

Development of multimodel ensemble based district level medium range rainfall forecast system for Indian region

S K ROY BHOWMIK* and V R DURAI

India Meteorological Department, Mausam Bhavan, Lodi Road, New Delhi 110 003, India.

India Meteorological Department has implemented district level medium range rainfall forecast system applying multimodel ensemble technique, making use of model outputs of state-of-the-art global models from the five leading global NWP centres. The pre-assigned grid point weights on the basis of anomaly correlation coefficients (CC) between the observed values and forecast values are determined for each constituent model at the resolution of $0.25^{\circ} \times 0.25^{\circ}$ utilizing two season datasets (1 June–30 September, 2007 and 2008) and the multimodel ensemble forecasts (day-1 to day-5 forecasts) are generated at the same resolution on a real-time basis. The ensemble forecast fields are then used to prepare forecasts for each district, taking the average value of all grid points falling in a particular district. In this paper, we describe the development strategy of the technique and performance skill of the system during summer monsoon 2009. The study demonstrates the potential of the system for improving rainfall forecasts at five days time scale over Indian region. Districtwise performance of the ensemble rainfall forecast reveals that the technique, in general, is capable of providing reasonably good forecast skill over most states of the country, particularly over the states where the monsoon systems are more dominant.

1. Introduction

There has been long demand from the user community for district level quantitative weather forecasts in short to medium range time scale. The quantitative rainfall forecast for smaller spatial distribution such as district level over highly complex inhomogeneous region like India is a very challenging task. For the generation of district level quantitative rainfall forecasts, one has to depend on the forecasts from dynamical Numerical Weather Prediction (NWP) models. During the last two decades, NWP methods have acquired greater skills and are playing an increasingly important role in the operational weather forecasting. But rainfall prediction skill of NWP models is still not adequate to satisfactorily address detailed aspects of Indian summer monsoon. This is because of large spatial and

temporal variability of rainfall and some inherent limitations of NWP models. There are various factors like topography, prevailing synoptic situation and its interaction with mesoscale systems, lack of observations, etc., are some of the key factors which pose difficulties for numerical weather prediction of any region, and so Indian region is not an exception.

Considering the need of farming sector, India Meteorological Department (IMD) has upgraded the Agro-Meteorological Advisory Service from agro climate zone to district level. As a major step, IMD started issuing district level weather forecasts from 1 June 2008 for meteorological parameters such as rainfall, maximum and minimum temperature, relative humidity, surface wind and cloud octa up to 5 days in quantitative terms (Roy Bhowmik *et al* 2009). These forecasts are

Keywords. Multimodel ensemble; district level forecast; rainfall analysis; medium range quantitative forecast; Indian summer monsoon; rainfall prediction skill.

generated through multi-model ensemble (MME) system making use of model outputs of state-of-the-art three global models from the leading global NWP centres. These forecasts are made available on the national website of IMD (www.imd.gov.in). During summer monsoon 2009, the number of ensemble members is increased from three to five. In the present study, we describe the development strategy of the MME technique, used for high resolution rainfall forecasts over Indian region and demonstrate the prediction skill of the technique during summer monsoon 2009.

In our previous study (Roy Bhowmik and Durai 2008, 2010), performance skill of MME at 50 km horizontal resolution for district level short range rainfall forecasts during summer monsoon 2007 was demonstrated from the use of four coarser grid models namely (i) IMD limited area model at 75 km horizontal resolution, (ii) IMD MM5 at 45 km horizontal resolution, (iii) National Centre for Medium Range Weather Forecasting (NCMRWF) MM5 at 30 km resolution, and (iv) NCMRWF T-80 (grid space ~ 156 over the tropics). At 50 km resolution, MME could cover only 250 districts. The encouraging results of the study motivated authors for further research to increase the forecast period and model resolution using improved rainfall outputs of state-of-the-art high resolution global models from leading NWP centres to meet the operational requirement of farming community.

2. Formulation of the operational multimodel ensemble suite

The strategy for multimodel super ensemble (Krishnamurti *et al* 1999, 2000), involves two phases. In the first phase, known as training period, utilizes the direct model outputs and the corresponding observed fields to derive the statistics. The weight for each model is generated from the least square minimization of the difference between the analysis (observed field) and model output utilizing a training period of 120 days. In the work of Krishnamurti *et al* (1999, 2000), daily rainfall analysis generated on the basis of Tropical Rainfall Measurement Mission (TRMM) and Special Sensor Microwave Imager (SSM/I) dataset are used as observed fields. These data are also used by them for physical initialization of Florida State University (FSU) Global model. The second phase, called the forecast phase, utilizes the multimodel forecasts and aforementioned weights to obtain the final ensemble forecast.

In the present study, weight for each constituent model at each grid point is derived on the basis

of anomaly correlation coefficients (CC) between the observed values (analysis fields) and forecast values based on the training period datasets. For computation of anomaly CC, observed climatology (on the basis of gridded observed rainfall fields of 15 years) is used.

Five NWP models considered for this development work are: (i) National Centre for Medium Range Weather Forecasting (NCMRWF T-254), (ii) European Centre For Medium Range Weather Forecasting (ECMWF T-799), (iii) Japan Meteorological Agency (JMA T-959), (iv) United Kingdom Meteorological Office (UKMO), and (v) National Centre for Environmental Prediction Global Forecast System (NCEP GFS T-382). Under a joint collaborative research project, IMD has been receiving global model outputs (in the GRIB format) of ECMWF and JMA. The outputs (GRIB) of NCEP GFS are available freely from the Internet. The model outputs of these models are post-processed using GRIB decoder. IMD receives NCMRWF T-254 and UKMO model outputs online from NCMRWF, Noida.

As the model outputs available are at different resolutions, in the first step, model outputs of the constituent models are interpolated at the uniform grid resolution of $0.25^\circ \times 0.25^\circ$ lat./long. for the domain covering from lat. 0° – 40° N and long. 60° – 100° E. In the second step, the weight for each model at each grid is determined objectively by computing the anomaly correlation coefficient between the predicted rainfall and observed rainfall. High resolution (lat./long.) gridded data (Rajeevan *et al* 2006; Rajeevan and Bhat 2009) produced operationally at National Climate Data Centre of IMD Pune are used for the development and validation of forecasts. The gridded data with the use of high dense land rain gauge observations has the capability to capture large scale as well as mesoscale features of monsoon precipitation systems, and is superior to other rainfall products like, Climate Prediction Centre (CPC), Merged Analysis of Precipitation (CMAP), Global Precipitation Climate Product (GPCP) data to bring out characteristic features of Indian monsoon (Roy Bhowmik and Das 2007).

