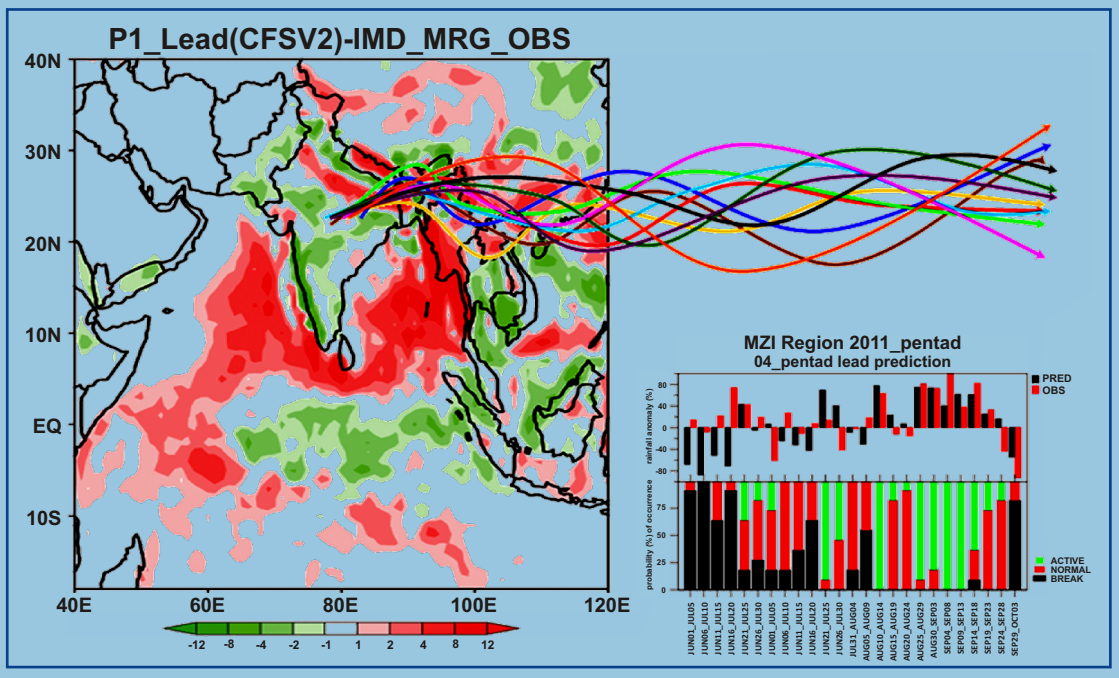
ISSN 0252-1075
Contribution from IITM
Research Report No. RR-128

Performance of an Ensemble Prediction System based on CFSV2 for the Extended Range Prediction of Active-Break Spells of Indian Summer Monsoon Rainfall during 2011



Abhilash S., Susmitha Joseph, R. Chattopadhyay, S. Pattnaik, P. Murali Krishna, S. Dey, A.K. Sahai and B.N. Goswami

December 2012

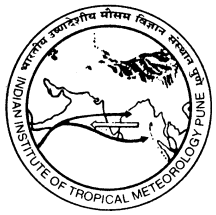
Indian Institute of Tropical Meteorology
Pune - 411 008, India

Performance of an Ensemble Prediction System based on CFSV2 for the Extended Range Prediction of Active-Break Spells of Indian Summer Monsoon Rainfall during 2011

Abhilash S., Susmitha Joseph, R. Chattopadhyay, S. Pattnaik*,

P. Murali Krishna, S. Dey, A. K. Sahai and B.N. Goswami

December 2012



Indian Institute of Tropical Meteorology
Dr. Homi Bhabha Road, Pashan Pune - 411 008
Maharashtra, India

E-mail : lip@tropmet.res.in Fax : 91-020-25865142
Web : http://www.tropmet.res.in Telephone : 91-020-25904200

*Dr Sandeep Pattnaik, Assistant Professor, School of Earth, Ocean and Climate Sciences, Indian Institute of Technology Bhubaneswar, Samantapuri, Bhubaneswar -751 013, Odisha, India

Content

Performance of an Ensemble Prediction System based on CFSV2 for the Extended Range Prediction of Active-Break Spells of Indian Summer Monsoon Rainfall during 2011

Abstract

This study reports the extended range forecast skill of a state-of-the-art operational Ensemble Prediction System (EPS) in an Ocean-Atmospheric coupled modeling frame work during 2011 monsoon season. The ability of the EPS system to produce probabilistic forecasts for three categories of rainfall namely, active, break and normal has also been evaluated. To sample the uncertainty in the initial condition as well as model imperfection, an EPS has been developed based on the NCEP CFS Version2 (T126L64) model, which produces 11 member ensemble forecasts for 45 days lead time. Each ensemble member was generated by slightly perturbing the initial atmospheric conditions with random matrix generated from a random seed. We perturbed the wind, temperature and moisture fields and the amplitude of perturbation for the variables are scaled according to the magnitude of each variable at a given vertical level.

The EPS is able to capture the spatial distribution of rainfall reasonably well, although the rainfall intensity has been slightly underestimated. The pentad lead seasonal mean rainfall structure shows large overestimation over equatorial Indian Ocean and underestimation over Indian land mass and adjoining eastern Arabian Sea and BoB especially at pentad lead 3 and 4. It is noticed that the EPS performed better in the later phase of the monsoon during August and September than the early phase during June. Though the pentad lead forecast captured the seasonal cycle, it relatively underestimated the rainfall over Monsoon Zone of India region for pentad 1 to 4 leads during June, July and August. However, during September 2011, the rainfall is comparable in magnitude to the observations for all lead pentads. The EPS exhibits better skill in 2011 compared to the 10years hindcast skill in the 4^{th} pentad lead.

The skill of the EPS is also compared in capturing the phases of MISO by obtaining the MISO indices as defined by Suhas et al., (2012). It is found that the model overestimates the amplitude of strong MISOs cases. In contrast to the relatively lower skill in actual rainfall prediction during June, the error growth for MISO index is relatively lower during June and increases thereafter.

1. Introduction

Extended range prediction of active/break spells of Indian Summer Monsoon rainfall has a variety of applications depending on the user community ranging from agriculture to planning. Prediction of Indian Summer Monsoon on sub-seasonal time scales, especially three to four weeks in advance is a challenging task, as this time scale falls between the normal predictability limit of deterministic forecast of weather phenomena and the seasonal mean. The extended range time scale is too long that, at this scale, most of the atmospheric memory could be lost. Additionally, it is too short that the oceanic variability does not have considerable influence. The deterministic and statistical approaches alone exhibit relatively lower skill beyond 10 days (Sahai et al, 2008). Hybrid approaches combining the deterministic and probabilistic forecasts from a dynamical ensemble prediction system has shown improved prediction skill over Indian region on sub-seasonal and seasonal time scales (Pattanaik and Kumar, 2010; Rai and Krishnamurthy, 2011; Abhilash et al, 2012).

The intraseasonal Oscillations (ISOs) are manifested as active-break cycles of ISMR (Lau et al. 1998; Waliser et al. 2003; Goswami 2005). The mean spatial structure of rainfall and circulation fields associated with active and break conditions (Ramaswamy 1962; Ramamurthy 1969; Krishnamurti and Subrahmanyam 1982; Krishnamurti et al. 1985; Webster et al. 1998; Krishnan et al. 2000; Annamalai and Sperber 2005; Goswami 2005; Joseph et al. 2009) have very large spatial scale extending far beyond the Indian continent. An important characteristic of these intraseasonal spells is the repeated northward propagation of the zonally oriented cloud band from south of equator to about 25°N (Sikka and Gadgil 1980; Yasunari 1979). Further, the nonlinear relationship between the precipitation and the large-scale circulation indicates that the active-break spells are influenced by air-sea interactions (Sengupta et al. 2001; Fu et al. 2003; Waliser et al. 2003) and are related to a convectively coupled oscillation consistent with theory (Jiang et al. 2004; Wang 2005; Wang et al. 2009; Goswami et al. 2011). Therefore a coupled ocean-atmosphere climate model may be essential for predicting

active-break spells on extended range. Forecast skill in the extended range is largely determined by the model's ability to represent the ISO and associated variabilities.

