME is based on (i) an atmospheric model with bias corrected SST boundary conditions, a unique combination which no operational agency uses as of now; (ii) a combination of both high and low resolution forecasts thus correcting the orographic bias and (iii) effectively combining the coupled air–sea interaction from the CFSv2 coupled model forecast with atmospheric model component. The MME approach thus paved the way for the development of a synergistic approach for skilful forecast based on a 'same model different combination' strategy. This essentially preserves and highlights the original MME concept of Krishnamurti *et al.*[32] while adding the realistic effort of bias correction within the uniform model code framework. We believe that the present strategy makes the operational forecasting technically simplified as well as theoretically consistent as it eliminates the dependency on multiple models using

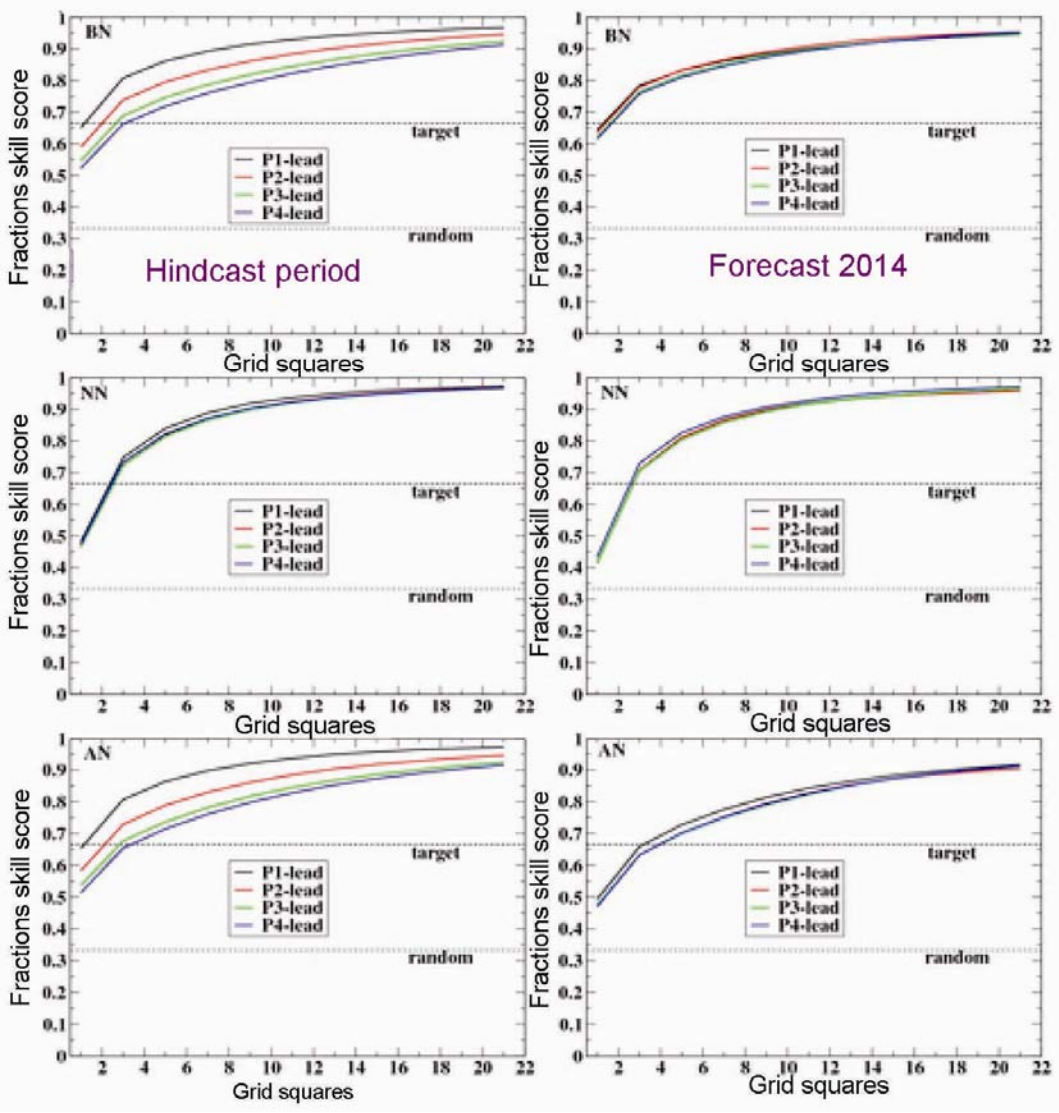

Figure 12. The fractional skill score of verification plotted for the 10 years (2001–2010) of hindcast (left column) for the below normal (BN; top panels), near normal (NN; middle panels) and above normal (AN; bottom panels) categories. (Right column) Same for the 2014 monsoon season.

statistical (dominantly regression based) bias correction. This article thus paves the way for real-time implementation of the IITM–MME approach for extended range prediction of the Indian summer monsoon on different spatial and temporal lead-time scales.

1. Vitart, F., Monthly forecasting at ECMWF. *Mon. Weather Rev.*, 2004, **132**, 2761–2779.
