Lead time for medium range prediction of the dry spell of monsoon using multi-models

A Jayakumar[2,*], Vinay Kumar[1] and T N Krishnamurti[1]

[1]Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL 32306, USA.
[2]Department of Meteorology, Florida State University, Tallahassee, FL 32306, USA.
*Corresponding author. e-mail: jkumar@fsu.edu

The purpose of this study is to address prediction of the start date and the duration of breaks in the summer monsoon rains using multi-model superensemble. The availability of datasets from the 'observing system research and predictability experiment (THORPEX)' initiated a forecast data archive, called THORPEX interactive grand global ensemble (TIGGE), makes it possible to use forecasts from a suite of individual ensemble prediction systems (member models) and to construct multi-model superensemble forecasts that are designed to remove the collective bias errors of the suite of models. Precipitation datasets are important for this study, we have used high resolution daily gridded rainfall dataset of India Meteorological Department (IMD), in addition to rainfall estimates from tropical rainfall microwave mission (TRMM) satellite and the CPC morphing technique (CMORPH). The scientific approach of this study entails the use of a multi-model superensemble for forecast and to verify against the rainfall information during a training phase, as well as during a forecast phase. We examine the results of forecasts out to day-10 and ask how well do forecast strings of day-1 through day-10 handle the prediction of the onset and duration of the breaks in the summer monsoon rains. Our results confirm that it is possible to predict the onset of a dry spell, around week in advance from the use of the multi-model superensemble and a suite of TIGGE models. We also examine trajectories of the parcels arriving in India in such forecasts from member models and from the multi-model superensemble to validate the arrival of descending dry desert air from the Arabian region during the dry spells and its mode of transition from wet spell. Some phenological features such as a shift in the latitude of the tropical easterly jet and changes in its intensity during break periods are additional observed features that are validated from the history of multi-model superensemble forecasts. Invariably this multi-model superensemble performs better than any single model in proving the better forecasts during our experiment period.

1. Introduction

Prediction of monsoon's wet and dry spell is very crucial to the Asian continents for the mitigation of disaster and agriculture managements. The dry spell of a monsoon here refers to the so called 'Break in the Indian monsoon' during a summer monsoon season, which last for periods of the order of a week or two (Ramamurthy 1969; Krishnan et al. 2000). Wang and Xu (1997) carried out an interesting study on the climatological dates of breaks of the Indian summer monsoon. Their analysis, based on calendar day composites, shows that the preferred dates of major breaks in the summer monsoon over India occur primarily at mid-July, secondarily at mid-September and during mid-October for the third time. Rajeevan et al. (2010) identified core monsoon zone in the central India

as a critical area for rainfall fluctuations between the active and dry spell for July and August months. The breaks we have studied here fit in these spatial and time frames of major break of climatological dates. The onset and duration of the dry spell is important in the perspective of both scientific as well as operational issue.

Major improvements in medium range numerical weather prediction have emerged from improvements in model resolution, data assimilation, physics and surface boundary conditions (Krishnamurti 2005). Multi-model ensembles have contributed to further growth in the skill of forecasts (Krishnamurti *et al.* 1999; Weigel *et al.* 2008). In recent years yet further improvement has come from the use of a multi-model superensemble (SE, e.g., Krishnamurti *et al.* 1999, 2000; Mishra and Krishnamurti 2007), and it runs through training and forecast phases. In the training phase, the recent past forecasts of member models of the multi-model suite are subjected to forecast validation and collectively all the forecasts are statistically regressed against the best observational estimates to derive statistical weights for the collective best forecasts. The forecasts from the multi-models during the forecast phase utilize the weights from the training phase to obtain a consensus, i.e., the SE forecast (Krishnamurti *et al.* 1999). The number of models, the number of grid locations in the horizontal and in the vertical, the number of variables and the number of member models lead to nearly 10^7 statistical weights, usage of these weights tends to improve the local forecasts in comparison with forecast from single deterministic member models (Krishnamurti *et al.* 1999).

The issue of predictability of dry spells has not been previously addressed in such a modeling context because the predictability of tropical rains is of the order of a few days (e.g., Waliser *et al.* 2003). For medium range numerical weather prediction, major improvements were noted in the reduction of RMS errors, spatial correlations and the equitable threat scores from the construction of the SE (Mishra and Krishnamurti 2007). The monsoon applications for numerical weather prediction were recently addressed by Krishnamurti *et al.* (2009), who noted medium range forecast improvements of precipitation for timescale of roughly five days. These promising studies lead us to ask for more specific phenomenological applications for the monsoon, such as the occurrences of dry spells of the monsoon and the lead time needed for such forecasts.

Indian subcontinent recently received severe dry spells in the year 2009 and recorded a deficit of 22% of its long term mean. Many studies have investigated the underlying dynamics for this deficit (e.g., Krishnamurti *et al.* 2010; Neena *et al.* 2011 and

references therein). Krishnamurti *et al.* (2010) noted that these dry spells of the monsoon bring in desert air incursions from the northwest at all vertical levels between 850 and 400 hPa. That feature is absent during the wet spells of the monsoon. That study also alludes to several antecedents for these dry spells that can be seen west and northwest of India. These include a blocking high over Arabia and various dynamical instabilities that precede those events over the southern Mediterranean. This made us to do much of the forecasting framework analysis focused more on the dry spell of 2009 in the current study.

The observing system research and predictability experiment (THORPEX) initiated a forecast data archive called THORPEX interactive grand global ensemble (TIGGE, Bougeault *et al.* 2010). This data archive provides ensemble forecasts from major forecast centers around the world. Several groups are currently active in research using multi-model ensemble forecast techniques based on the TIGGE data for medium range prediction (e.g., Matsueda and Endo 2011; Zhi *et al.* 2012). Therefore, we have put the focus of this study on the predictability of the dry spells using TIGGE models and the performance of the SE over the Indian domain. We also provide a comparative analysis on the characteristics of wind fields associated with the monsoonal dry spell using same model framework.

2. Datasets and methodology

2.1 *Observed datasets*

We have mainly used NOAA CPC Morphing Technique (CMORPH) Precipitation Product for representing monsoon spells over the Indian domain. This precipitation product combines data from different passive microwave radiometers aboard the US Airforce satellites DMSP 13, 14 & 15 (SSM/I), the NOAA satellites NOAA-15, 16, 17 and 18 (AMSU-B) as well as AMSR-E and TMI aboard NASA's Aqua satellites using different NASA algorithms (Kummerow *et al.* 1996; Ferraro 1997; Hong *et al.* 1999). The algorithm makes use of cloud motion features, at frequent intervals such as every three hours, from geostationary satellites. The rain rate information from TRMM and DMSP satellites are thus morphed (i.e., interpolated at higher time intervals) with this additional information to provide a space time continuity (Joyce *et al.* 2004). This global precipitation analyses combine the superior retrieval accuracy of passive microwave estimates and the higher temporal and spatial resolutions of the available IR datasets. This dataset covers a period from December 2002 to the present.

The current work utilizes daily values that are derived from the 3-hourly CMORPH estimates at a spatial resolution of 0.25° covering the 2008–2011 periods. In addition to the CMORPH, we have also made use of tropical rainfall measuring mission (TRMM) 3B42 datasets for model validation. The 3B42 is based on the retrieval technique by Huffman *et al.* (1995). This global data is available at the spatial resolution of 0.25° for every 3-hour from 1998 to present over tropical regions.