It is also conceivable that the extended range prediction skill may also depend on different phases of monsoon ISO's. Xavier and Goswami (2007) found that the transition from break to active is much more chaotic than that from active to break and hence the break phases of monsoon ISO's are intrinsically more predictable than the active phase. Uncertainty in the dynamical extended range prediction mainly arises from the uncertainty in the initial conditions and also from model uncertainty in terms of incomplete representation of the physics of the problem. Ensemble approach is usually used to sample these two sources of errors (Palmer et. al.1993; Richardson 1998; Harrison et al. 1999). The operational Ensemble Prediction System (EPS) by European Centre for Medium Range Weather Forecasting (ECMWF) produce largest ensemble sets of 51 members in real time (Molteni et al. 1996, Buizza et al. 2007, 2008). Global Ensemble System (GES) of NCEP produce ensemble of 21 members in their real time operational analysis/forecast cycle (Toth and Kalney 1997). Rashid et al. (2010) assessed the Madden-Julian-Oscillation (MJO) prediction skill using a 10 member ensemble of hindcast from POAMA (Australian Bureau of Meteorology coupled ocean-atmosphere seasonal prediction system). They found that MJO can be predicted to about 21 days, as measured by bivariate correlation exceeding 0.5. Vitart and Molteni (2009) have shown that 15 member ensembles of dynamical forecast using ECMWF Variable Resolution Ensemble Prediction System (VarEPS) display some prediction skill during onset phase of Indian monsoon with one month lead time.

Importance of extended range prediction has been recognized by the Ministry of Earth Sciences, Government of India and it has emphasized on the development of an ensemble System for Dynamical Extended Range Prediction of Active-Break Spell (ERPAS) of Indian Summer Monsoon Rainfall (ISMR) under the National Monsoon Mission (NMM) program. Extended Range Prediction group of IITM has provided real-time operational forecast of the Active-Break spells of ISMR with 4 pentad lead using an indigenously developed ensemble prediction system (EPS) based on Climate Forecast

System Version 2 (CFSV2). In this report, the general performance of the EPS in monsoon simulation and extended range operational prediction during 2011 monsoon season has been evaluated. Additionally, this report discusses about the developmental aspects of an Ensemble Prediction System (EPS) based on NCEP Climate Forecast System.

2. Data and Methodology

2.1 Data

Both atmospheric and oceanic initial conditions were obtained from National Climate Data Centres (NCDC) NOAA Operational Model Archive and Distribution System (NOMADS) on-line model data server (http://nomads.ncep.noaa.gov). For verification purpose, we have used the gridded rainfall dataset (Rajeevan et al, 2006) and the TRMM-gauge merged rainfall dataset (Mitra et al, 2009) from India Meteorological Department (IMD).

2.2 Model

Here, we have used the NCEP climate forecast system version 2 (CFSv2; Saha et al, 2012) for the real-time extended range forecasts. The atmospheric component of the model is NCEP Global Forecast System (GFS) at T126 resolution (approximately 100km) with 64 vertical levels, which is coupled to the GFDL MOM4 ocean model, sea-ice model and Noah land model. For its ocean component, the CFSv2 uses the GFDL MOM version 4p0d (MOM4; Griffies et al. 2004) with zonal resolution of 1/2° and meridional resolution of 1/4° between 10°S and 10°N, gradually increasing through the tropics to 1/2° poleward of 30°S and 30°N. There are 40 layers in the vertical with 27 layers in the upper 400 m of the ocean, and the maximum depth is approximately 4.5 km. The vertical resolution is 10 m from the surface to 240 m depth, gradually increasing to about 511 m in the bottom layer. The initial conditions have been prepared from coupled data assimilation system (CDAS) with T574L64 resolution atmospheric assimilation and MOM4 based oceanic assimilation.

2.3 Perturbation method

Dynamical extended range prediction is subject to various sources of errors, of which the errors arising from the uncertainties in the initial conditions and model are important. To eradicate such errors, ensemble prediction approach is one of the best options and has been attempted by various researchers. The EPS of ECMWF evolved from 50 initial perturbations generated using singular-vector technique and an unperturbed control run (Buizza and Palmer 1995). At NCEP, the ensemble of initial perturbations are generated in a similar way as at ECMWF, but breeding vectors (Toth and Kalnay, 1993) are used instead of singular vectors. Later Magnusson et al (2008) reported that breeding vector based EPS system outperforms compared to singular vector based EPS over the tropics. At Meteorological Service of Canada (MSC) an Ensemble Kalman Filter (EnKF) with perturbed observations and different combinations of parameterization schemes has been used to generate ensemble of initial conditions for their medium range prediction (Houtekamer et al. 1996, 2005). Though there are several approaches to generate ensembles of different initial conditions, we use an approach which is similar to the "complex-and-same-model environment group" as classified in Buizza et al. (2008).

It is well known that the skill and spread of an EPS essentially depends on the ensemble size (Richardson 2001; Reynolds et al. 2011). However, running large ensemble members in real-time basis require huge computational power and most of the operational centers limit the ensemble size to few tens. In this analysis, an ensemble of 10 perturbed atmospheric initial conditions has been developed in addition to one actual initial condition.

Each ensemble member is generated by slightly perturbing the initial atmospheric conditions with a random matrix (random number at each grid point) generated from a random seed. In order to make the perturbation size consistent with analysis variance of each variable, the amplitudes of perturbations are adjusted to ensure sufficient spread in the forecast fields. It is also ensured that amplitude of perturbation varies in accordance with the uncertainty in the analysis. In order to make the ensemble mean to be centered

around the unperturbed analysis, fraction of the difference between the short range forecast and the analysis of different model variables are added to or subtracted from the unperturbed analysis, with random perturbation between -1 and +1 times the difference between the short term forecast and analysis, so that the perturbations follows a Gaussian distribution. We perturb the wind, temperature and moisture fields and the amplitude of perturbation for all variables are scaled according to the magnitude of each variable at a given vertical level. Main advantage of this system is its potential to generate infinite number of ensembles. Also the amplitude of perturbations can be adjusted by changing the tuning factor. In the present formulation, there is option to select any one of the variable, u-component of wind (u), v-component of wind (v), temperature (t) or humidity (q). This approach also helps to study the individual impact of perturbing each variable on the final forecasts.

Fig 1 a, b and c shows the actual 850 hPa zonal wind (u850) analysis, perturbed initial condition, and perturbation valid for the same analysis time. It is noted that there is not much difference between the perturbed and actual analysis fields. Note that the differences are rather small over most parts, except some places where large climatological variance of daily analysis field exists. There exist noticeable difference between the perturbed and actual analysis (see lower panel of Fig 1) around extra tropics. This is due to the fact that magnitude of the wind field is higher in the extra tropics, and that we adopted a perturbation strategy in which the magnitude of perturbation is proportional to the difference between the short range forecast and analysis.

The impact of perturbation factor on forecast field is assessed by plotting the time series of precipitation over core MZI covering central India for 45 day forecast lead. Four distinct initial conditions have been selected here to evaluate the growth of spread under two different atmospheric basic states. Fig 2 a -d shows time series of the area averaged precipitation forecasts over MZI for 45 days starting from 16 May, 15 June, 15 July and 4 August 2011. The top left panel of Fig 2 represents the growth of spread from forecast starting from an initial atmospheric condition much before the onset of monsoon. Figure 2 presents the forecast started from an entirely different atmospheric state in

which monsoon circulations were already established. The solid lines represent the ensemble mean and spread among members is given as vertical bars. Obviously, the growth of spread and hence the predictability varies with different atmospheric initial conditions. For 16 May initial condition, the dispersion started growing during early forecast hours at a lead of 15 days, whereas from other initial condition, the dispersion in the forecast started increasing from a forecast lead, much earlier than 10 days. Another contributing factor influencing the small spread during early forecast leads from 16 May initial condition is the relatively less amount of rainfall over central India during May and early June.