The weight ($W_{i,j,k}$) for each member model (k) at each grid (i, j) is obtained from the following equation:

$$W_{i,j,k} = \frac{C_{i,j,k}}{\sum_{k=1}^5 C_{i,j,k}}, \quad i = 1, 2, \dots, 161; \\ j = 1, 2, \dots, 161$$

$C_{i,j,k}$ is the correlation coefficient between rainfall analysis and forecast rainfall for the grid (i, j) of model (k). For computational consistency, $C_{i,j,k}$ is taken as 0.0001 in case $C_{i,j,k} \leq 0$.

In order to apply the technique, grid-point weights are computed at the resolution of $0.25^\circ \times 0.25^\circ$ utilizing dataset of two seasons (1 June–30 September, 2007 and 2008). The ensemble forecasts (day-1 to day-5 forecasts) are generated at the $0.25^\circ \times 0.25^\circ$ resolution. The ensemble forecast fields are then used to generate district level forecasts by taking average value of all grid points falling in a particular district.

Very recently, Mitra *et al* (2011) reported results of experimental MME forecasts at the horizontal resolution of $1^\circ \times 1^\circ$ for medium range forecast during summer monsoon 2008 and compared the results with simple ensemble mean and bias corrected ensemble mean. Models used as ensemble members for the experimental MME are: NCMRWF, NCEP, UKMO and JMA. For the experimental MME, the scheme used is very similar to that of Krishnamurti *et al* (2000) and Mishra and Krishnamurti (2007). The study (Mitra *et al* 2011) noted that the bias corrected ensemble mean produces the best result, compared to simple mean ensemble and MME over the Indian region. They concluded that up to day-3, the MME products were always better than individual member models.

As the MME technique, applied in this study is different from the method used by other authors (such as Krishnamurti *et al* 1999, 2000, etc.), the spatial distribution of model weights of each member model is discussed in section 3. Performance of MME forecasts with reference to each member model is discussed in section 4. An evaluation of performance statistics of MME based district level forecasts is made in section 5. Finally, conclusions and outlook for the future are discussed in section 6.

3. Analysis of model weights

Spatial distribution of weight coefficient for each model at day-1 to day-5 forecasts are presented in figures 1–5 respectively. At day-1 forecast, ECMWF shows higher weights (more than 0.3) over northwest and east central India, but lesser weight for the southern Peninsula. UKMO has higher weights, in general, well distributed all over the country. JMA has higher weights over north-east India, along the east coast and over some parts of central India. Both UKMO and JMA have the higher weights over south Peninsula. Along the Western Ghats, both ECMWF and JMA carry higher weight. NCEP model in general shows less weights all over the country. NCMRWF shows higher weights over a north–south oriented region starting from central India to north India. Similar distribution is noticed in the day-2 forecast also.

For the north-east India, JMA gets higher weights at day-1 to day-5 forecasts and ECMWF carries higher weight at day-2 to day-5 forecasts.

On the day-3 forecast, UKMO, ECMWF and JMA show increased weights over most parts of the country, whereas NCEP and NCMRWF indicate decrease in weight all over the country. In day-4 and day-5, UKMO shows considerable increase in weights and spread over most parts of the country, whereas NCEP shows considerable decrease in the weight as well as in the area coverage.

Similar features are reflected in country mean weight (figure not shown). Up to day-2 forecast, all models show equal weights. But after day-2, UKMO shows higher weights followed by ECMWF and JMA. NCEP gets the less weight followed by NCMRWF.

4. Demonstration of MME forecast skill

The country mean spatial CC is computed taking into account the CC between observed value and corresponding forecast value at each grid point over the land areas of India. Figure 6 presents an inter-comparison of country mean spatial CC of rainfall forecasts by MME, mean ensemble and individual models for the summer monsoon 2009. For the NCEP, CC ranges from 0.31 to 0.21 for day-1 to day-5 forecasts, for ECMWF it ranges between 0.35 and 0.27, for JMA between 0.40 and 0.28, for T254 between 0.35 and 0.25 and for UKMO between 0.40 and 0.32. For the MME, CC lies between 0.41 and 0.35. The results show that MME is superior to each member model at all the forecasts (day-1 to day-5). The result of mean ensemble is found to be closed to the MME results. Among the individual model, UKMO and JMA are found to be superior, followed by ECMWF, NCMRWF, JMA and NCEP.

Similar results are found for the summer monsoon 2010, as shown in figure 7. In the MME forecast, CC ranges from 0.47 at the day-1 forecast to 0.37 at the day-5 forecast. The curve of mean ensemble is close to the MME curve. Among the individual model, JMA is having highest CC for the forecast up to day-3 forecast and then decreases at the day-4 and day-5 forecasts. The spatial CC curves of UKMO and ECMWF are found to be very close. The performance of NCEP is found to be the worst. NCEP based GFS T-382 was made operational at IMD New Delhi during summer monsoon 2010. For the comparison purpose, results of IMD GFS (which is not a member of MME) is also included in figure 7. Both NCMRWF and IMD CC curves are found to be similar.

An inter-comparison of spatial distribution of cumulative rainfall of the season, based on these