A comparison of July–August climatology from both TRMM and CMORPH are shown in figure 1(a and b). A high coherence in the spatial variability of monsoon rainfall is noted in both of these datasets, these carry a spatial correlation (SC) of ∼0.8 over the Indian domain. There is, however, a considerable difference over western Indian hills, called the Western Ghats. This difference may be due to the distinct handling of deep convective precipitation estimates for CMORPH and TRMM, as it was noted for the precipitation rates over mountainous terrain (Xie *et al.* 2007; Gopalan *et al.* 2010). Average rainfall amount over the box (figure 1, 70°–85°E; 10°–25°N), shows a good phase agreement, with some differences in the peak values during the convective phase, as compared to dry phase (figure 1c). Rainfall amounts averaged over this box is considered as the monsoon core region (MCR) in the present work, and rainfall estimated from aforementioned rainfall products are generally consistent with rain guage based gridded product from India Meteorological Department (IMD, Rajeevan *et al.* 2006). Estimation of rainfall amount over India per day carry the daily average from as many as 2300 rain gauges. Temporal Correlation Coefficient (CC) with IMD and TRMM shows values around 0.91, whereas with CMORPH carries a value of 0.88, but the higher variance in the TRMM in the wet spell gives RMSE of 3.07 mm/day whereas the CMORPH
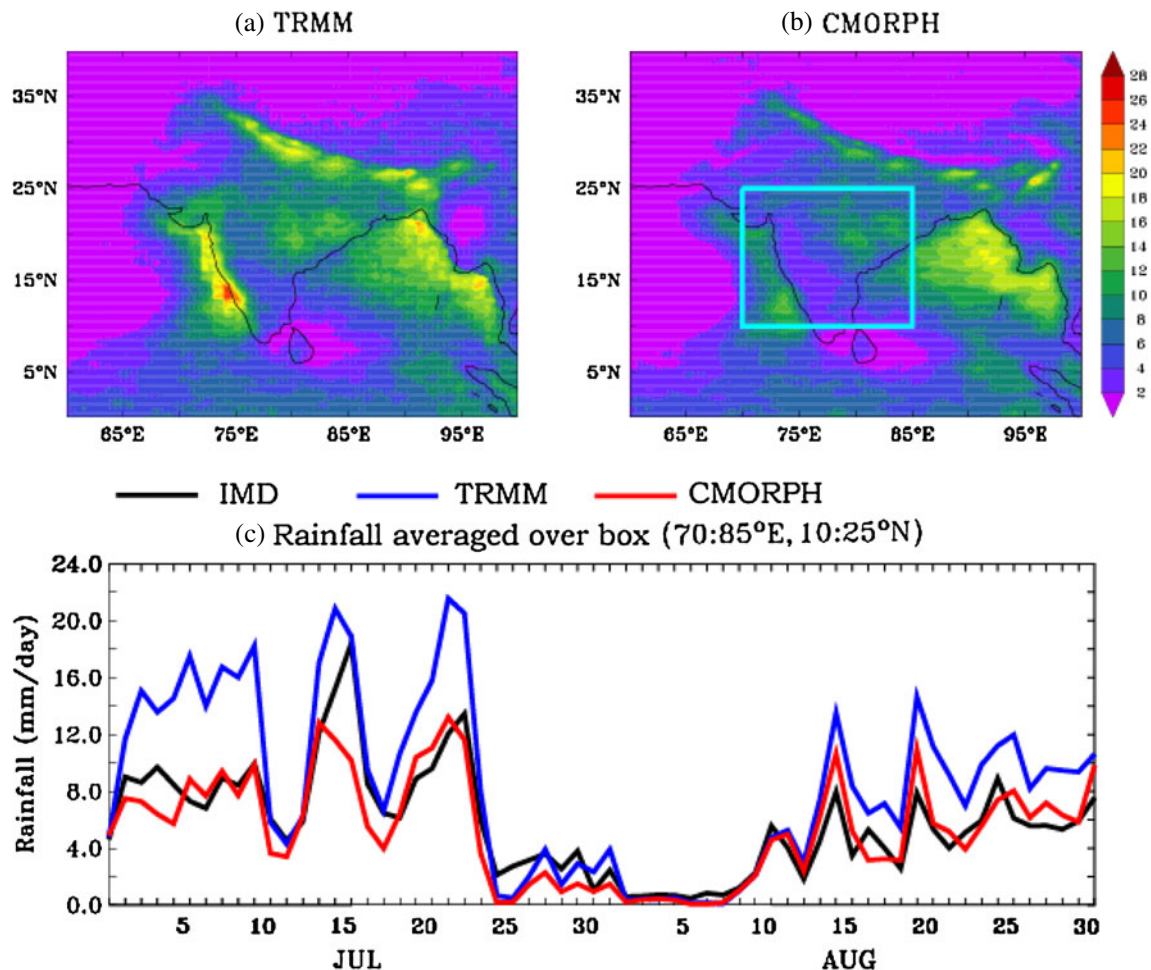

Figure 1. July–August average precipitation [mm/day] for (**a**) TRMM and (**b**) CMORPH for 2008–2010. Grey box shows the monsoon core region (MCR) considered for current study. The observed IMD (Rajeevan *et al.* 2006, black) TRMM (blue) and CMORPH (red) rainfall averaged for the box, 70°–85°N; 10°–25°E for 2009 July–August period. Oceanic region is masked in TRMM and CMORPH before taking box average.

gives a better value (RMSE = 1.82 mm/day) providing a better confidence for selecting it as observed climatology in this study.

Observed winds used here is from the ECMWF, that is based on the four-dimensional variational analysis (4DVAR) (Dee *et al.* 2011). This study also includes the HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) model of NOAA using semi-lagrangian trajectory algorithm for computing back trajectory analysis (Draxler and Rolph 2012).

2.2 *Details of methods used in current study*

The daily forecasts of precipitation and winds were extracted at 1200 UTC from the archives of the TIGGE portal. The SE employed in this study contains the ensemble prediction system from CMA (China Meteorological Administration), CMC (Canadian Meteorological Center), CPTEC (Centro de Previsao de Tempoe Estudos Climaticos, Brazil), ECMWF (European Centre

for Medium-range Weather Forecasts, UK), NCEP (National Centers for Environmental Prediction, USA), and UKMO (United Kingdom Meteorological Office). Detailed information about these models is provided in table 1. As we wanted to evaluate the skill for the lead time of prediction of the occurrence of dry spells for n days, where n lies anywhere between 1 and 10, only those models could be included in the study that ran 10-day forecasts through the period 2008–2011. This was necessary, since we wanted to evaluate the skills for the lead time of prediction for the occurrence of dry spell for n days, where n lies anywhere between 1 and 10 days. Each member model's forecast for each day was interpolated to a common resolution of 25 km, using 3rd order Bessel interpolation method (Carl de Boor 1978). During a training phase of the SE, the forecasts are statistically downscaled to 25 km resolution. Outlier in the downscaled data were removed by the Robust linear fit method (Press *et al.* 1986). Figure 2 gives the idea of improvement in the

Table 1. *Operational global ensemble prediction models details from TIGGE data portal.*

	CMA, China	CMC, Canada	CPTEC, Brazil	ECMWF, Europe	NCEP, USA	UKMO, UK
Resolution	TL213L31	0.9degL28	T126L28	TL639L62	T126L28	0.5555° (lat.) × 0.8333° (long.)
Initial UTC	00, **12**	00, **12**	00, **12**	00, **12**	00, 06, **12**, 18	00, **12**
Member/run	15	21	15	51	21	24
Forecast length (Days)	10	16	15	15	16	15
Available period	2007.05.15. 00-Till present	2007.10.03. 00-Till present	2008.02.01. 00-Till present	2006.10.01. 00-Till present	2007.03.05. 00-Till present	2006.10.01. 00-Till present