2. Vitart, F., Woolnough, S., Balmaseda, M. A. and Tompkins, A. M., Monthly forecast of the Madden–Julian oscillation using a coupled GCM. *Mon. Weather Rev*., 2007, **135**, 2700–2715.
3. Seo, K.-H. *et al.*, Evaluation of MJO forecast skill from several statistical and dynamical forecast models. *J. Climate,* 2009, **22**, 2372–2388.
4. Rashid, H. A., Hendon, H. H., Wheeler, M. C. and Alves, O., Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system. *Climate Dyn.*, 2010, **36**, 649–661.
5. Fu, X., Wang, B., Lee, J.-Y., Wang, W. and Gao, L., Sensitivity of dynamical intraseasonal prediction skills to different initial conditions. *Mon. Weather Rev*., 2011, **139**, 2572–2592.
6. Fu, X., Lee, J.-Y., Wang, B., Wang, W. and Vitart, F., Intraseasonal forecasting of the Asian summer monsoon in four operational and research models. *J. Climate,* 2012, **26**, 4186–4203.
7. Sahai, A. K. *et al.*, Simulation and extended range prediction of monsoon intraseasonal oscillations in NCEP CFS/GFS version 2 framework. *Curr. Sci*., 2013, **104**, 1394–1408.
8. Hudson, D., Marshall, A. G., Yin, Y., Alves, O. and Hendon, H. H., Improving intraseasonal prediction with a new ensemble generation strategy. *Mon. Weather Rev*., 2013, **141**, 4429–4449.
9. Gadgil, S., Srinivasan, J., Nanjundiah, R., Kumar, K. K., Munot, A. and Kumar, K. R., On forecasting the Indian summer monsoon: the intriguing season of 2002. *Curr. Sci*., 2002, **83**, 1307–1309.
10. Gadgil, S., Rajeevan, M. and Nanjundiah, R., Monsoon prediction – why yet another failure? *Curr. Sci*., 2005, **88**, 1389–1400.
11. Gadgil, S. and Srinivasan, J., Understanding and predicting the Indian summer monsoon. *Curr. Sci*., 2010, **99**, 1184–1186.