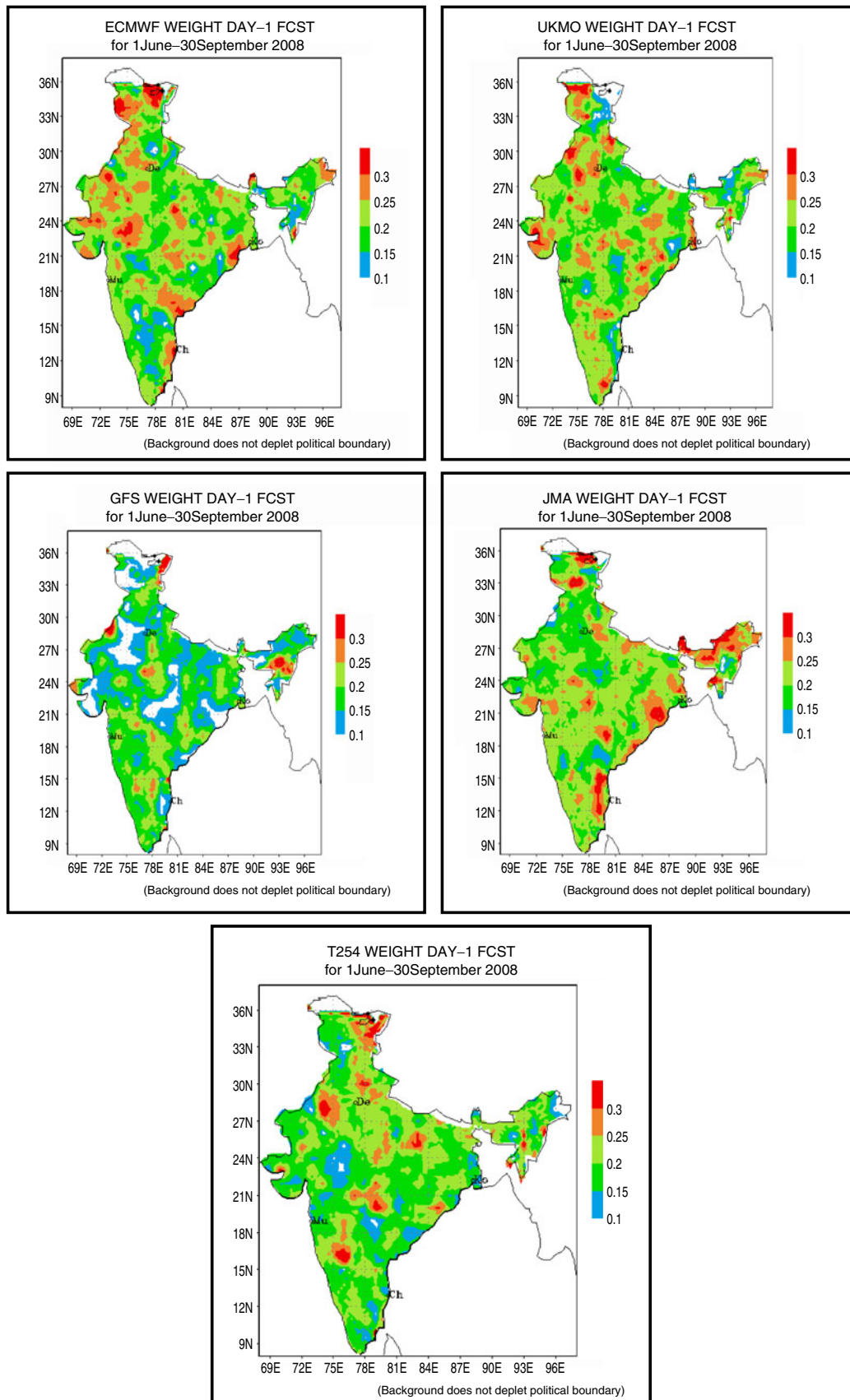


Figure 1. Spatial distribution of weight coefficient for each model at day-1 forecast.

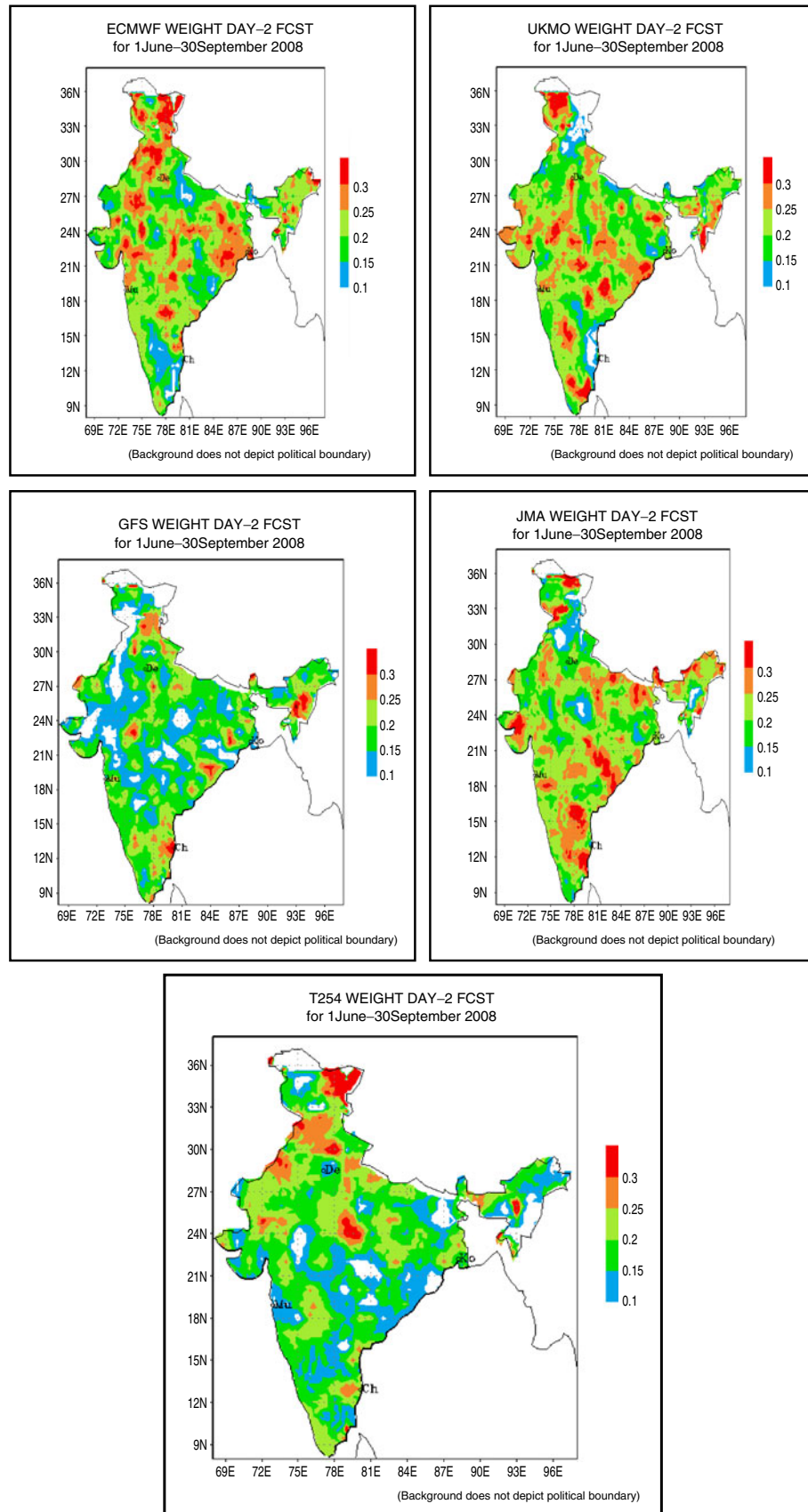


Figure 2. Same as figure 1 except for day-2 forecast.