List of operational medium-range ensemble forecast system by numerical weather prediction (NWP) centers:– CMA: China Meteorological Administration, CMC: Canadian Meteorological Center, CPTEC: Centro de Previsão de Tempo e Estudos Climáticos, Brazil, ECMWF: European Centre for Medium-range Weather Forecasts, UK, NCEP: National Centers for Environmental Prediction, USA and UKMO: United Kingdom Meteorological Office, UK.
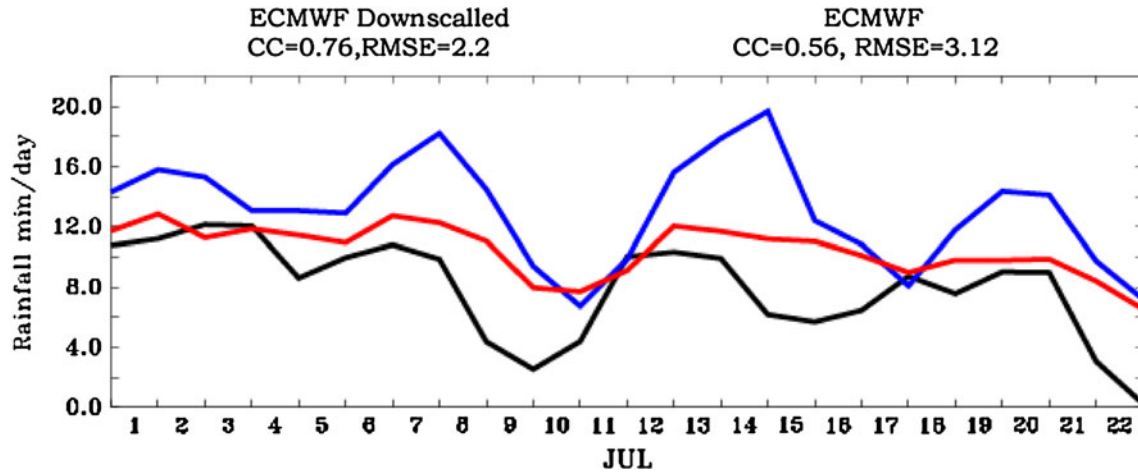

Figure 2. Averaged precipitation over MCR for July 2009 from CMORPH (black), day-1 forecast of ECMWF model after Bessel interpolation (blue) and after downscalling (red). Values of RMSE and CC curves are indicated in top of the panel.

phase (CC ~ 0.8) and amplitude of downscaled precipitation from the Bessel interpolated a member model (e.g., ECMWF model) through reduction of RMSE ~ 2.2. This step is very important for improving the statistical weights of the SE, since improved forecast skills were noted from such improved statistical weights.

The well known SE methodology for multi-model forecasts, discussed in detail in a series of publications by Krishnamurti *et al.* (e.g., 1999, 2000), is used in the present study. This technique entails two phases, with the first being a 'training phase' and the second being a 'forecast phase'. The training phase calculates weights (a_i) based on the least square fit of the recent past forecasts regressed towards the observation based fields, and this provides a_i for the forecast phase for constructing the SE forecasts. The training phase contains a period of ~ 186 days in July and August, which is independent of the forecast data phase. More clearly if we are going to predict the dry spell in the year 2011, we prepared the training sets using the forecast data from July–August of 2008–2010 and in a similar fashion for predicting other dry spells in different years. The temporal model anomalies of the variable are regressed against observation, and multiplied with a_i for the construction of SE forecast, which is given by

$$S = \overline{O} + \sum_{i=1}^{N} a_i (F_i - \overline{F}_i) \qquad (1)$$

where S is the SE prediction, $\overline{O}$ is the observed climatology over the training period, F_i and $\overline{F}_i$ are the forecast and forecast mean for a model i for training period, and N is the number of member models and weight a_i can be calculated by minimizing the error term G is obtained as in equation (2)

$$G = \sum_{t=1}^{N_{\mathrm{train}}} (S'_t - O'_t)^2 \qquad (2)$$

where N_{train} is the number of time samples in the training phase, S' and O' are the SE and observed fields, respectively at the training time t. In addition to superensemble, we also used an ensemble mean (EM), where weights are partitioned equally to all the employed models.

3. Predictability of monsoonal dry spells

The skill of the individual member models and the SE for the period from 16 July to 15 August 2009 (cover both active to break and break period) can be assessed by investigating Taylor diagrams for day-1, day-5 and day-7 of forecasts (figure 3a, b, c) over MCR (box indicated in figure 1). The
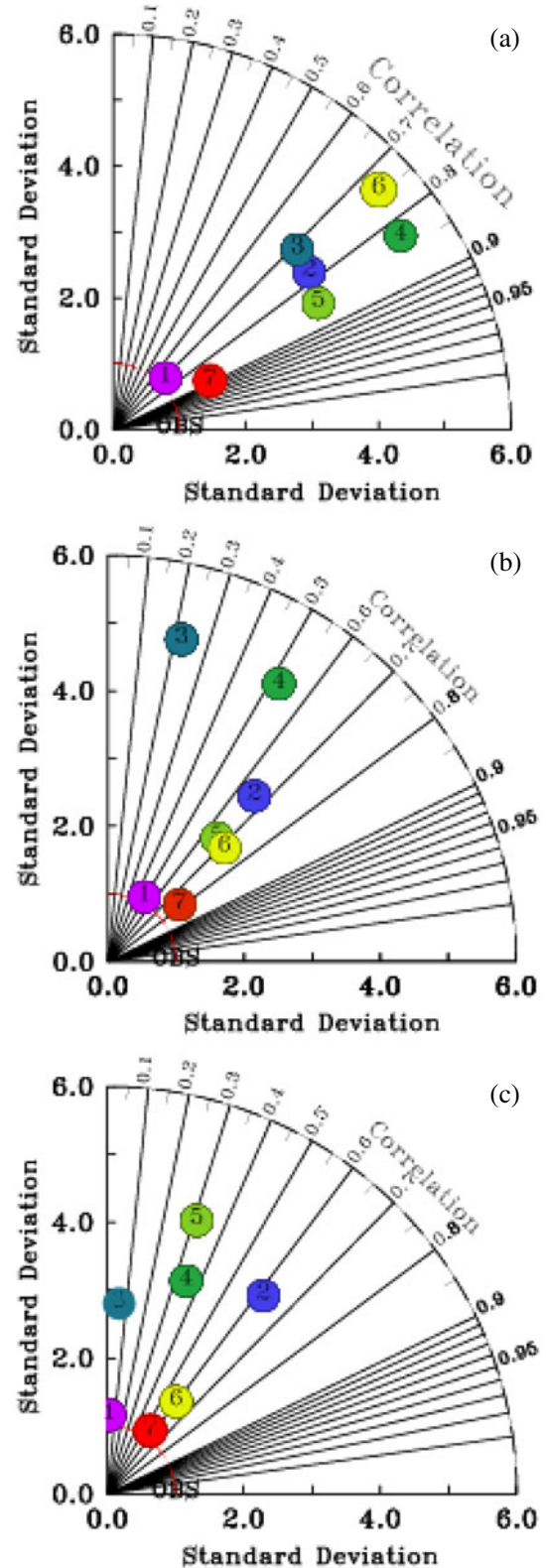

Figure 3. Taylors diagram for the models. 1) CMA, 2) CMC, 3) CPTEC, 4) ECMWF, 5) NCEP, 6) UKMO and 7) SE for the (**a**) Day-1, (**b**) Day-5 and (**c**) Day-7 valid over MCR for the period of 16 July–15 August 2009.

Taylor diagram (Taylor 2001) is often used in the ranking of a suite of multi-models for weather and seasonal climate forecasts. This is a polar diagram

with the standard deviation (of performance) as the radial leg and the forecast pattern correlation as the azimuthal coordinate. A perfect score for a forecast carries standard deviation of 1 (since it is normalized with respect to the observed value) and a correlation of 1.0. The Taylor diagram displays how far off is a particular model's forecast from that perfect score location of this polar diagram. The best forecasts invariably come from the SE (figure 3, red circle). That point shown in red is closest to the position of the perfect forecast for days 1, 5 or 7, whereas the positions for all member models of our suite show larger errors. Those are largely systematic errors that are reduced by the SE which removes the collective bias errors of all member models. This can be seen for the forecasts (figure 3a, b and c). The NCEP/GFS (light-green circle) carries the largest errors for 7-day of forecasts compared to its 1 and 2 day forecast, whereas UKMO (yellow circle) shows a smaller error for the 7-day forecast as compared to the other member models. CMA model had relatively good performance for the standard deviations compared to other member models, but the correlation indicates lower skill, especially after 5-day forecast string.