12. Abhilash, S. *et al.*, Improved spread-error relationship and probabilistic prediction from CFS based grand ensemble prediction system. *J. Appl. Meteor. Climatol.*, 2015, 1569–1578.
13. Abhilash, S. *et al.*, Does bias correction in the forecasted SST improve the extended range prediction skill of active-break spells of Indian summer monsoon rainfall? *Atmos. Sci. Lett.*, 2014, **15**, 114–119.
14. Sahai, A. K. *et al.*, High-resolution operational monsoon forecasts: an objective assessment. *Climate Dyn.*, 2014, 1–12.
15. Abhilash, S., Sahai, A. K., Pattnaik, S., Goswami, B. N. and Kumar, A., Extended range prediction of active–break spells of Indian summer monsoon rainfall using an ensemble prediction system in NCEP Climate Forecast System. *Int. J. Climatol.,* 2014, **34**, 98–113.
16. Borah, N., Sahai, A. K., Chattopadhyay, R., Joseph, S., Abhilash S. and Goswami, B. N., A self-organizing map-based ensemble forecast system for extended range prediction of active/break cycles of Indian summer monsoon. *J. Geophys. Res. Atmos.*, 2013, **118**, 9022–9034.
17. Palmer, T. N., Branković, Č., Molteni, F. and Tibaldi, S., Extended-range predictions with ecmwf models: interannual variability in operational model integrations. *Q.J.R. Meteorol. Soc.*, 1990, **116**, 799–834.
18. Kalnay, E., *et al.*, The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.*, 1996, **77**, 437–471.
19. Mitra, A. K., Bohra, A. K., Rajeevan, M. N. and Krishnamurti, T. N., Daily Indian precipitation analysis formed from a merge of rain–gauge data with the TRMM TMPA satellite-derived rainfall estimates. *J. Meteorol. Soc. Jpn Ser II*, 2009, **87A**, 265–279.
20. Saha, S. *et al.*, The NCEP Climate Forecast System Version 2. *J. Climate*, 2014, **27**, 2185–2208.
21. Griffies, S., Harrison, M., Pacanowski, R. and Rosati, A., A technical guide to MOM4. GFDL Ocean Group: NOAA GFDL, 2004; http://www.gfdl.noaa.gov/bibliography/related_files/smg0301.pdf?PHPSESSID=c94ddc382b93e57e39c8f976c83e970d.
22. Saha, S. *et al.*, The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, 2010, **91**, 1015–1057.
23. Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S. and Schlax, M. G., Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, 2007, **20**, 5473–5496.
24. Roberts, N., Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol. App.*, 2008, **15**, 163–169.
25. Roberts, N. M. and Lean, H. W., Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.*, 2008, **136**, 78–97.
26. Duc, L., Saito, K. and Seko, H., Spatial-temporal fractions verification for high-resolution ensemble forecasts. *Tellus A*, 2013, **65**, doi: 10.3402/tellusa.v65i0.18171.
27. Pattnaik, D., Kumar, A. and Tyagi, A., Development of empirical–dynamical hybrid forecasts for the Indian monsoon rainfall using the NCEP Climate Forecast System. IMD Monograph, 2010, 11/2010.
28. Pattanaik, D. R., Meteorological subdivisional-level extended range forecast over India during southwest monsoon 2012. *Meteorol. Atmos. Phys.*, 2014, **124**, 167–182.
29. Pattanaik, D. R., Pai, D. S. and Mukhopadhyay, B., Rapid northward progress of monsoon over India and associated heavy rainfall over Uttarakhand: a diagnostic study and real time extended range forecast. *Mausam*, 2015, **66**, 1–18.
30. Rajeevan, M., Gadgil, S. and Bhate, J., Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.*, 2010, **119**, 229–247.
31. Xavier, P. K., Marzin, C. and Goswami, B. N., An objective definition of the Indian summer monsoon season and a new perspective on the ENSO–monsoon relationship. *Q. J. R. Meteorol. Soc.*, 2007, **133**, 749–764.
32. Krishnamurti, T. N. *et al.*, Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, 2000, **13**, 4196–4216.

ACKNOWLEDGEMENTS. IITM, Pune is fully supported by the Ministry of Earth Sciences, Govt. India, New Delhi. The model runs have been carried out at the IBM Prithvi High Performance Computing System installed at IITM. We thank D. R. Pattnaik (IMD, New Delhi) for providing us the operational forecast data of JMA and NCEP(US)-CFS.

Received 5 May 2015; revised accepted 4 August 2015

doi: 10.18520/v109/i10/1802-1813