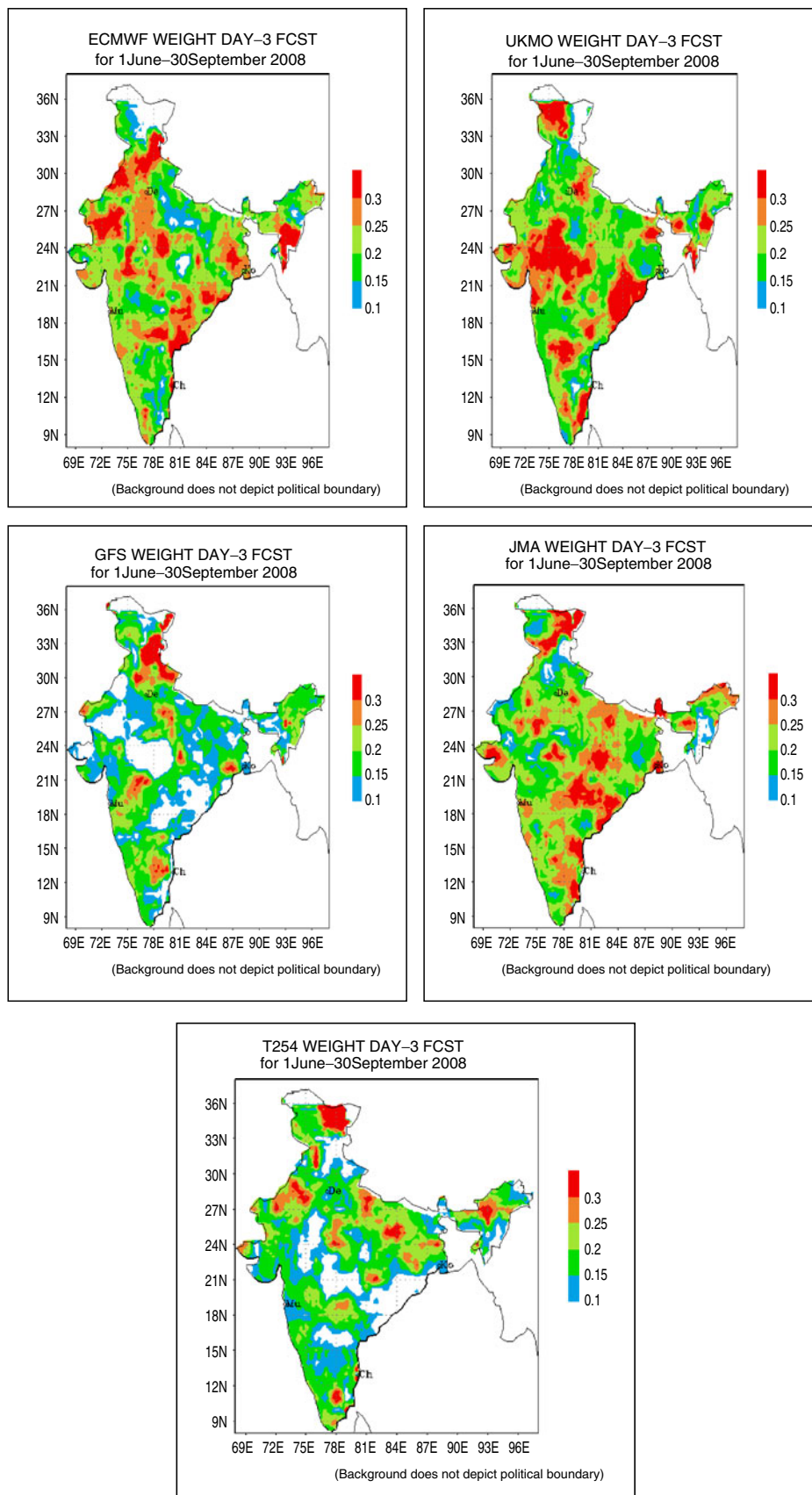


Figure 3. Same as figure 1 except for day-3 forecast.

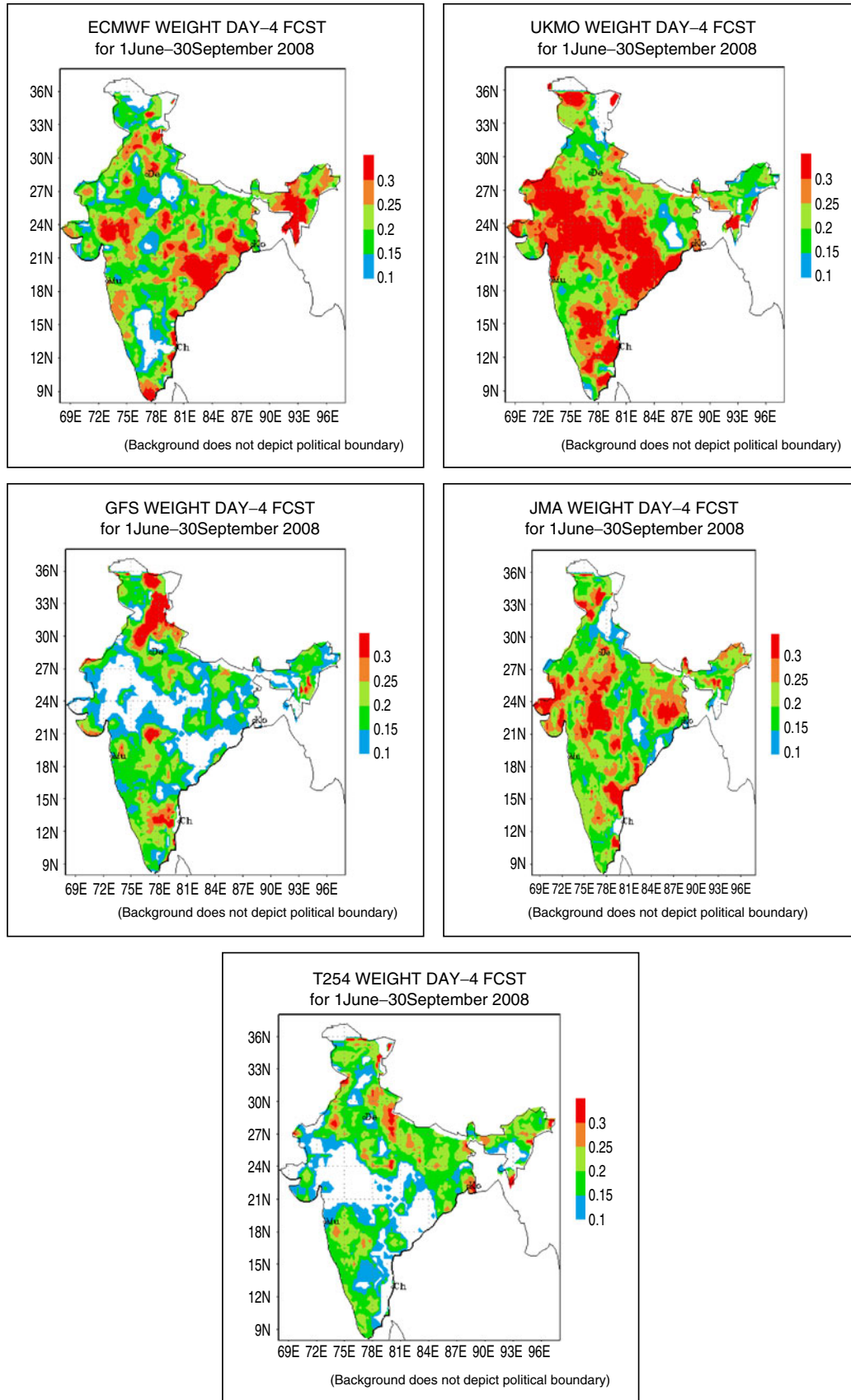


Figure 4. Same as figure 1 except for day-4 forecast.