We have selected three major dry spell events covering the months July and August during 2009–2011 for examining the predictability of monsoon spells (table 2). Based on the method described by Ramamurthy (1969); Gadgil and Joseph (2003); and Rajeevan *et al.* (2010), the following days were identified as break days in the monsoon season: 25 July–6 August 2009, 5–10 August 2010 and 24 July–8 August 2011. Model-predicted skills for the above break days are given in table 2. Table shows the skill of models in predicting the lead time for forecasts of the dry spells. The SE carries a correlation coefficient ∼0.5 (day-1 to day-7) indicating its useful predictive skill. In this table, the member models show lower skills for the RMSE and for the CC compared to the SE for all of the break cases covering the years 2009, 2010 and 2011 that were included here. The member models with the highest and the lowest skills for the prediction of breaks respectively are the UKMO and the CPTEC respectively. The consistently higher skill of the SE strongly suggests that for operational practice such predictions of break events can be useful.

More specifically, figure 4(a, b, c) shows an event that has an initial state during a wet period and leads to a dry spell cycle during the 2009 summer monsoon season. Here the observed benchmark for precipitation comes from CMORPH, which is shown as a dark line in figure 4(a, b, c) and covers the period from 16 July through 12 August, 2009, also shown are the TIGGE forecast values from all member models. The day-1 forecasts, shown in

Table 2. *Model predicted skills (CC = correlation coefficient, RMSE = Root Mean Square Error) for the break days on the aforementioned years.*

| Year	CMA		CMC		CPTEC		ECMWF		NCEP		UKMO		EM		SE	
	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE
2009																
Day-1	0.58	1.24	0.64	0.96	0.67	1.47	0.83	0.73	0.76	1.62	0.67	1.18	0.74	0.89	0.78	0.75
Day-5	0.12	2.12	0.57	1.06	0.76	1.05	0.72	1.05	0.59	1.68	0.59	1.19	0.63	0.97	0.69	0.86
Day-7	0.05	3.70	0.40	3.24	0.30	2.84	0.36	3.17	0.38	3.30	0.63	1.80	0.48	1.90	0.64	1.29
2010																
Day-1	0.34	2.1	0.80	1.38	0.37	1.68	0.36	1.28	0.80	1.25	0.52	1.49	0.68	1.09	0.91	0.72
Day-5	0.41	2.4	0.37	5.18	0.06	5.32	0.63	3.8	0.39	3.89	0.71	1.82	0.57	3.92	0.73	1.64
Day-7	0.21	3.4	0.18	6.0	0.37	4.0	0.46	5.6	0.14	6.09	0.61	3.53	0.4	4.29	0.63	3.34
2011																
Day-1	0.58	2.38	0.75	2.0	0.76	2.34	0.72	2.21	0.51	3.2	0.71	2.04	0.74	2.01	0.89	1.38
Day-5	0.49	2.24	0.05	2.84	0.30	3.24	0.39	2.57	0.04	3.2	0.64	1.97	0.45	2.29	0.67	1.85
Day-7	0.19	3.03	0.02	4.48	−0.19	5.23	0.44	2.64	0.50	3.19	0.61	2.39	0.36	2.53	0.62	1.88
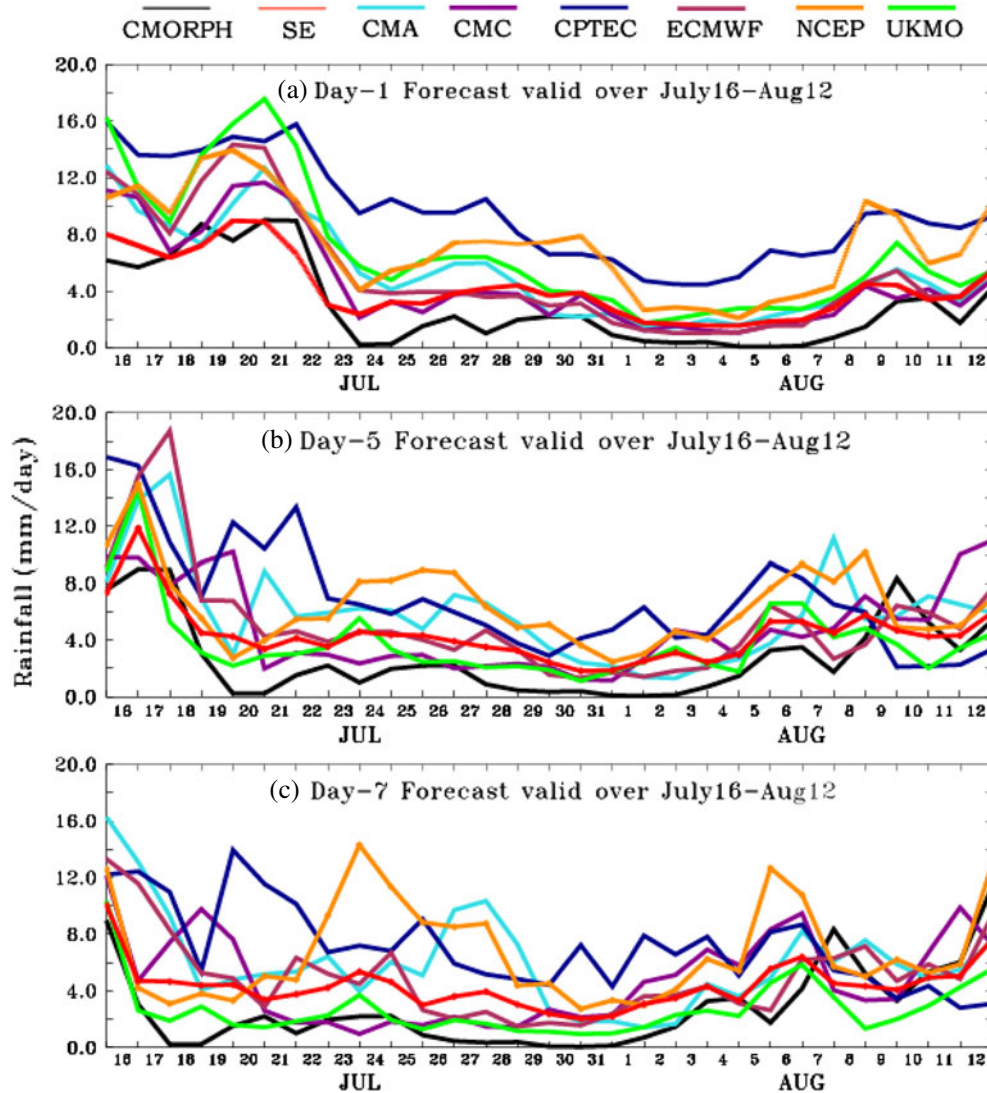
Figure 4. (**a**) Day-1, (**b**) day-5 and (**c**) day-7 forecast of the precipitation (mm/day) valid over July 14–August 13 from CMA, CMC, CPTEC, ECMWF, GFS, UKMO and SE over MCR region. Corresponding observed rainfall from CMORPH for the aforementioned forecasted data is depicted as black curve.

figure 4(a), carries results of many forecasts starting from July 16 (this is a string of day-1 forecasts). This shows a dry spell that lasted from roughly 24th of July to August 8th. This implies that the superensemble carries very high skill for forecasting at day-1 skill for the occurrence of a dry spell (meaning a dry spell starts one day after the start of the forecast). The forecasts for the strings of day-5 and day-7 are illustrated in figure 4(b and c). Basically all these forecasts show that the curve closest to the observed (black curve) is the SE. When the performance of the individual member models are examined, it is noted that a much delayed start date for the onset of the break period. Furthermore, most models carried rains of the order 4 mm/day or higher for days 1 through 7 of forecasts even though these were to represent their dry spells. The CPTEC model had rains of

the order of 10 mm/day even on July 26 (day-5 forecast string), the observed start date for the break was in 25 July. The CPTEC model carried rains in excess of 4 mm/day through all days of its forecasts. All models other than the UKMO model carried large errors in their prediction of this break spell of the monsoon. The GFS (CC = 0.23, RMSE = 3.8) shown by an yellow curve carries far lower skill for precipitation on day-7 forecast over India compared to ECMWF, CMC and UKMO (CC = 0.5, RMSE = 2.5) generally. The SE appears to be the best precipitation forecasts product for one week forecasts in a rather consistent manner.