2.4 MISO index method

The MISO index method has been proposed by Suhas et al. (2012) to monitor the temporal evolution and amplitude of ISO as it evolves from Equatorial Indian Ocean to foothills of Himalayas. This is an EOF analysis similar to Wheeler and Hendon (2004) but with a difference that the EOF analysis is performed on an extended data matrix. The extended data matrix is constructed by temporal embedding of a number of lagged copies of the data that are appended to the original data to create the extended matrix. The number of lags to be embedded depends on the dominant periodicity needed to be resolved and will filter out higher frequency variability. Since the dominant periodicity of the MISO in rainfall lies between 24 and 40 days, a lag of 15 days with 1-day delay was found to be useful to resolve the smooth evolution of the MISO. The EOFs are constructed based on rainfall data averaged over the longitudinal band 60°-90°E spanning the latitudes 12°S-30°N for the 122 days of years 1998-2011. The EOF1 and EOF2 explain a combined variance of 23% and is separated from rest of the EOFs. The two principal Components (MISO index 1 and MISO index 2) are obtained by projecting the daily data matrix with delay embedding as defined earlier. We will use here the MISO indices for the year 2011 only. For further details on EOF analysis of ISOs, Suhas et al., (2012) and Wheeler and Hendon (2004) are referred.

3. Results and discussion

3.1 Hindcast skill

This section is devoted exclusively for the various verification measures to quantify the extended range prediction skill of the present EPS. Though, we developed a global EPS, our forecast skill analysis mostly focused over Indian region at pentad intervals. The pentad prediction skill may be considered as total intraseasonal variability (ISV) prediction skill, and is a more rigorous way of evaluating the model's hindcast skill. The verification skill scores are shown for ensemble mean forecast as deterministic and also considering the individual member forecast for probabilistic skill measures.

Various skill scores were also evaluated and presented in figure 3. The Kupper skill score (KSS) and bias score proposed by Hanssen and Kuiper (1965) is shown in figure 3 (a, b) and found that normal category is always over forecasted for all lead pentads. It is also evident that skill in predicting break is higher followed by active category and then normal. The normal category shows lowest skill compared to other two categories. Analysis of Heidke skill score (HSS) and Gerity score shows that CFSV2 show slightly better fractional improvement over random chance (fig 3 c) and skill is significant even at pentad 4 lead. However, the correlation coefficient of the ensemble deterministic forecast shows that the correlation values are less significant beyond pentad 3 lead (fig 3 d). It is found from the month wise skill analysis shown in figure 3(e) that the EPS has relatively little skill in the month of July and August and higher skill during the early and later phases of monsoon season during June and September.

An assessment of value added forecast skill also developed based on the dynamical ensemble forecasts. Categorical forecast skill score is computed by using discrete forecasts grouped into different classes based on a threshold value. Different possibilities are thus obtained from a range of several threshold values from a 2x2 contingency table in which forecast-observation pairs are classified into four different groups (Wilks, 1995). Traditional precipitation forecast verification scores such as false alarm rate (FAR) and hit rate is used to analyze categorical forecast skill over MZI region. Here we choose relative operating characteristic curves (ROC) to asses the forecast skill by defining the dichotomous event such as occurrence or non occurrence of

the precipitation above or below certain thresholds. Basically ROC combines the hit rate and false alarm rate. The ROC has been generated from 11 member ensemble forecast. The ROC has been gained widespread acceptance as a tool for probabilistic as well as ensemble forecasts verification (Mason and Graham 1999, Hamill et al. 2000, Hamill and Juras, 2006).

In the present analysis, ROC for three observed categories is evaluated. The three observed categories are defined in similar manner as discussed in the previous sections. The 11 member ensembles prediction of area averaged percentage departure is sorted from lowest to highest and which is then converted to yes/no forecast by comparing it with the observation and assigned a binary value (1 for correct forecast case and 0 for not forecasted). Here three categories of observations are also converted to binary value of 1 or 0 according to the percentage departure of precipitation. The area under the ROC curve can be used for the calculation of a probabilistic skill score. Then separate 2x2 contingency tables are calculated for each sorted ensemble member with different probabilities. In this study the forecast distribution from ensemble system is arranged into 10% wide probability range bins so that total 10 probability classes are obtained. As for the deterministic ROC score, the Probabilistic ROC score is simply calculated using the Hit Rate (HR) and False Alarm Rate (FAR). The HR and FAR are calculated for each probability interval. In this case an event is said to be forecast at a point if the forecast probability for an event occurred within the probability range. (e.g., a forecast for above median rainfall that had a 43% probability would fall in the 40-50% probability range.) Observed occurrences (i.e., Hits) are then the number of times that a forecast probability fell in that bin and subsequently that event occurred (in the example case above median rainfall occurred), while the observed non-Occurrences (i.e., Misses) are the number of times a forecast was made for that probability bin but the forecast was "incorrect" (in the example case below median rainfall occurred). More details on generating probabilistic ROC score can be found from Kharin and Zwiers (2003). The curve against hit rate and false alarm rate (ROC) for all 10 probability thresholds for 1^{st} to 4^{th} pentad lead forecast is plotted in figure 4.

The area under the ROC is evaluated for three categories namely, break, normal and active cases. It is evident from the figure 4 (a) that at all lead pentads, the EPS has

higher skill in predicting break followed by active and then normal category. It is also to be noted that the skill in predicting break for different probabilities are almost same and different probabilities are clustered around. Skill of predicting active and normal at different probabilities shows a range of values with slightly less skill compared to break prediction. One interesting thing is that curve for normal for pentad 3 and 4 leads more or less coincides with the diagonal and in other words, it has equal chance of hit and false alarm rate and considered as little or no skill at all probability ranges. At pentad 4 lead, the skill in predicting break and active exhibits similar hit and false alarm rate. Obviously there is relatively little skill at pentad 4 forecast for all three categories (see Fig. 4d). However, the EPS shows that the hit rate is slightly higher than false alarm rate and probably there exists some skill in the probabilistic prediction from ensemble system. It may be concluded that skillful probabilistic categorical forecast can be generated from the EPS even at pentad 4 lead. Of course the skill is likely to probably depend on the ensemble size (Kumar et al., 2001). In general the forecast skill deteriorates with the forecast lead.

3.2 Onset forecast using the EPS

Despite the late arrival of monsoon over Andaman Sea, monsoon onset over Kerala (MOK) took place on 29th May 2011, according to IMD bulletin. Operationally, IMD has been determining the MOK in a qualitative and subjective approach by considering various factors like persistent wide spread rainfall spatially over Kerala for two days and the lower tropospheric westerly wind. Following a criteria similar to the IMD, we used the 45 day deterministic dynamical forecasts from the EPS for identifying the date of MOK qualitatively. We used two indices namely rainfall over Kerala (ROK), the area averaged rainfall covering Kerala and adjacent eastern Arabian Sea (74°-78°E, 5°-10°N) and zonal wind over eastern Arabian Sea (U850_ARB) defined as the area averaged zonal wind at 850 hPa over eastern Arabian Sea (60°-70°E,0°-10°N). In this approach, the onset date is defined as the day in which the ROK and U850_ARB index exceeds 60% of model climatological value, provided the forecasted ROK is greater than 5 mm and U850_ARB also exceeds 12 knots consecutively for two or more days. As the

long-term mean date of MOK is around 1st June with a standard deviation of about 8 days (Ananthakrishnan and Soman, 1988), we fixed 16th May Initial condition for identifying the date of MOK. Figure 5 shows the time series of ROK and U850_ARB starting from 16th May 2011. It is evident from the figure that the forecasted date of MOK was 31st May 2011. Our onset forecast suffered an error of 2 days, as the actual onset date declared by IMD was 29th May 2011.

3.3 Extended range forecast and verification.

As the main objective in this study is to provide the real-time extended range prediction, forecast has been issued at five day intervals up to 4 pentad lead. Figures 6 and 7 shows some selected snap shots of the pentad rain anomaly forecast over Indian region along with the area averaged pentad rain anomaly over MZI region. The top two panels in the figures represent the spatial rain anomaly maps for up to 4 pentad lead. Here we choose 2 initial conditions arbitrarily in which monsoon circulations were active. Based on synoptic reports, there had been a steady advance of monsoon during 15-26 June in association with the formation of deep depression (16-22 June) over the northwest Bay of Bengal (BoB) and its gradual west-northwestward movement. This feature is well captured in P1 lead forecast (Figure 6) and positive rain anomalies were forecasted over MZI. The depression over land during 22-23 July was well predicted in P1 lead (Figure.7). Detailed discussion on the spatial verifications for the pentad lead forecasts is given in the next section.