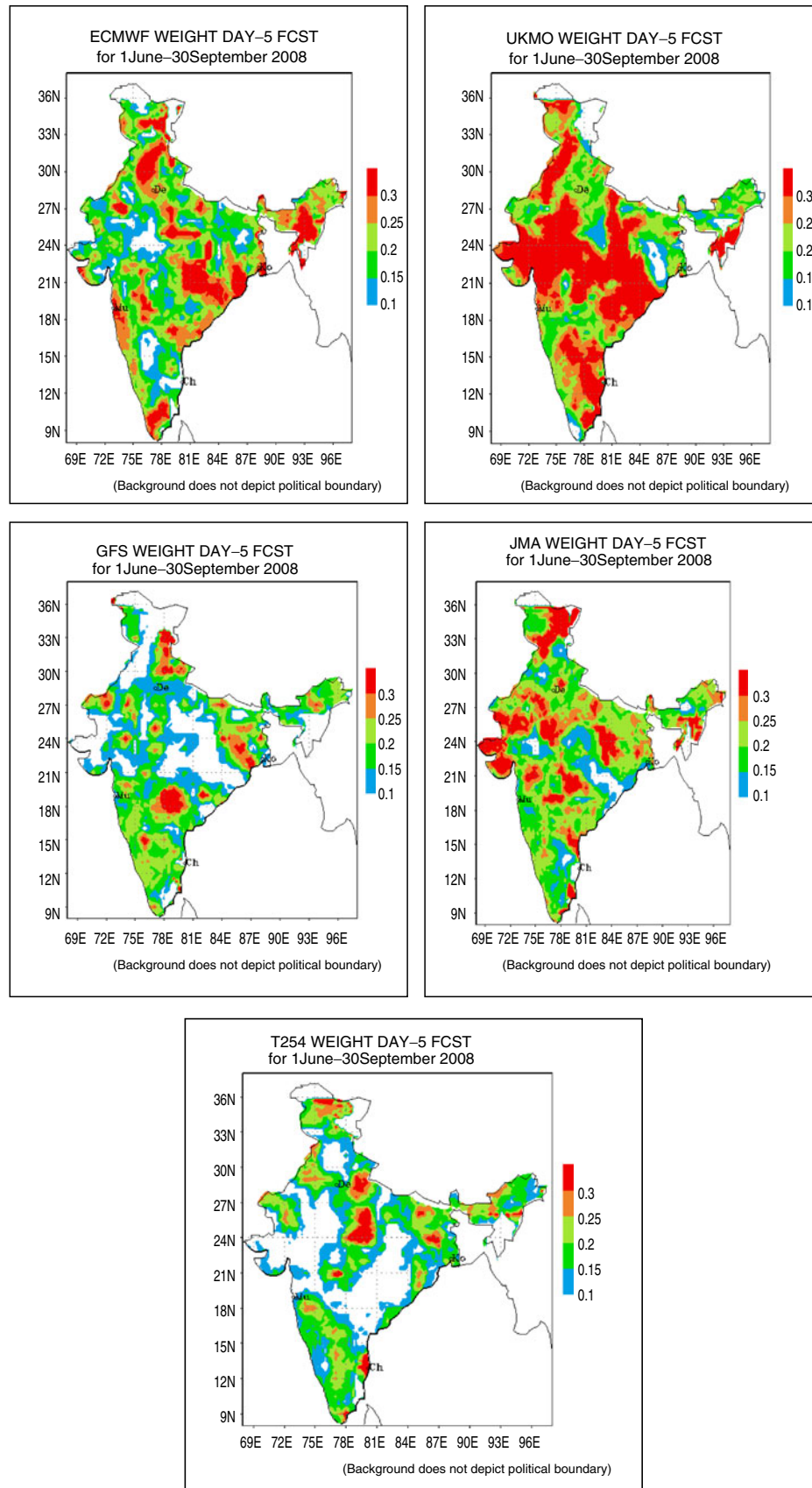


Figure 5. Same as figure 1 except for day-5 forecast.

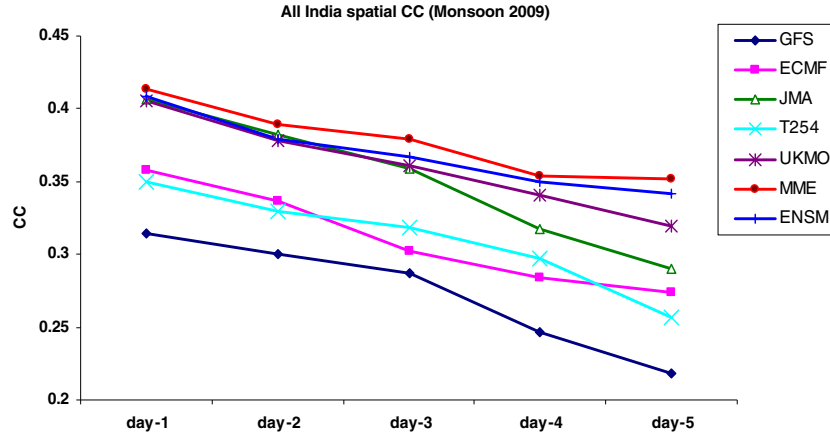


Figure 6. An inter-comparison of country mean spatial CC between the forecast and observed values for day-1 to day-5 forecasts of rainfall by NCEP, ECMWF, JMA, NCMRWF, UKMO, mean ensemble and MME for summer monsoon 2009.