Figure 5 shows the spatial variability of the predicted rainfall amount in 7 days in advance, commencing on 24 July 2009 for each model and estimated spatial correlation (SC) is indicated above each model panels. The dates are chosen
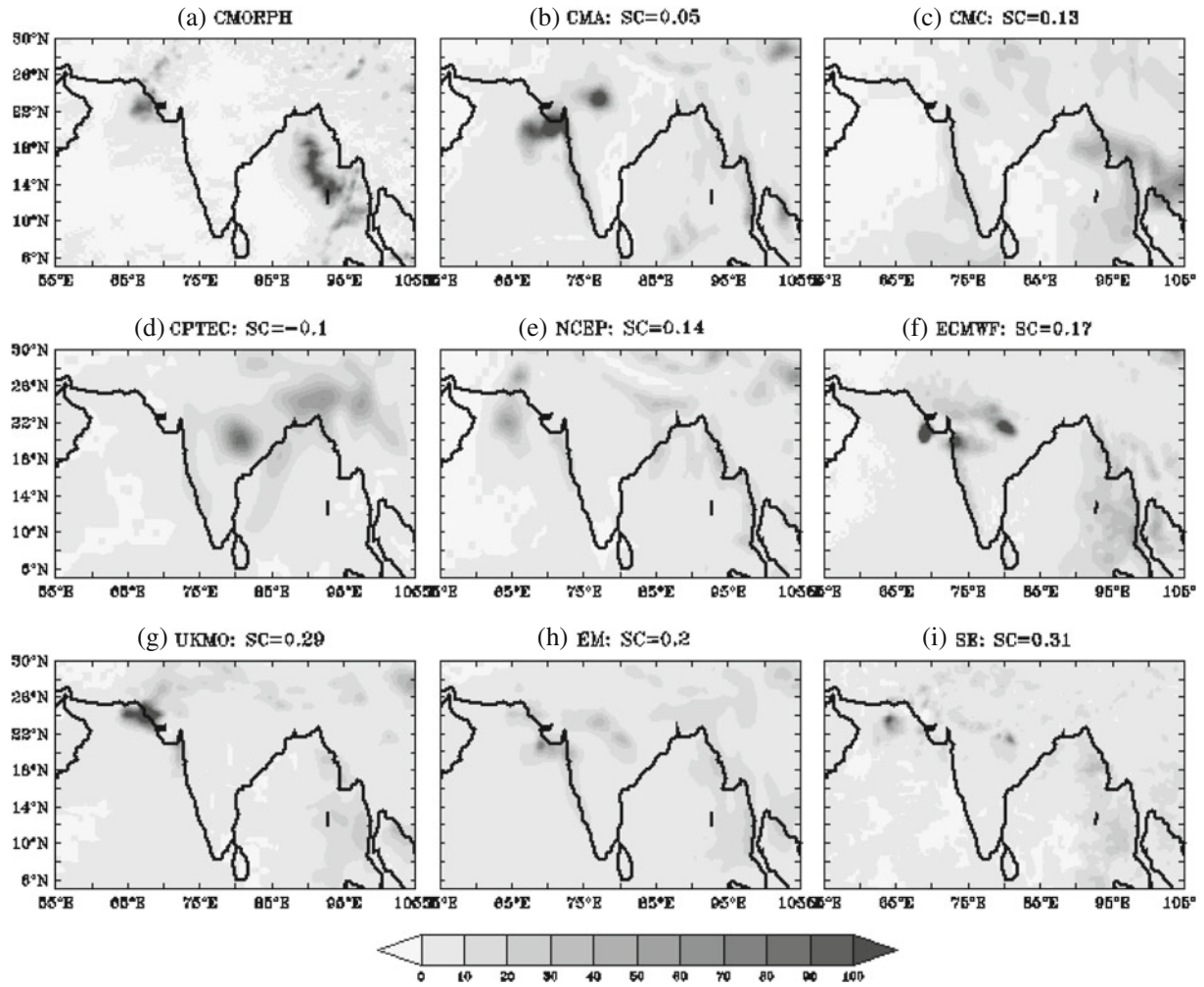
Figure 5. Day-7 forecast of total precipitation (mm/day) (shaded) valid for 24 July 2009 and the corresponding spatial correlation (SC) value indicated above each panel.

such a way as the dry spell begins for the aforementioned case. There were two rainfall maxima observed, one is above North Bay of Bengal and another is over north-east Arabian Sea. Structure of spatial variability of rainfall shows comparatively good agreement with the ECMWF and UKMO, while CPTEC shows maxima over central India. SC of member model ranges from 0.10 to 0.29, excluding CPTEC, whereas much improvement is shown in the structure of predicted rainfall in the superensemble (SC = 0.31).

4. Observed wind circulation features and its predictability associated with monsoon spells

Dynamic interaction between rainfall and the associated divergent winds have an important role in modifying the monsoon behaviour and the intraseasonal variations related to these dry and

wet spells. This is addressed in the following subsections.

4.1 *Local Hadley Cell for the Indian monsoon domain for monsoon spell*

Basically the formulation of the Global Hadley Cell is adopted for limited span of longitudes. The method of construction of the Hadley Cell is the following: we start from the zonally averaged mass continuity equation using pressure as the vertical coordinate: (averaged here over a local belt of longitudes).

$$\frac{\partial \, [\overline{v}]}{a \partial \phi} - \frac{[\overline{v}] \tan \phi}{a} + \frac{\partial \, [\overline{\omega}]}{\partial p} = 0 \qquad (3)$$

where v the north/south (meridional) velocity, ω the vertical p-velocity $(\mathrm{d}p/\mathrm{d}t)$, a the earth's radius and ϕ the latitude. The square bracket is the zonal mean over a limited longitudinal belt, and the

overbar is the time mean. This is re-expressed in the form:

$$\frac{\partial \left[\overline{v}\right]}{a\partial\phi}\cos\phi + \frac{\partial \left[\overline{\omega}\right]}{\partial p}\cos\phi = 0. \qquad (4)$$

We next define a Hadley Cell stream function (ψ) by the pair of equations

$$\frac{\partial \left[\overline{\psi}\right]}{a\partial\phi} = -\left[\overline{\omega}\right]\frac{2\pi a}{g}\cos\phi = 0 \qquad (5)$$

$$\frac{\partial \left[\overline{\psi}\right]}{\partial p} = -\left[\overline{v}\right]\frac{2\pi a}{g}\cos\phi = 0 \qquad (6)$$

and it carries the units kg s^{-1}.

Note that this stream function satisfies the mass continuity equation. Given the datasets for the meridional wind, it is first averaged zonally and next in time over a period and a region of interest. All one needs is to vertically integrate equation (6) with the boundary condition $\psi = 0$ at the earth's surface. That provides the stream function at all vertical levels. The local Hadley Cell is simply the same construction over the Indian Monsoon longitude that is here used between 70° and 85°E longitudes.