Pentad rain anomaly forecast is verified against both merged gridded IMD rainfall analysis combining TRMM observations and rain gauge (IMD_MRG). The pentad rain anomaly forecast up to 4 pentad leads valid for same forecasted pentad were compared with corresponding rain anomaly from both IMD and TRMM observations. Figure 8 shows the pentad rain anomaly forecast valid for 20-24 June, during which a deep depression has been formed over central India. The model could predict the rainfall associated with the deep depression well up to pentad 2 lead. However, in pentad 3 lead forecast rainfall is more shifted towards south peninsular region and pentad 4 lead

forecast predicted more false alarms over peninsular India, Arabian Sea and Bay of Bengal region.

The formation of a low pressure system over south Chhattisgarh and adjoining Telangana region, and an offshore trough extending from Gujarat coast to Kerala coast helped southwest monsoon to cover the entire country by 9th July, almost 6 days in advance of the normal date. Figure 9 shows the pentad rain anomaly forecast valid for 5-9 July. This feature is reasonably well captured in pentad 1 lead forecast. On the other hand, all other lead forecast from pentad 2 to pentad 4 could not predict any such signal as seen in the observations. Figure 10 shows another case of depression formed over land during 22-23rd July. The first two lead pentad forecasts are able to capture both land and oceanic convection over south BoB. However, pentad 3 and 4 lead forecasts well captured the land rainfall, but predicted some false alarms over Arabian Sea and south BoB. Figure 11 represents another case of a low pressure area during 8-11 August which formed over the western end of monsoon trough over northwest Madhya Pradesh and neighborhood, leading to wide spread rainfall over northwest Madhya Pradesh, Gujarat and Rajasthan. This system is well captured in pentad 1 lead forecast followed by pentad 2 lead with slightly lower intensity and a northwest shift in the rainfall location. Nevertheless, pentad 3 and 4 lead forecasted rainfall were shifted more towards south east Madhya Pradesh. Thus, from figures 8-11, it may be noted that the depressions were well predicted in pentad 3 and sometimes pentad 4 leads, whereas, lows are least predicted even in pentad 2 lead. Additional rigorous statistical analyses are required to arrive at a solid conclusion.

Although the ensemble members are limited to 11, we made an attempt to issue the probability forecast from the EPS. As deterministic predictability limit in NWP range reaches within 10 days, issuing deterministic forecast alone will not add value to forecast in the extended-range time scale beyond 10 days. Figure 12 a-d presents the ensemble mean deterministic forecast and corresponding verification from IMD analysis (upper panel of each figure) and probability for forecast for tercile categories from pentad 1 to 4 lead (lower panel of each figure). The categories of forecast are defined in such a way

that if the area averaged rainfall is greater than 40% of its long term mean then it is considered as active period, if it is less than -40% then it is break period, and otherwise a normal period. These categories are defined independently for EPS forecast and observations. It is clear that the probability of three categories favored the ensemble mean rainfall anomaly. It is noted from the figure that the for pentad 1 and pentad 2 leads, the probability forecasts favors ensemble mean forecast as spread among members are small up to 10 days. Pentad 3 and 4 lead forecast presents wide range of probability values especially when the ensemble mean forecasted normal category and higher or 100% probability when the ensemble mean forecasted categories are either active or break. This means that all the members favor ensemble mean forecast when deterministic ensemble mean forecast was on two extreme sides. It is speculated that small ensemble size of 11 members or initial perturbation size might have an impact on the probability forecast. However, a wide range of probability values can be achieved only by increasing the ensemble size.

3.4 MISO Index Monitoring and Hindcast.

The **MISO** index obtained based on Suhas et al., (2012) defines the eight-phase evolution of ISO similar to that of Madden-Julian Oscillation, described in Wheeler and Hendon (2004). The eight-phase evolution of ISO is shown in Figure 13. The phase space representation of ISO for the year 2011 from 15June-30September, based on the two MISO indices is shown in Figure 14. From the figure, two strong MISO spells (when both the indices are greater than one standard deviation) could be identified: the first one from mid-June to last week of July and the other one during first half of September. However, it is to be noted that the peak amplitudes are not present over central India. Most of the strongest amplitude of the MISO indices occurs over northern and foothills of India and southern peninsular India. Thus the ISO does not have strong projection over central India and the rainfall amplitude over this region. Thus the contribution of net rainfall over MZI region through ISO is not significant in 2011. Another important feature is that the strong spells have small curvature with smooth turning whereas the weak spells within or close to the unit circle shows abrupt turning and has large

curvature. The existence of large curvature implies sharp jumps in rainfall spatial patterns within a short span of time in the regions used for the construction of EOF (here 12°S-40°N). Due to this, it may be difficult to predict the days associated with such evolution pattern. On the other hand, strong events have small curvature and the temporal evolution may be smooth enough to be predicted with good skill.

In order to see the skill of CFSv2 in capturing the phases of ISO and the forecast of MISO indices for the year 2011, the longitude averaged forecasted rainfall from any forecast start day is projected onto the EOFs obtained from observation for each of the ensemble members and for the ensemble mean. Then the next three-week evolution of the indices and the ISO phases are compared with observation. The evolution of MISO from some selected initial dates is depicted in Figure 15. The top panels show the evolution identified as *best* forecast while the bottom panel shows the evolution identified as *worst* forecast. The *best* and *worst* forecasts are categorized as follows: we computed the amplitude of MISO for both the observation and forecast i.e. square root of [$index1^2$ +$index2^2$]. The RMSE of 21 days evolution for each of the 18 forecast initial conditions are computed and the 18 RMSE values are ordered in ascending order starting from lowest RMSE value. So the 3 forecasts with lowest (highest) RMSE value are categorized as *best (worst)* forecast. It is interesting to note that both *best* and the *worst* forecasts follow the evolution of the observed ISO i.e. they are over the same region in the graph, however they significantly differ in intensity in the *worst* cases with large overestimation of MISO amplitude by the CFS forecast. In order to see the spatial propagation of rainfall for the three *best* and *worst* cases we plot the time-latitude Hovmuller plots from model and observation in Figures 16 and 17 respectively. Interestingly, the *best* cases show weak propagation and *worst* cases show better propagation. However, this becomes clear if we compare with Figure 15. It may be seen that all the *best* forecast cases show weak MISO amplitude (close to the unit circle). The *worst* cases show real northward propagation (observed panel in the right) and the amplitude of MISO for these initial dates are under-predicted (over estimation of amplitude by the CFS forecast). Thus, the model shows clear amplitude bias for the strongly propagating MISOs.

In order to see the growth of ISO forecast error, we plot the correlation of MISO Index-1 (M1) and MISO index-2 (M2) as a function of lead-time calculated for all the

forecasts (Figure 18a). It may be seen that the forecast shows significant correlation (>0.5) for 15 days with observation. Also we plot the absolute error of the indices defined as: $ABS[\left[M1^2 + M2^2\right]_{OBS} - \left[M1^2 + M2^2\right]_{CFS}]$ in Figure 18b. This quantity brings out the growth of errors from each initial condition (abscissa) as a function of lead day (ordinate). It may be seen interestingly that the error growth is low for a month as we start from June initial condition and then increases rapidly towards the end of the season. Thus it may be concluded that there are some initial conditions within the season from which the error growth is large and there are some from where the error growth is less. Such inhomogeneous distribution of error growth needs further investigation.

3.5 Simulation of seasonal monsoon rainfall

As the main objective of the study is to evaluate the performance of an extended range ensemble prediction system at pentad scale during 2011, seasonal mean monsoon precipitation at pentad lead is computed and compared with IMD_MRG. Figure 19 shows the seasonal mean monsoon precipitation from observation and its difference up to 4 pentad lead during 2011 monsoon season. The JJAS mean precipitation from IMD_MRG in the top panel of the figure shows that there exist few local pockets of precipitation maxima over north and eastern Bay of Bengal, the west coast of India, south of equatorial Indian Ocean and over maritime continent and western Pacific. Figures 19 b-g presents the difference of the JJAS mean precipitation at pentad lead forecast and observation up to pentad 4 lead. At pentad 1 lead, large overestimation of precipitation is found over west coast of India, Bay of Bengal (BoB) and along the mean monsoon trough region along Gangetic plains near foothills of Himalaya. While large underestimation is found over Indian land mass and central equatorial Indian Ocean. At pentad 2 and 3 lead, the overestimation over Arabian Sea shifted to more west and patch of overestimation over south of the Indian and SriLankan coast shifted to the equatorial central Indian Ocean. It is also noted that pockets of little overestimation appeared over peninsular India. At pentad 4 lead, it is found that there exists a dipolar structure in the precipitation with large overestimation over equatorial Indian Ocean and underestimation over Indian land mass and adjoining eastern Arabian Sea and BoB.