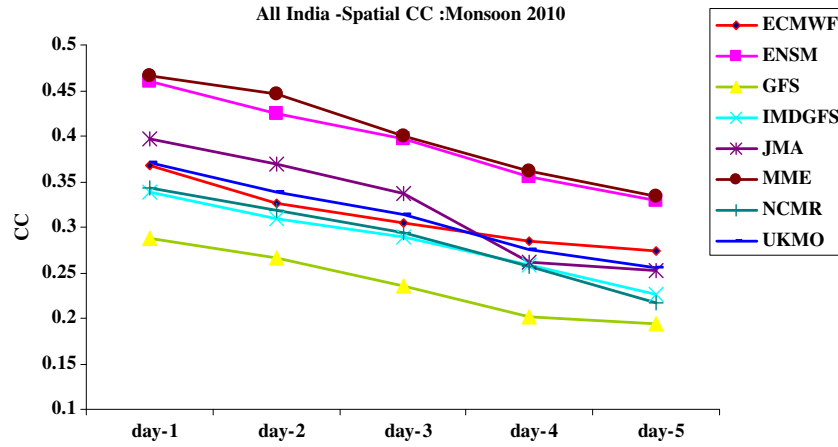


Figure 7. Same as figure 6, except for the summer monsoon 2010. A result of IMD GFS, which is not a member of ensemble is also included in the figure.

forecasts against the observations is also included in this study. The spatial distribution of forecast rainfall (figure not shown), in general, could reproduce the heavy rainfall belts along the west coast and over the domain of monsoon trough and foothills of the Himalaya. However, some spatial variations in magnitude are noticed. The UKMO shows location of monsoon trough more to the north. Except JMA, all member models considerably overpredicts rainfall amount along the foothills of Himalaya. As a result MME also shows higher rainfall along the foothills of Himalaya. However, the inter-comparison reveals that MME forecasts are closer to the corresponding observed field. The pattern remained broadly similar at all forecast days (day-1 to day-5).

In figure 8, the spatial distribution of root mean square errors (RMSE) of rainfall for day-5 forecasts by these models are presented. RMSE of the order

to 15–25 mm over the domain of Western Ghats, northeast India, monsoon trough has been produced by all these member models over the regions where occasionally monsoon heavy rainfall occurs. RMSE generally remained within 10 mm over the rest of the country. Magnitude as well as spatial coverage of errors are found to be considerably reduced in the MME forecasts. Similar features are noticed for the forecasts day-1 to day-5.

5. District level forecasts

In order to evaluate the performance of district level forecasts, skill score – Probability of Detection (POD) is considered. POD is defined as:

$$\text{POD} = \frac{H}{H + M}$$

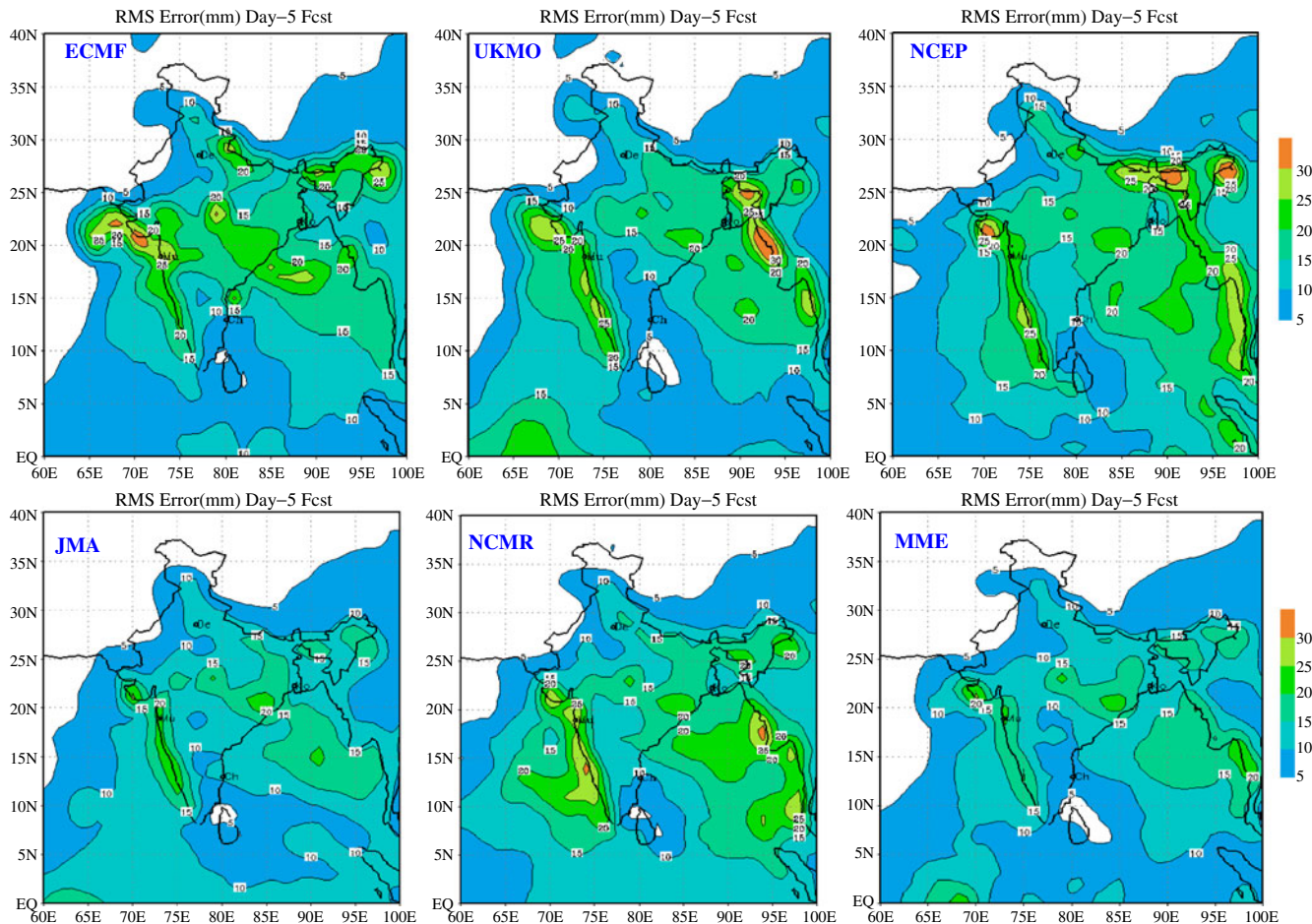


Figure 8. Spatial distribution of RMSE of rainfall (mm) based on day-5 forecast by NCEP, ECMWF, JMA, NCMRWF, UKMO and MME forecasts for the period from 1 June–30 September 2009.

where H indicates hits and M for missing events for the following rainfall categories:

- Rain or no rain
- Light rain: 0–10 mm
- Moderate rain: >10 and <65 mm
- Heavy rain: >65 mm

Statewise performance of district level rainfall forecasts for day-1 and day-5 forecasts for some selected states like Orissa, Rajasthan, Maharashtra, Gujarat and Kerala, which respectively represent east-central India: the domain of monsoon low; northwest India: region of less monsoon rainfall; west India: region of mid-tropospheric circulation and extreme southeast Peninsula are illustrated in figures 9 and 10. For the day-5 forecast, results of Madhya Pradesh (central India) is also included.

The results show that performance skill of day-1 forecast of the district level rainfall (figure 9) for the rainfall amount of moderate range is reasonably good for all these states, where POD is more than 0.4. For rain/no rain case, POD for these states are more than 0.6. For the heavy rainfall event, POD

significantly deteriorates. Among these states, Kerala has the best performance, followed by Orissa, Gujarat, Maharashtra and Rajasthan. POD of Kerala for moderate rainfall is close to 0.9. At the day-5 forecast (figure 10), performance deteriorates but pattern remains the same.