During the active monsoon periods the largest ascent for the local Hadley Cell generally resides at around 23°N latitude over the Gangetic belt of north India (figure 6). The descending lobe of the
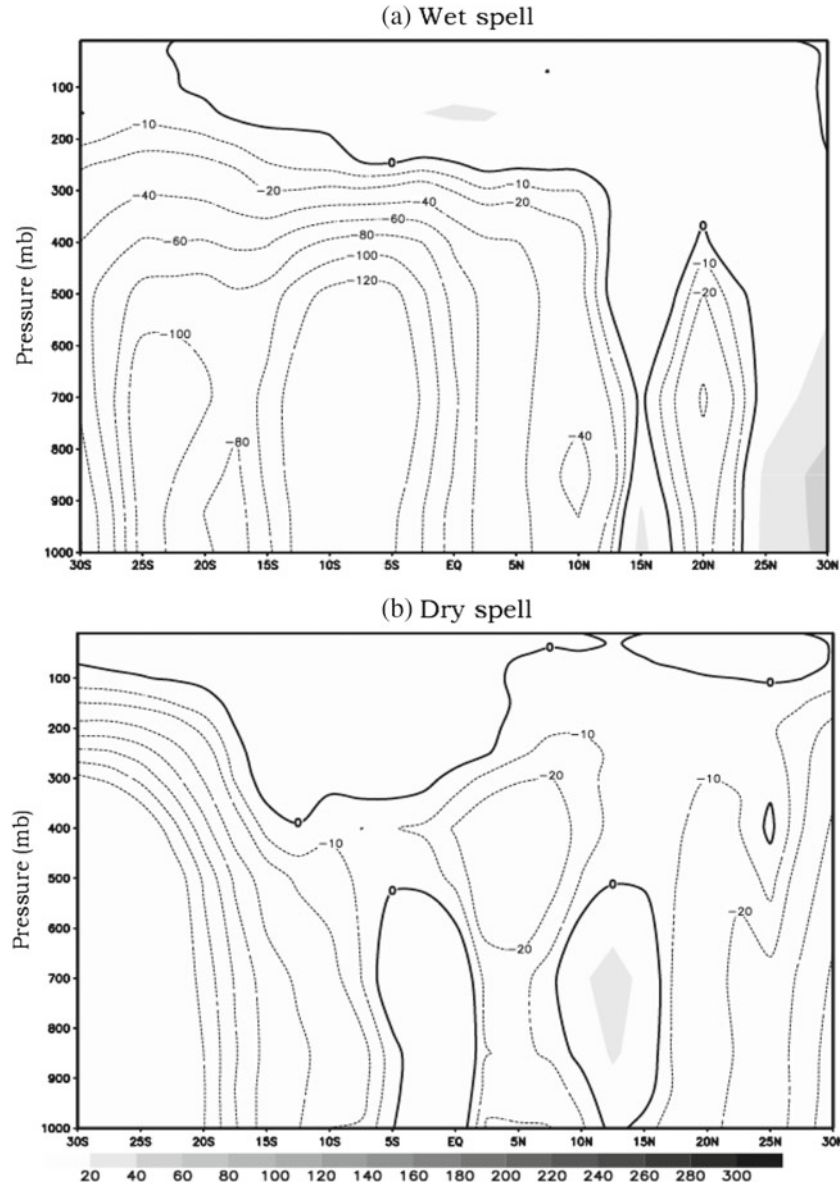

Figure 6. The local Hadley cell stream function (10^9 m^2 s^{-2}) (**a**) for the wet spells (18th July 2009) and (**b**) for the dry spells (30th July 2009) over central India. Positive values of stream function is shaded and negative value is contoured.
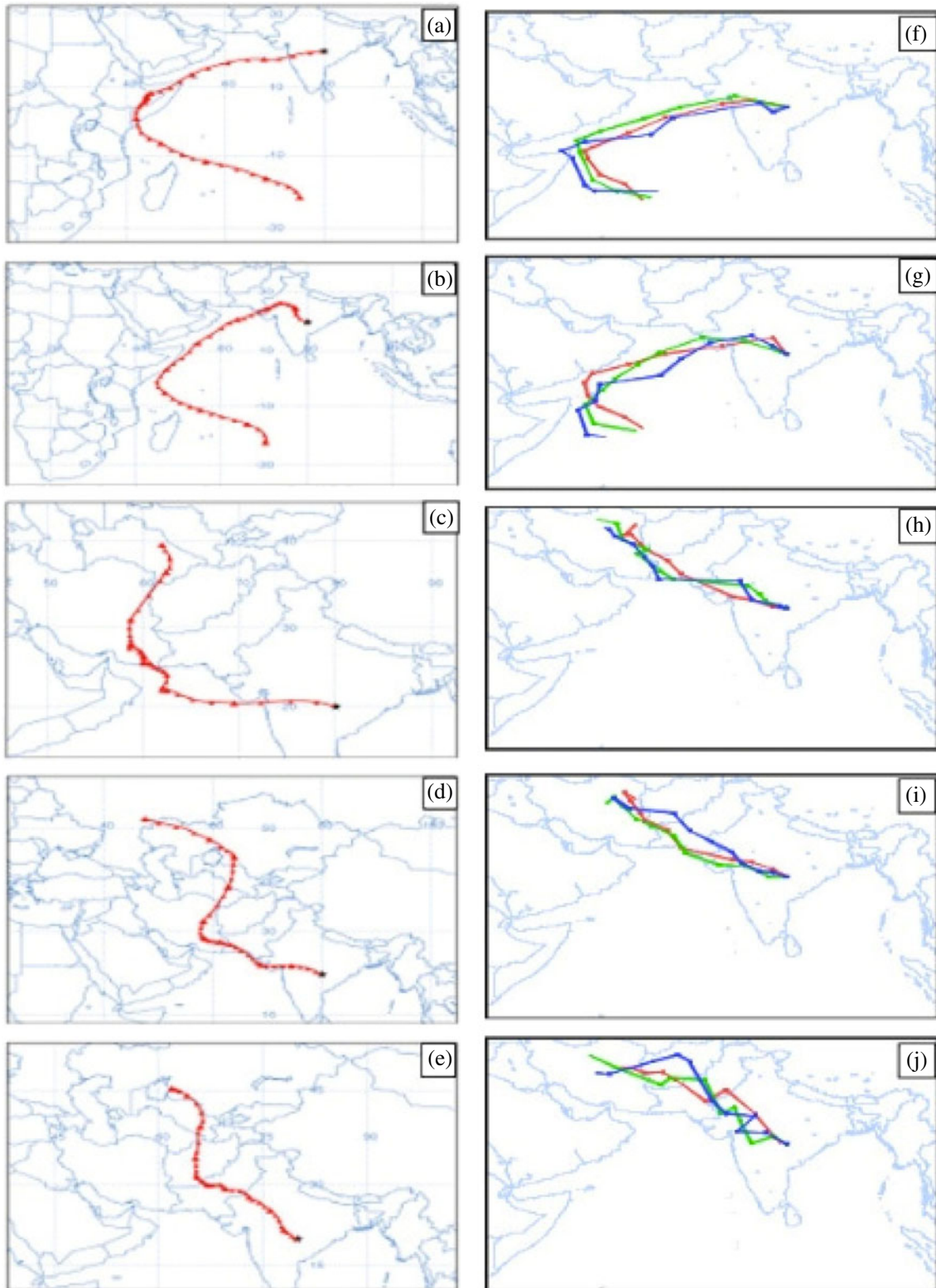
Figure 7. Ten-day back trajectories from central India (80°E, 20°N) from HYSPLIT model (left panels) and 1-day (red), 5-day (green) and 7-day (blue) forecast of superensemble (right panels) for trajectories terminating at 700 hPa on (**a**, **f**) 18 July 2009, (**b**, **g**) 22 July 2009, (**c**, **h**) 26 July 2009, (**d**, **i**) 30 July 2009 and (**e**, **j**) 4 August 2009.

primary Hadley cell during these periods is located near 28°S latitude. During dry spells the rainfall belt shifts northwards to near 28°N. Consequence of the northward shift of the rain belt is a longer span of latitudes over which the northerly flow of the upper tropospheric branch of the Hadley cell is found. Conservation of angular momentum for these air parcels moving south from 28°N calls for a stronger upper tropospheric easterlies and a strengthening of the tropical easterly jet during the dry spells can thus be expected.