Figure 20 shows the departure in percentage of the seasonal rainfall during 2011 JJAS season from observation and CFSV2 forecast for pentad 1 to 4 leads. Following the official IMD criteria for classifying the seasonal rainfall into different categories, the seasonal rainfall departure is classified as scanty (% departure between -60 and -100), deficient (% departure between -20 and -59), normal (% departure between -19 and +19) and excess (% departure above + 19). It is found that p1 lead forecast very well captures the spatial pattern of rainfall departure and is comparable to IMD merged analysis. At pentad 2 lead, the region of excess rainfall departure over central and western India expands in area, whereas the deficient departure region over peninsular India shrink in area. At pentad 3 lead, the overestimation further spread to peninsular India and the deficient region has completely disappeared. However at pentad 4 lead, the spatial rainfall departure pattern improved compared to pentad 3 lead. The observed rainfall departure pattern over central and western India is very much similar to pentad 1 lead and better than pentad 2 and 3 lead, as compared to the observation.

3.6 Simulation of northward propagation

In this subsection, we examine how well the model forecasted the northward propagation of active/break spells occurred during JJAS 2011. The active and break spells were identified from figure 21, which shows the pentad rain anomaly over MZI along with +1/-1 standard deviation shows in yellow shades. Figures 22-25 portrays the latitude-time cross-section of rainfall anomalies, averaged over Indian longitudes 70°-90°E, corresponding to the three active spells (16-25 Jun, 10-14 Aug and 25 Aug-08 Sep) and one break spell (01-05 Jul) respectively. It is clear from the figures that the model could forecast the temporal evolution of all these spells realistically, one pentad earlier. In the case of first two active spells (starting from 16 June and 10 August respectively), the model could predict the spells in almost all pentad leads (Figure 22 and 23). Regarding the third active spell, the model failed to capture the spell 2-3 pentads in advance (Figure 24 d and e). The model could predict the break spell starting from 01 July in almost all pentad leads, except the third one (Figure 25).

3.7 Seasonal cycle of monsoon indices at pentad forecast leads

The seasonal cycle of pentad lead forecast up to 4 pentads were compared with observation and shown in Figure 26. We selected two precipitation indices which exhibit reasonable active-break cycle as described in Krishnamurthi and Shukla (2001) and Chattopadhyay et al (2008). Two monsoon rainfall indices were calculated using analysis and also for different pentad forecast leads. To asses the precipitation forecast over India and adjoining oceanic regions, following two rainfall indices were used, monsoon zone India rainfall index (MZI) defined as the area averaged rainfall over MZI region (Rajeevan et al, 2006) and extended Indian monsoon Rainfall (EIMR) index defined as the precipitation area averaged over India and adjoining oceanic region covering the Bay of Bengal and Arabian sea (70°E-110°E, 10°N-30°N; Goswami et al., 1999).

Pentad lead seasonal cycle from EPS, plotted in Figure 26 along with observation, could capture the seasonal cycle of both rainfall index MZI and EIMR. It is found that during June to early July, rainfall forecast over MZI is underestimated as compared to analysis (solid line in figure 26 a). Though the pentad lead forecast could capture the seasonal cycle, it relatively underestimated the rainfall over MZI region for pentad 1 to 4 leads during June, July and August. However, it is comparable in magnitude for all lead pentads during September. The EIMR index is overestimated during the early phase of monsoon during June and underestimates in the later phase during September. This is due to the fact that the EIMR index covers a large portion including Indian land area, eastern Arabian Sea and Bay of Bengal and deficit rainfall forecast over Indian land area is compensated by relatively overestimation over Arabian Sea and Bay of Bengal (see figure 19). The underestimation is due to the above normal active spells of monsoon rainfall during the entire September month of 2011 monsoon season and the inherent dry bias of the model.

3.8 Pentad forecast skill

Forecast verification and skill analysis is an essential component of the ensemble prediction system. Correlation coefficient (CC), a common measure of the deterministic forecast skill has been computed over MZI region. Model anomaly is calculated from model forecast climatology corresponding to each start date (e.g. 1 May, 6th May, 11 May, etc). The correlation coefficient of area averaged rainfall over MZI is presented in figure 27 along with the corresponding RMSE. As a base line comparison, the hindcast skill is also included in the figure in black shade. The CC for precipitation is computed against the IMD merged rainfall analysis interpolated to model resolution. During 2011 monsoon season, the EPS exhibits significant correlation values during the first two lead pentads. However, the CC has decreased to 0.2 at pentad 3 lead and increased to 0.4 in the 4^{th} pentad lead. This is also evident from the RMSE values which increases with forecast leads and reaches a large value and then decreased in the 4^{th} pentad lead forecast.

Figure 28 presents the anomaly pattern correlation (left panel) and the corresponding forecast errors in terms of RMSE (right panel). Correlation coefficients significant at 0.05 level is highlighted in shades. At pentad 1 lead, most of the Indian region shows significant correlation values. However, ACC decreases sharply during pentad 2 lead, however parts of central, northwest India, northeast India and southern peninsular India shows significant correlation values. At pentad 3 lead, except for north India, all other regions show relatively small correlation which is not significant. Negative correlation coefficients are found over central India, which is also evident from the area averaged ACC presented in the previous figure. At pentad 4 lead, patches of positive values are found over northwest and southeast of Central India and region of negative values shifted more towards southern peninsular India. Hence over central India, pentad 4 lead exhibits higher correlation values than pentad 3 lead during 2011 monsoon season.

Large RMSE values are found over eastern Arabian Sea and Bay of Bengal. It is found that over these two regions, the forecast error increases with forecast lead and

saturates by pentad 2 lead. This structure is consistent with mean bias characteristics of the CFSv2 model. It is also seen from the figure that, the RMSE over Indian region increases with lead and saturates by pentad 3 lead and thereafter it decreases. This may be another contributing factor to the higher correlation values even at pentad 4 lead forecast. The deterministic forecast skill of the EPS during 2011 monsoon season is higher in first two pentads and 4th pentad lead and lower in 3rd pentad lead forecast.

4 Conclusions and Roadmap

An ensemble prediction system (EPS) has been developed in a coupled-atmosphere-ocean modeling frame work using CFSv2 model at T126 resolutions. The main advantages of the EPS developed in the CFSv2 model are that it is a reliable and flexible system, which is simple to implement in operational environment and capable of generating as many members, provided adequate computer resources are available. Several sensitivity experiments have been performed to test the accountability of the perturbation technique.

The seasonal cycle in precipitation as well as its northward propagation is captured reasonably well by the EPS system. The pentad lead seasonal mean rainfall errors shows large overestimation over equatorial Indian Ocean and underestimation over Indian land mass and the adjoining eastern Arabian Sea and BoB especially at pentad 3 and 4 lead. The observed JJAS seasonal rainfall departure pattern over central and western India is very much similar to pentad 1 lead and better than pentad 2 and 3 lead as compared with observation. Though the pentad lead forecast captures the seasonal cycle, it relatively underestimates the rainfall over MZI region for pentad 1 to 4 leads during June, July and August and comparable in magnitude for all lead pentads during September. However, the EIMR index is overestimated during the early phase of monsoon in June and relatively underestimates during later phase in September.

The forecast skill of EPS has been analyzed by computing the correlation coefficients and RMSE. The EPS could capture the spatial distribution of rainfall reasonably well, but the rainfall intensity is slightly underestimated. Analysis of various skill scores for categorical forecast suggests that the skill in predicting break is higher followed by active category and then normal. The ROC analysis of probabilistic

categorical forecast also confirms that the EPS has higher skill in predicting break followed by active and then normal category. It is found from the hindcast skill that that the EPS has relatively little skill in the month of July and August and higher skill during the early and later phases of monsoon season during June and September. The EPS exhibits good skill even up to 4th pentad lead and the forecast skill for 2011 monsoon season is slightly higher in the 4th pentad lead than the hindcast skill. The EPS performed better in the later phase of the monsoon during August and September than the early phase during June.