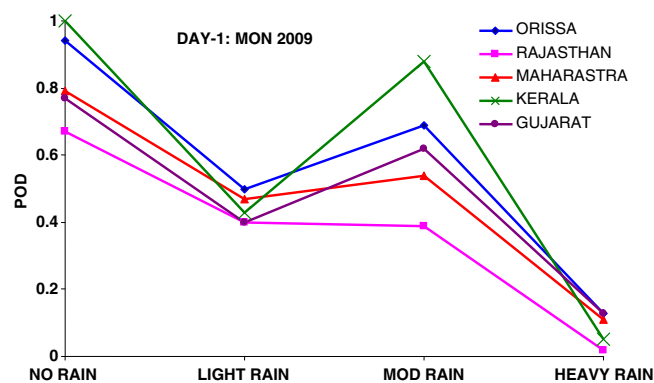


Figure 9. Statewise performance of district level day-1 forecasts for some selective states during summer monsoon 2009 (1 June–30 September).

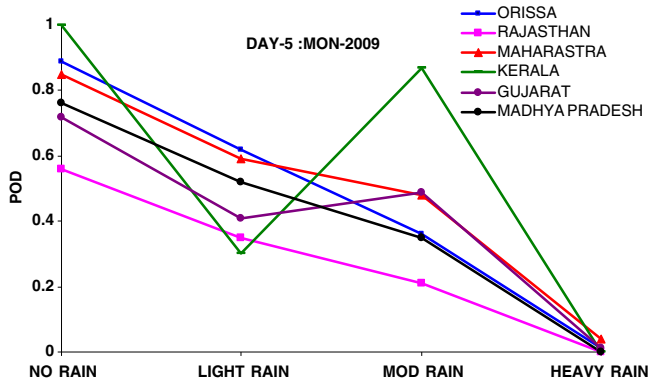


Figure 10. Statewise performance of district level day-5 forecasts for some selective states during summer monsoon 2009.

Districtwise performance of day-1 to day-5 rainfall forecasts for the districts of Orissa, Rajasthan and Kerala are presented in figure 11. For Orissa,

POD for rain/no rain case has been above 0.8 at all the districts, for light rain it is around 0.6, for moderate rainfall it is between 0.3 and 0.4 and for heavy rainfall it is close to zero. For Rajasthan, POD for rain/no rain case has been between 0.4 and 0.7 at most of the districts, for light rain it ranges between 0.3 and 0.5, for moderate rainfall it is between 0.1 and 0.3 and for heavy rainfall it is zero. For the state of Kerala, POD for rain/no rain case has been 1.0 at all the districts, for light rain it is around 0.3, for moderate rainfall it is around 0.9 and for heavy rainfall close to zero.

In order to examine the performance of the system in providing quantitative rainfall forecasts in a day-to-day manner, time series of daily day-5 rainfall forecast against the corresponding observed rainfall for the Bolangir district of Orissa during summer monsoon 2009 is illustrated in figure 12. The CC between the observed and forecast values is found to be 0.63. The result shows that the method

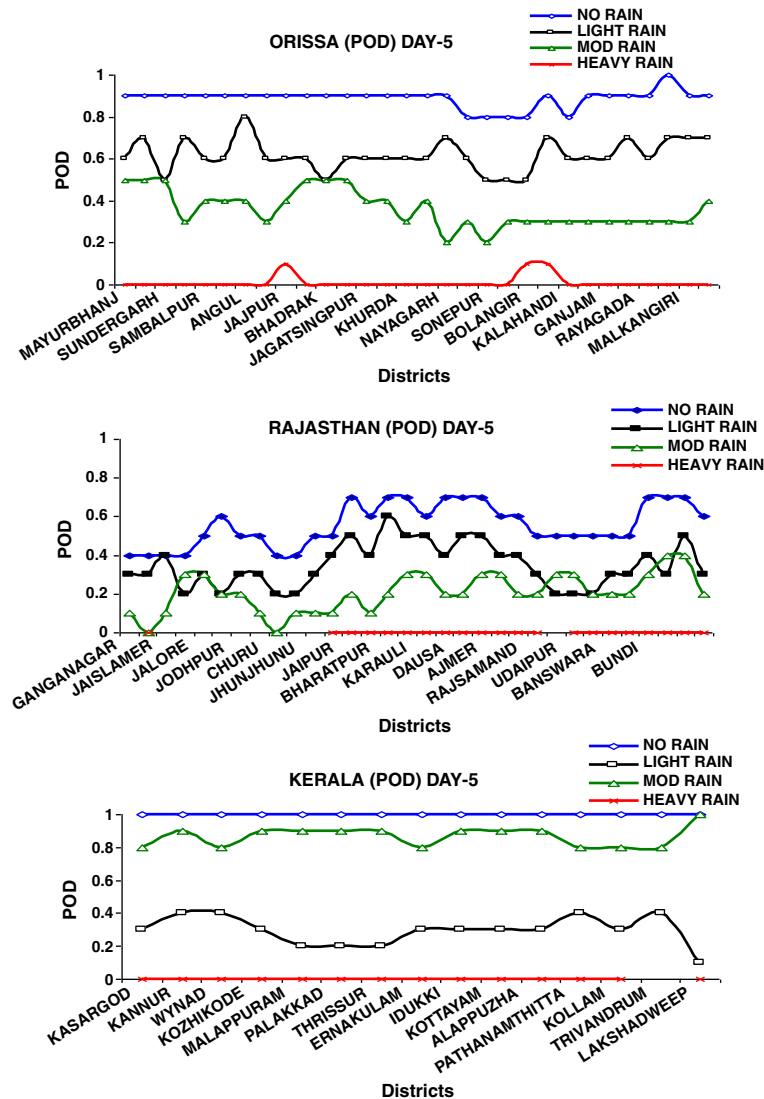


Figure 11. District level performance of day-5 forecasts for some selective states during summer monsoon 2009.