4.2 *Transition mode of trajectories*

CMORPH precipitation (figure 1c) conveys the periods of dry and wet for the summer monsoon season of 2009 and is consistent with the IMD gridded rainfall (Rajeevan *et al.* 2006); both of these datasets confirmed a major wet spells between 18 July and 10 August. In between these wet spells

is a period of dry spell that last from roughly 25 July to 4 August. The dynamics of dry spells was addressed in a recent study by Krishnamurti *et al.* (2010). They noted a number of dynamical antecedents that preceded the dry spell. They showed that during dry spells, the antecedents included a deep blocking high over Arabia that extended from 700 to 400 hPa levels, descending very dry air over the Arabian Sea and towards central India. The marine air of oceanic origin was limited between the surface and the 850 hPa levels. In contrast during wet spells of summer monsoon over India, air of oceanic origin prevailed at all vertical levels between the surface and the 400 hPa levels. Figure 7(a, b) shows typical trajectories using HYSPLIT that terminate over central India during wet spells of the year 2009. These are trajectories that terminate at the 850 hPa level on July 18 and 22. These clearly confirm that air of oceanic origin terminates over central India at 850 hPa
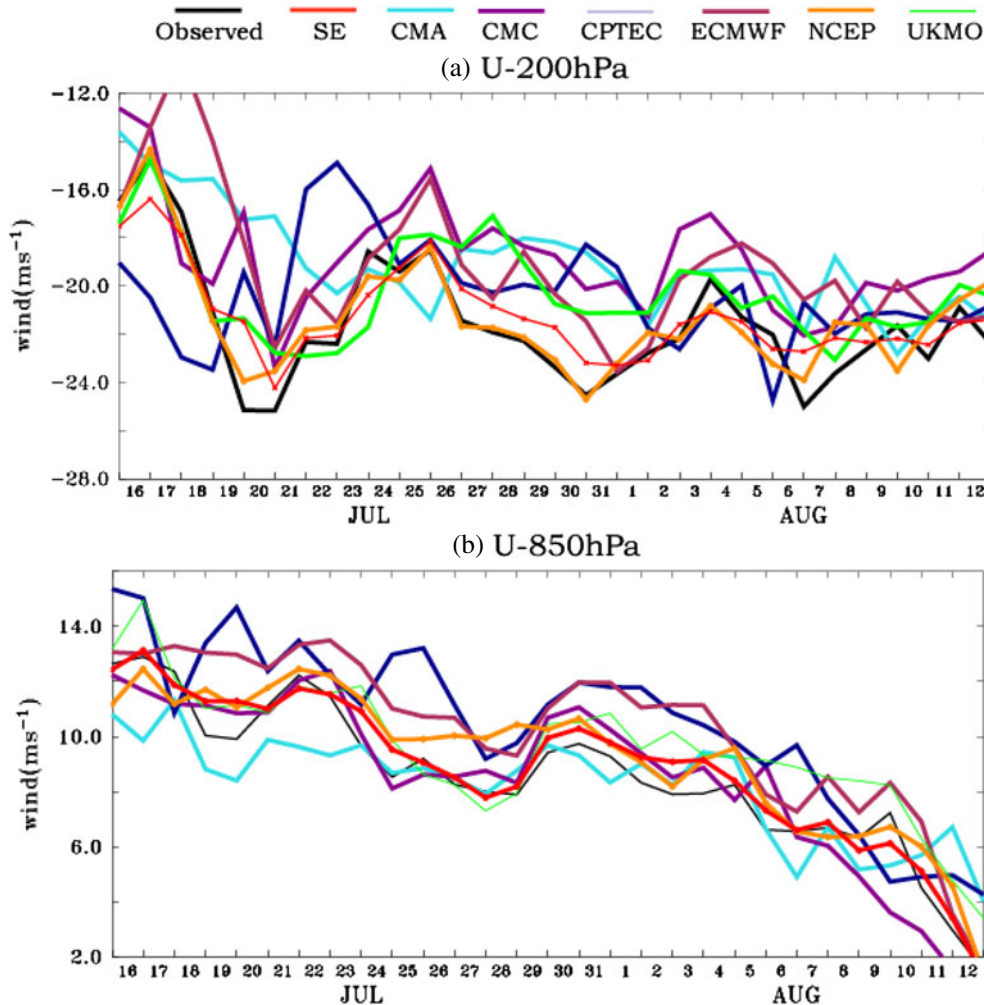

Figure 8. Five-day of the zonal wind (ms^{-1}) valid over July 14–August 13, 2009 from CMA, CMC, CPTEC, ECMWF, GFS, UKMO and SE over MCR for (**a**) 200 hPa and (**b**) 850 hPa. Corresponding observed wind field for the aforementioned forecasted data is depicted as black curve.

level during this wet spell. Figure 7(c, d, e) shows the dry spell trajectories that terminate at the 850 hPa level over central India. These descending parcels arrive from the northwest enter central India around 700 hPa level. This mode of transition of trajectories was noted during the initiation of wet (18 July 2009, figure 7a) and dry spells (26 July 2009, figure 7c).

Trajectory computations make use of the horizontal wind components u and v, and the vertical velocity ω, at all vertical levels from all member models and SE. The trajectory analysis for the string of day-1 forecasts (red curve) using SE is shown in the right panel of figure 7(f) through (j). Two similar for day-5 (green) and day-7 (blue) forecasts are illustrated in similar fashion in figures 7(f–j), these essentially confirm the passage of dry air from the northwest over India during this dry spell. These illustrations support the possibility for monitoring the dry spell for days 1 though 10 using wind fields employed in SE. Similar trajectory analysis for each member model can be illustrated to check the performance of forecast wind field, helping in guidance for the eventual occurrence of a dry spell (figure not shown).

4.3 *Prediction of lower and upper tropospheric winds during dry spells*

Predictability of aforementioned local Hadley cell circulation changes and mode of variability in the trajectory during dry spell depends on the performance of forecast wind in the member models. Figure 8 shows the predicted winds, for day-5 of forecasts, for the upper troposperic easterly winds (U-200 hPa) and for the lower tropospheric westerly winds (U-850 hPa) over MCR. As discussed in section 4.1, strengthening of the upper tropospheric zonal winds corresponding to the dry spell period is clearly evident in both model forecast and observations. In the upper level winds, our computations show that the initial 24-hour errors at the 200 hPa level winds for ECMWF are as large as 2 ms^{-1} (figure not attached), those are still of the order of 5 ms^{-1} at 120-hour forecast for the date of 17 July 2009 (figure 8). These errors are much reduced by the SE. Seasonal reduction in the strength of the lower tropospheric winds from July to August with intraseasonal oscillation is also evident in figure 8(b). All the models show this seasonal variability of wind field, the SE carries the best forecasts since that follows the observations very closely. There is a strong tendency for model's inability to carry a long dry spell in its forecasts (all computations are not indicated); the tendency to revive the monsoon somewhat quickly contributes to these increasing errors.

5. Summary and discussion

This study addresses the number of days for possible prediction of the breaks in the monsoon rains over India. In order to address this problem we have used both CMORPH and the multi-model ensemble datasets from the TIGGE stream for SE construction. The SE for precipitation forecasts invariably outperforms all other member models in predicting the dry spells of the monsoon through day-7 of forecasts, based on different statistical analysis (Taylor's diagram, SC, RMSE). The important question is one of definition of the start and duration of a dry spell over India, we have followed the definition of India Meteorological Department in the present work which is: periods during which the standardized rainfall is less than −1.0, consecutively for three days over the monsoon core zone during the peak monsoon months of July and August (Rajeevan *et al.* 2010). Figure 9 shows the reduction of the spatial correlation for the lead time of starting date of break from day-1 string to day-10. Worst model (blue) in our SE methodology (based on RMSE and CC) shows a negative spatial correlation after 6–7 days in advance over Indian monsoon domain, indicate the limit of predictability of our NWP models.

The lower and upper tropospheric winds, associated with monsoonal dry spells over India are also being predicted much better by the SE compared to all member models. Strengthening of Tropical easterly jet associated with the northward shift of the Hadley cell during the break spell is predicted in week advance by many of the member models forecasted. The TIGGE dataset-based forecasts of the dry and wet spells carry trajectories that are close to those computed from observed datasets. The horizontal wind field and the vertical velocity ω, together portray trajectories that clearly illustrate the mode of transition of trajectories from wet (entirely oceanic below 400 hPa levels) to dry spell (that conveys oceanic trajectories to India from the oceans only below the 850 hPa levels).