We also compared the MISO amplitudes obtained from CFS and observation. The MISO amplitude gives an idea of the amplitude of the dominant low frequency mode present in the atmosphere. Most of the strongest amplitude of the MISO indices occurs over northern and foothills of India and southern peninsular India. Thus for 2011 the observed ISO does not have strong projection over central India and the fluctuation of rainfall amplitude over this region may also be related to non-ISO components. Another important feature for this year 2011 is that the strong spells have small curvature with smooth turning whereas the weak spells within or close to the unit circle shows abrupt turning and has large curvature. Thus it indicates the change of rainfall amplitude is relatively sudden over certain region. The CFS forecast of MISO amplitude suggests some dependence of skill on the initial condition. June initial condition shows relative lesser growth rate of error than other months. Also the forecast skills of the 2 MISO indices (M1 and M2) are significantly higher to 2 weeks in advance. The temporal evolution of MISO amplitude for three cases defined as "*worst*" cases shows large overestimation of the amplitude. Analysis of skill for both MISO and MZI index reveals that CFS shows better skill in predicting large scale MISO as compared to actual rainfall forecast over MZI region for forecast leads of 2 pentad and above. However, as far as stake holders is concerned, predicting actual rainfall over MZI is very much important and confidence in predicting MISO should be translated in terms of actual rainfall both qualitatively and quantitatively.

Preliminary results show positive impact of EPS on the extended range prediction of Active-Break cycle of ISMR. The same procedure will be extended with different combinations of physical parameterizations. Each set of experiments will be run with

same 10 perturbed initial conditions by keeping same model configuration and physical parameterization but with two different cumulus schemes and also with stochastic physics perturbation. By doing so, the model uncertainty will be taken into account before producing the final forecast. Future work will also investigate the impact of the resolution in CFS version 2 and number of ensemble members on the EPS prediction skill, impact of various ensemble creation techniques on the design of EPS and finally the development of different types of hybrid forecasting strategies combining the dynamical and statistical approaches including probabilistic forecast based on EPS. Also the spread-error characteristics and confidence information from different EPS techniques will be investigated to design an optimum ensemble prediction system for the extended range prediction of active-break cycle of the ISMR. The forecast skill will be further evaluated by examining relevant dynamical and moist convective features associated with ISO.

Acknowledgment:

IITM is fully supported by the Ministry of Earth Sciences, Government of India, New Delhi. India. We are thankful to Dr. K. Ashok and Dr. A. Suryachandra Rao for the constructive comments and suggestions.

References:

Abhilash, S., Sahai A.K., Pattnaik, S., Goswami, B.N. and Kumar A., 2012 A dynamical ensemble prediction system for the extended range prediction of active-break cycle of Indian summer monsoon using a coupled model (Under review).

Ananthakrishnan, R. and Soman, M.K., 1988, The onset of the southwest monsoon over Kerala 1901-1980, Journal of Climatology, 8, 283-296.

Annamalai, H. and Sperber K.R., 2005, Regional heat sources and the active and break phases of boreal summer intraseasonal (30-50 day) variability, Journal of Atmospheric Sciences, 62, 2726-2748.

Buizza R, Bidlot J-R, Wedi N., Fuentes M., Hamrud M., Holt G. and Vitart F., 2007, The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System), Quarterly Journal of Royal Meteorological Society, 133, 681–695.

Buizza R., Palmer T.N., 1995, The singular-vector structure of the atmospheric global circulation, Journal of Atmospheric Sciences, 52, 1434–1456.

Buizza, R., Leutbecher M. and Isaksen, L., 2008, Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System, Quarterly Journal of Royal Meteorological Society, 134, 2051–2066

Chattopadhyay, R., Sahai A.K. and Goswami B.N., 2008, Objective identification of nonlinear convectively coupled phases of monsoon intraseasonal oscillation: Implications for prediction, Journal of Atmospheric Sciences, 65, 1549-1569.

Fu, X., Wang, B., Li, T., McCreary, J., 2003, Coupling between northward propagating intraseasonal oscillations and sea-surface temperature in the Indian Ocean, Journal of Atmospheric Sciences, 60(15),1733–1753.

Goswami, B.N., Krishnamurthy, V. and Annamala,i H., 1999, A broadscale circulation index for the interannual variability of the Indian summer monsoon, Quarterly Journal of Royal Meteorological Society, 125, 611–633, doi:10.1002/qj.49712555412.

Goswami, B.N., 2005, South Asian Monsoon: Intraseasonal Variability of the Atmosphere-Ocean Climate System, Eds. William K.M. Lau and Duane E. Waliser Chapter 2, Praxis, Springer, Berlin, Heidelberg, 19-61.

Goswami, B.N., Wheeler M.C., Gottschalck J.C. and Waliser D.E., 2011, Intra- seasonal Variability and Forecasting: A Review of Recent Research, The Global Monsoon System: Research and Forecast, 2nd Edition, World Scientific Publication Company in collaboration with WMO, vol. 5, 389-407

Griffies, Stephen M., Matthew J. Harrison, Ronald C. Pacanowski and Anthony Rosati, 2004, A Technical Guide to MOM4, GFDL Ocean Group Technical Report No. 5, Princeton, NJ, NOAA/Geophysical Fluid Dynamics Laboratory, 342 pp.

Hamil, T.M. and Juras, J., 2006, Measuring forecast skill: is it real skill or is it the varying climatology, Quarterly Journal of Royal Meteorological Society, 132, 2905–2923.

Hamill, T.M., Snyder, C. and Morss, R.E., 2000, A comparison of probabilistic forecast from bred, singular-vector and perturbed observation ensembles, Monthly Weather Review, 128, 1835–1851

Hanssen, A.W. and Kuipers W.J.A., 1965, On the relationship between the frequency of rain and various meteorological parameters, Meded. Verh., 81, 2–15.

Harrison, M.S.J., Palmer, T.N., Richardson, D.S., and Buizza, R., 1999, Analysis and model dependencies in medium-range ensembles: two transplant case studies, Quarterly Journal of Royal Meteorological Society, 125, 2487-2515.

Houtekamer P.L., Lefaivre L., Derome J., Ritchie H., Mitchell H.L., 1996, A system simulation approach to ensemble prediction. Monthly Weather Review, 124, 1225–1242.

Houtekamer P.L., Mitchell H.L., Pellerin G., Buehner M., Charron M., Spacek L. and Hansen B., 2005, Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations, Monthly Weather Review, 133, 604–620.

Jiang, X., Li, T. and Wang B., 2004, Structures and mechanisms of the northward propagating boreal summer intraseasonal oscillation, Journal of Climate, 17, 1022-1039.

Joseph, S., Sahai A.K. and Goswami B.N., 2009, Eastward propagating MJO during boreal summer and Indian summer monsoon droughts, Climate Dynamics, 32, 1139-1153, doi:10.1007/s00382-008-0412-8.

Krishnamurthi, V. and Shukla, J., 2001, Observed and model simulated interannual variability of the Indian monsoon, Mausam, 52, 133–150.

Krishnamurti, T.N. and Subrahmanyam, D., 1982, The 30-50 day mode at 850 mb during MONEX, Journal of Atmospheric Sciences, 39, 2088-2095.

Krishnamurti, T.N., Jayakumar P.K., Sheng , J., Surgi, N. and Kumar A., 1985, Divergent circulations on the 30 to 50 day time scale, Journal of Atmospheric Sciences, 42, 364-375.

Krishnan, R.,. Zhang, C and Sugi, M., 2000, Dynamics of breaks in the Indian summer monsoon, Journal of Atmospheric Sciences, 57, 1354-1372.

Kumar, A., Barnston, A.G. and Hoerling, M.P., 2001, Seasonal predictions, probabilistic verifications, and ensemble size, Journal of Climate, 14, 1671-1676.

Lau, K.-M., Wu, H.-T. and Yang, S., 1998, Hydrologic processes associated with the first transition of the Asian summer monsoon: A pilot satellite study, Bulletin of American Meteorological Society, 79, 1871–1882.

Magnusson, L., Leutbecher, M. and Källén, E., 2008, Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system, Monthly Weather Review., 136, 4092–4104.