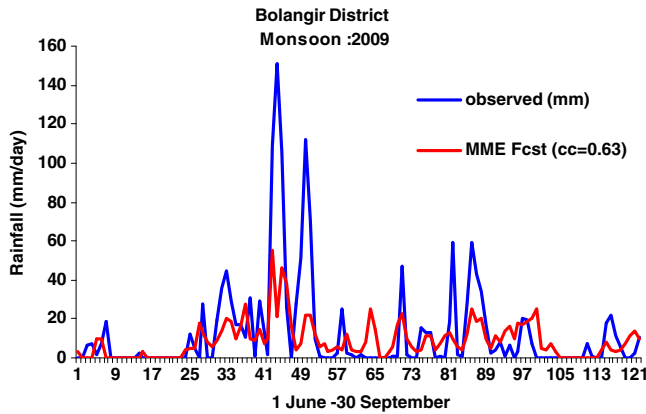


Figure 12. An example of time series of MME forecasts against the observations for the Bolangir district in Orissa during summer monsoon 2009.

is capable to capture daily ups and downs of rainfall spells. MME is able to capture light and no-rainfall days. However, for the heavy rainfall spells, though the model is able to capture the trends, but in most cases, magnitudes are underestimated. One reason for poor performance to predict heavy rainfall event is that in this study we considered the district level average rainfall (total rainfall divided by number of grids falling in a particular district) as the representative rainfall of a particular district. In reality, there may be a few isolated heavy rainfall events in a district, but when average is made, it may not be in the category of heavy rainfall. As such, heavy rainfall event (in this procedure) itself may be very less.

6. Concluding remarks and outlook for the future

One of the very specific requirement for the Integrated Agro-advisory Service of IMD has been to provide district level quantitative rainfall forecasts in the medium range time scale. For this purpose in this paper, we have documented the procedure adopted to generate MME based district level rainfall forecast and examined the performance skill of the real-time forecasts during summer monsoon 2009.

Results of the model weight analysis show that the UKMO has higher weights, in general, well distributed all over the country, except for south peninsula. In day-4 and day-5, UKMO shows considerable increase in the weights and spread over most parts of the country. JMA has the higher weights over south peninsula and over north northeast India. In general, UKMO shows highest weights followed by ECMWF and JMA. NCEP gets the less weight followed by NCMRWF.

Results of this study showed that all these models, in general, have the capability to capture large scale rainfall features of summer monsoon, such as heavy rainfall belt along the west coast, over the domain of monsoon trough and along the foothills of the Himalayas. It has clearly emerged from the results of the skill score that MME is superior to each member model.

For the district level forecast, the procedure has showed appreciable skill to predict occurrence and non-occurrence of rainfall, as well as for the rainfall category of moderate rainfall. But it fails to capture heavy rainfall events. Otherwise, performance of the district level forecast for most of the districts has been fairly good, particularly over the monsoon affected states.

One possible explanation for the poor performance to predict heavy rainfall event is that in this study we considered the district average rainfall (total rainfall divided by number of grids falling in a particular district) as the representative rainfall of a particular district. In reality, there may be a few isolated heavy rainfall events in a district, but when average is made, it may not be in the category of heavy rainfall. As a result, heavy rainfall event (in this procedure) itself may be very less.

MME based district level forecasts generated at the NWP division of IMD is disseminated daily to concerned meteorological centres spread all over the country. Value additions to the dynamical forecast is made by the concerned meteorological centres on the basis of synoptic, climatological and topographical knowledge of the district and from the consideration of the error structures of the previous day's forecast. The value added district level forecasts are the final forecast product provided for the Integrated Agro-advisory Service.

Though some significant improvement in accuracy and reliability of NWP product has been driven by adopting MME approach, however, limitations remain, particularly in the prediction of intensity and mesoscale rainfall features causing inland flooding. During recent years, Ensemble Prediction System (EPS) has emerged as a powerful tool for improving medium range weather forecasts. In the EPS, single model is used with multiple sets of initial conditions (Brooks and Doswell 1999) to obtain the final forecast. While Singular Vector and Bred Vector (BV) methods are still widely used in generating initial perturbations, Ensemble Transform of BV, Ensemble Transform Kalman Filter and Ensemble Data Assimilation are also implemented in various centres. Currently, 10 global centres operate EPS for medium range forecasts and they exchange model outputs at the native resolution among themselves. Very recently, with the commissioning of High Performance

Computing System, IMD has implemented NCEP based GFS T-382, with grid statistical interpolation as the global data assimilation. IMD has the plan to start an EPS in the near future and to become producer of EPS forecasts. This would allow IMD to get access to EPS outputs of other global centres and provide an opportunity for the development and implementation of a Probabilistic Forecast System (PFS) using EPS outputs of all available centres for better forecasting of monsoon heavy rainfall in short to medium range time scale. This would be one of our future priorities to improve medium range quantitative weather forecasts.

Acknowledgements

The authors are thankful to the Director General of Meteorology for the encouragement and providing all facilities to complete this work. We thankfully acknowledge the use of model outputs of NCM-RWF, UKMO, ECMWF, JMA and NCEP in this work.

References

- Brooks H E and Doswell C A 1999 New technology and numerical weather prediction; *Weather* **48** 173–177.
- Krishnamurti T N, Kishtawal C M, Larow T, Bachiochi D, Zhang Z, Willford E C, Gadgil S and Surendran S 1999 Improved weather and seasonal climate forecasts from multimodel super ensemble; *Science* **285** 1548–1550.
- Krishnamurti T N, Kishtawal C M, Shin D W and Willford E C 2000 Improving tropical precipitation forecasts from a multi analysis super ensemble; *J. Climate* **13** 4217–4227.
- Mishra A K and Krishnamurti T N 2007 Current status of multi-model super-ensemble operational NWP forecast of the Indian summer monsoon; *J. Earth Syst. Sci.* **116**(5) 1–16.
- Mitra A K, Iyengar G R, Durai V R, Sanjay J, Krishnamurti T N, Mishra A and Sikka D R 2011 Experimental real-time multi-model ensemble (MME) prediction of rainfall during monsoon 2008: Large scale medium range aspects; *J. Earth Syst. Sci.* **120**(1) 27–52.
- Rajeevan M, Bhate J, Kale J D and Lal B 2006 High resolution daily gridded rainfall data for Indian Region: Analysis of break and active monsoon spells; *Curr. Sci.* **91**(3) 296–305.
- Rajeevan M and Bhate J 2009 A high resolution daily gridded rainfall dataset (1971–2005) for meso-scale meteorological studies; *Curr. Sci.* **96**(4) 558–562.
- Roy Bhowmik S K and Das Ananda K 2007 Rainfall analysis for Indian monsoon region using the merged rain gauge observations and satellite estimates: Evaluation of monsoon rainfall features; *J. Earth Syst. Sci.* **116**(3) 187–198.
- Roy Bhowmik S K and Durai V R 2008 Multimodel ensemble forecasting of rainfall over Indian monsoon region; *Atmosfera* **21**(3) 225–239.
- Roy Bhowmik S K and Durai V R 2010 Application of multimodel ensemble techniques for real-time district level forecasts in short range time scale over Indian region; *Meteor. Atmos. Phys.* **106** 19–35.
- Roy Bhowmik S K, Durai V R, Das Ananda K and Mukhopadhyaya B 2009 Performance of IMD multi-model ensemble based district level forecast system during summer monsoon 2008; *Meteorological Monograph Synoptic Meteorology No. 8/2009, India Meteorological Department, New Delhi*, 43p.