We have found generally all three datasets CMORPH, TRMM and the IMD rain gauge collection quite useful for the validation of model performances. CMORPH seems to carry somewhat finer spatial structures because of the way it interpolates the data in space and time, however for the limited land area of India all three datasets seem quite comparable in their value for the validation of dry spells. The message from this study should determine the practical utility of the proposed method for the prediction of the start and the duration of the dry spell. In that context it appears that this method is indeed quite powerful to help real time operations. The TIGGE data are currently available quasi-operationally (2-day delay), that
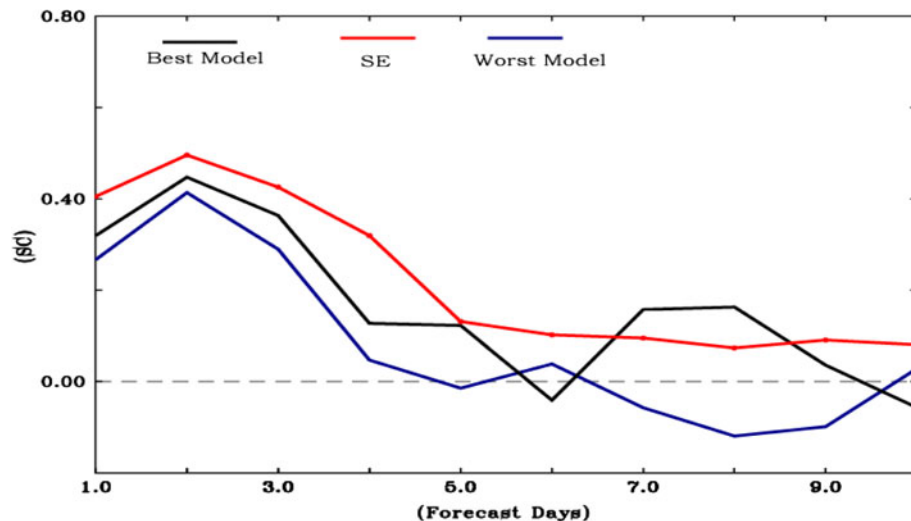
Figure 9. Spatial correlation (SC) of rainfall for each forecast days (1–10) valid over starting date of the dry spell of 2009 (24th July 2009) over MCR. Best model (black) and worst model (blue) among the models architecture used in the current study, plotted over superensemble (SE) rainfall.

access should be facilitated even more for the use of weather services. Given 10-day forecasts from the best operational models, it would be a relatively simple matter to construct the SE on a routine basis and search for the occurrence criteria for the onset and duration of a dry spell. Computationally, the post-processing for the construction of a SE is not a lengthy process. The entire forecast and the preparation of dry spell alerts can be easily made into an automatic dissemination process. Further improvements in the post-processing methodology for the construction of multi-model ensemble is an important area for further research. Temporal correlation of the CMORPH compared to the TRMM and rainguage derived data (Rajeevan *et al.* 2006) is slightly lesser during monsoon period, considered to be a caveat to current study. More number of member model included for the weight calculation in the training phase, will help to improve superensemble method much further.

References

Bougeault P and Coauthors 2010 The THORPEX interactive grand global ensemble; *Bull. Am. Meteorol. Soc.* **91** 1059–1072.

Carl de Boor 1978 A practical guide to splines; Rev. edn, *Appl. Math. Sci.* **27**, Springer-Verlag, New York Inc.

Dee D P and Coauthors 2011 The ERA-Interim reanalysis: Configuration and performance of the data assimilation system; *Quat. J. Roy. Meteorol. Soc.* **137** 553–597, doi: 10.1002/qj.828.

Draxler R R and Rolph G D 2012 HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY website (http://ready.arl.noaa.gov/HYSPLIT.php). NOAA Air Resources Laboratory, Silver Spring, MD.

Ferraro R R 1997 SSM/I derived global rainfall estimates for climatological applications; *J. Geophys. Res.* **102** 16,715–16,735.

Gadgil Sulochana and Joseph P V 2003 On breaks of the Indian monsoon; *Proc. Indian Acad. Sci. (Earth Planet. Sci.)* **112** 529–558.

Gopalan Kaushik, Nai-Yu Wang, Ralph Ferraro and Chuntao Liu 2010 Status of the TRMM 2A12 Land Precipitation Algorithm; *J. Atmos. Oceanic Technol.* **27** 1343–1354, doi: 10.1175/2010JTECHA1454.1.

Hong Y, Kummerow C and Olson W S 1999 Separation of convective/stratiform precipitation using microwave brightness temperature; *J. Appl. Meteorol.* **38** 1195–1213.

Huffman G J, Adler R F, Rudolf B, Schneider U and Keehn P R 1995 Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information; *J. Climate* **8** 1284–1295.

Joyce R J, Janowiak J E, Arkin P A and Xie P 2004 CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution; *J. Hydrometeorol.* **5** 487–503.

Krishnamurti T N 2005 Weather and seasonal climate prediction of Asian summer monsoon; *Review Topic B1a: Numerical Modeling-Forecast*, pp. 1–34.

Krishnamurti T N, Kishtawal C M, LaRow T E, Bachiochi D R, Zhang Z, Williford C E, Gadgil S and Surendran S 1999 Improved weather and seasonal climate forecasts from a multi-model superensemble; *Science* **285** 1548–1550.

Krishnamurti T N, Kishtawal C M, Shin D W and Williford C E 2000 Improving tropical precipitation forecasts from a multianalysis superensemble; *J. Climate* **13** 4217–4227.

Krishnamurti T N, Mishra A K, Chakraborty A and Rajeevan M 2009 Improving global model precipitation forecasts over India using downscaling and the FSU superensemble. Part I: 1–5-Day Forecasts; *Mon. Weather Rev.* **137** 2713–2735.

Krishnamurti T N, Thomas A, Simon A and Kumar V 2010 Desert air incursions, an overlooked aspect, for the dry spells of Indian summer monsoon; *J. Atmos. Sci.* **67** 3423–3441, doi: 10.1175/2010JAS3440.1.

Krishnan R, Zhang C and Sugi M 2000 Dynamics of breaks in the Indian summer monsoon; *J. Atmos. Sci.* **57** 1354–1372.

Kummerow C, Olson W S and Giglio L 1996 A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors; *IEEE Trans. Geosci. Rem. Sens.* **34** 1213–1232.

Matsueda M and Endo H 2011 Verification of medium-range MJO forecasts with TIGGE; *Geophys. Res. Lett.* **38** L11801, doi: 10.1029/2011GL047480.

Mishra A K and Krishnamurti T N 2007 Current status of multimodel superensemble and operational NWP forecast of the Indian summer monsoon; *J. Earth Syst. Sci.* **116** 369–463.

Neena J M, Suhas E and Goswami B N 2011 Leading role of internal dynamics in the 2009 Indian summer monsoon drought; *J. Geophys. Res.* **116** D13103, doi: 10.1029/2010JD015328.

Press W H, Flannery B P, Teukolsky S A and Vetterling W T 1986 Numerical recipes: The Art of Scientific Computing; Cambridge University Press, Cambridge, 818p.

Rajeevan M, Bhate J, Kale J and Lal B 2006 High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells; *Curr. Sci.* **91** 296–306.

Rajeevan M, Sulochana G and Bhate J 2010 Active and break spells of the Indian summer monsoon; *J. Earth Syst. Sci.* **119** 229–247.

Ramamurthy K 1969 Some aspects of 'break' in the Indian south-west monsoon during July and August. Forecasting Manual, India Meteorological Department Publication, FMU, **4**, 18.3. [Available from India Meteorological Department, Mausam Bhavan, Lodi Road, New Delhi 110 003, India.]

Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram; *J. Geophys. Res.* **106(D7)** 7183–7192, doi: 10.1029/2000JD900719.

Waliser D E, Stern W, Schubert S and Lau K M 2003 Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon; *Quat. J. Roy. Meteorol. Soc.* **129** 2897–2925.

Wang B and Xu X H 1997 Northern hemisphere summer monsoon singularities and climatological intraseasonal oscillation; *J. Climate* **10** 1071–1085.

Weigel A P, Liniger M A and Appenzeller C 2008 Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?; *Quat. J. Roy. Meteorol. Soc.* **134** 241–260, doi: 10.1002/qj.210.

Xie P, Mingyue C, Song Yang, Yatagai A, Tadahiro Hayasaka, Yoshihiro Fukushima and Changming Liu 2007 A gauge-based analysis of daily precipitation over east Asia; *J. Hydrometeorol.* **8** 607–626.

Zhi X, Haixia Q, Yongqing B and Lin C 2012 A comparison of three kinds of multi-model ensemble forecast techniques based on the TIGGE data; *Acta Meteorol. Sinica* **26(1)** 41–51, doi: 10.1007/s13351-012-0104-5.