Mason, S.J. and Graham, N.E., 1999, Conditional probabilities, relative operating characteristics, and relative operating levels, Weather and Forecasting, 14, 713–725.

Mitra, A.K., Bohra, A.K., Rajeevan, M.N. and Krishnamurti, T.N., 2009, Daily Indian precipitation analyses formed from a merged of rain-gauge with TRMM TMPA satellite derived rainfall estimates, Journal of Meteorological Society of Japan, 87A, 265-279

Molteni, F., Buizza, R., Palmer, T.N. and Petroliagis, T., 1996, The new ECMWF ensemble prediction system: methodology and validation, Quarterly Journal of Royal Meteorological Society, 122, 73-119.

Palmer, T.N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P. and Tribbia, J., 1993, Ensemble Prediction. Proc. ECMWF Seminar on Validation of Models over Europe, 7–11 September 1992, Reading, UK, Vol. 1, 21–66.

Pattnaik, D.R. and Kumar Arun, 2010, Prediction of summer monsoon rainfall over India using the NCEP climate forecast system, Climate Dynamics, Vol. 34, 557-572.

Rai, S. and Krishnamurthy V., 2011, Error growth in Climate Forecast System daily etrospective forecasts in South Asian monsoon, Journal of Geophysical Research, 116, D0310 doi:10.1029/2010JD014840.

Rajeevan, M., Bhate, J., Kale, J.D. and Lal, B., 2006, High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells, Current Science, 91, 296-306.

Ramamurthy, K., 1969, Monsoon of India: Some aspects of "break" in the Indian south west monsoon during July and August, Forecasting manual, Part IV, 18.3, India Meteorological Department, New Delhi, India, 13 pp.

Ramaswamy, C., 1962, Breaks in the Indian summer monsoon as a phenomenon of interaction between the easterly and subtropical westerly jet streams, Tellus, 14, 337-349.

Rashid, H.A., Hendon, H.H., Wheeler, M.C. and Alves, O., 2010, Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system, Climate Dynamics.

Reynolds, C.A., Mclay, J.G., Goerss, J.S., Serra, E.A., Hodyss, D. and Sampson, C.R., 2011, Impact of Resolution and Design on the U.S. Navy Global Ensemble Performance in the Tropics, Monthly Weather Review, 139, 2145-2155.

Richardson, D.S., 1998, The relative effect of model and analysis differences on ECMWF and UKMO operational forecasts. Proc. ECMWF Workshop on Predictability, ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Richardson, D.S., 2001, Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size, Quarterly Journal of Royal Meteorological Society, 127, 2473-2489.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Pan, H.L., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M., Ek, M., Meng, J., Yang, R., Dool, H.V.D., Zhang, Q., Wang, W., Chen, M., 2012, The NCEP Climate Forecast System Version 2, To be submitted to the Journal of Climate.

Sahai, A.K., Chattopadhyay, R. and Goswami, B.N., 2008, A SST based large multi-model ensemble forecasting system for Indian summer monsoon rainfall, Geophysical Research Letters, 35, L19705, doi:10.1029/2008/GL035461.

Sengupta, D., Goswami, B.N., Senan, R., 2001 Coherent intraseasonal oscillations of ocean and atmosphere during the Asian summer monsoon, Geophysical Research Letters, 28(21), 4127–4130.

Sikka, D.R., Gadgil, S., 1980, On the Maximum Cloud Zone and the ITCZ over Indian longitudes during the Southwest Monsoon, Monthly Weather Review, 108, 1840–1853.

Suhas, E., Neena, J.M. and Goswami, B.N., 2012, An Indian monsoon intraseasonal oscillations (MISO) index for real time monitoring and forecast verification, Climate Dynamics, doi:10.1007/s00382-012-1462-5.

Toth, Z. and Kalnay, E., 1993, Ensemble forecasting at NMC: The generation of perturbations, Bull. American Meteorological Society, 74, 2317–2330.

Toth, Z. and Kalnay, E., 1997, Ensemble forecasting at NCEP and the breeding method, Monthly Weather Review, 125, 3297–3319.

Vitart. F. and Moltreni, 2009, Dynamical Extended-Range Prediction of Early Monsoon Rainfall over India, Monthly Weather Review, 137, 1480-1492.

Waliser, D.E., Stern, W., Schubert, S. and Lau, K.M., 2003, Dynamic Predictability of Intraseasonal Variability Associated with the Asian Summer Monsoon, Quarterly Journal of Royal Meteorological Society, 129, 2897-2925.

Wang, B., 2005, Theory intraseasonal variability in the Atmosphere-Ocean climate System, Stern, W., Lau, K.-M. and Waliser, D.E., Eds., Springer-Praxis, 307-360.

Wang, W., Chen, M. and Kumar, A., 2009, Impacts of ocean surface on the northward propagation of the boreal summer intraseasonal oscillation in the NCEP climate forecast system, Journal of Climate, 22, 6561-6576.

Webster, P.J., Magaña, V.O., Palmer, T.N., Shukla, J., Tomas, R.A., Yanai, M. and Yasunari, T., 1998, Monsoons: Processes, predictability, and the prospects for prediction, Journal of Geophysical Research, 103, 14 451-510.

Wheeler, M.C., Hendon, H.H., 2004, An all-season real-time multivariate MJO Index: development of an index for monitoring and prediction, Monthly Weather Review, 132, 1917–1932.

Wilks, D.S., 1995, Statistical Methods in the Atmospheric Sciences. International Geophysical Series, Vol. 59, Academic Press, 464 pp.

Xavier, P.K. and Goswami, B.N., 2007, An analog method for real time forecasting of summer monsoon subseasonal variability, Monthly Weather Review, 135, 4149-4160.

Yasunari, T., 1979, Cloudiness fluctuation associated with the northern hemisphere summer monsoon, Journal of Meteorological Society of Japan, 57, 227–242.

List of figures:

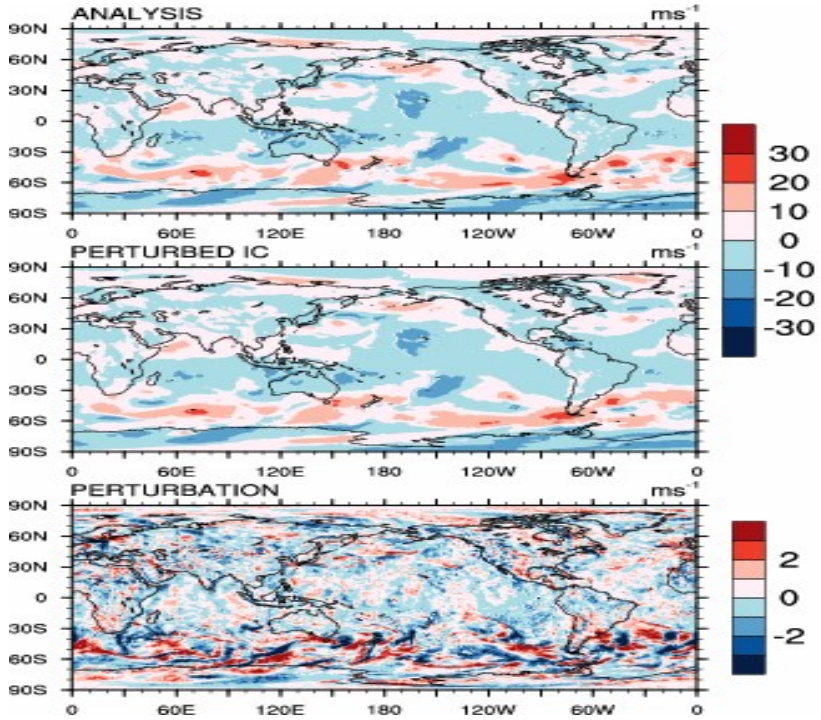


Figure 1: Zonal wind at 1000 hPa from (a) Analysis, (b) Perturbed initial condition and (c) Perturbation.

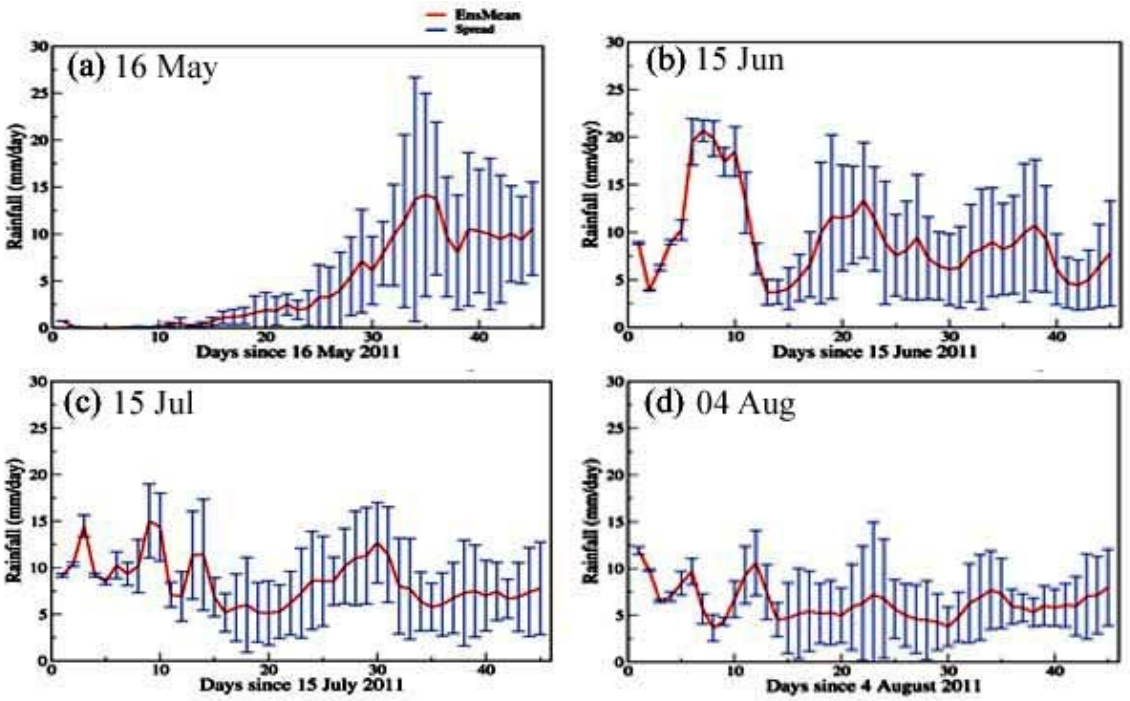


Figure 2: Time series of rainfall evolution starting from (a) 16th May, (b) 15th June, (c) 15 July and (d) 4th August 2011. The vertical bar represents the spread among members.

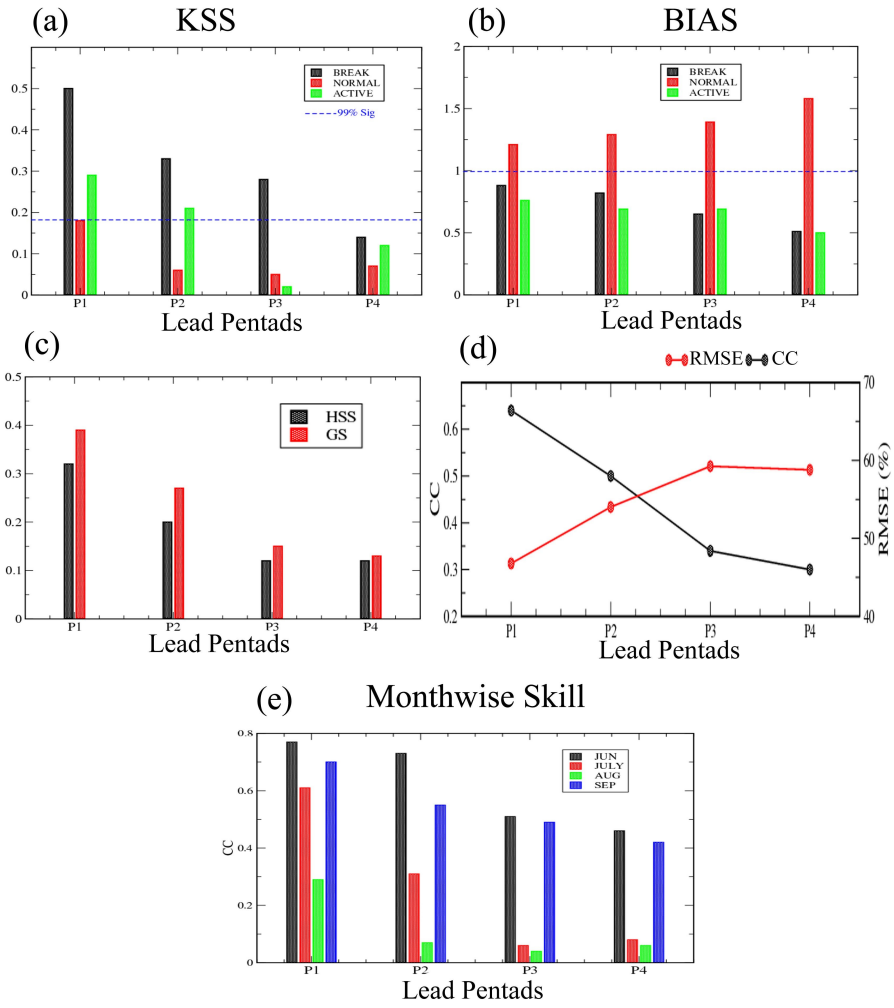


Figure 3: Pentad lead skill scores (a) KSS, (b) Bias, (c) HSS and Gerity score and (d) Month wise correlation coefficient

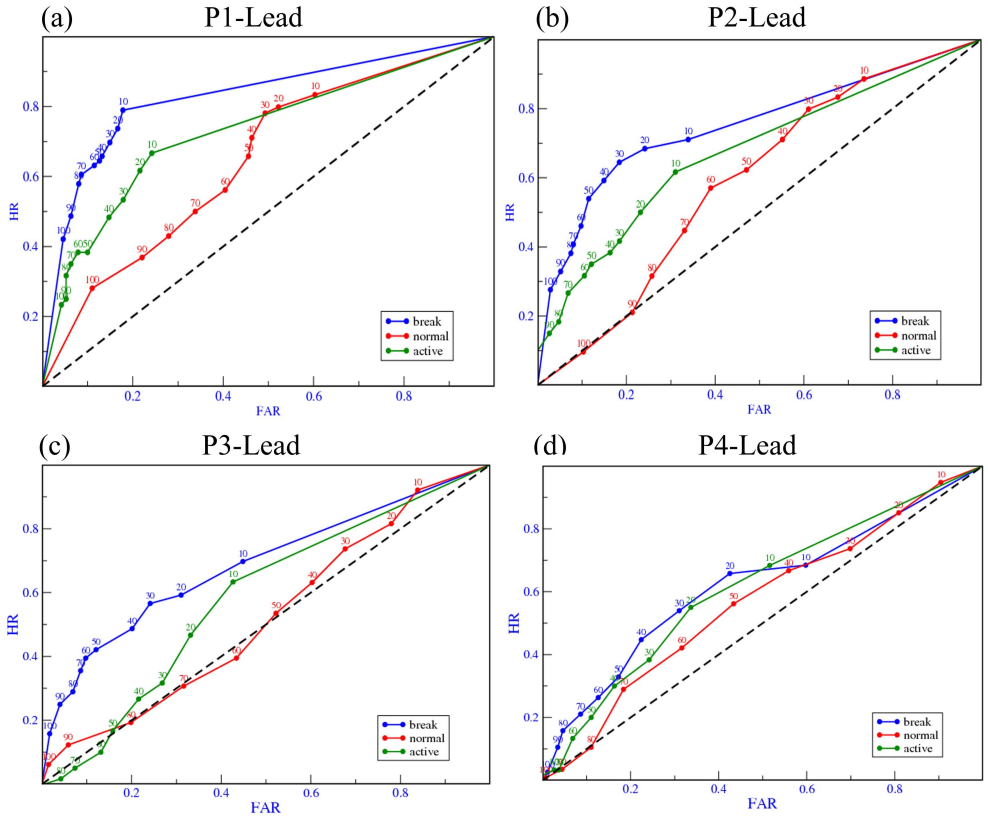


Figure 4: Pentad lead ROC for categorical probabilistic forecast

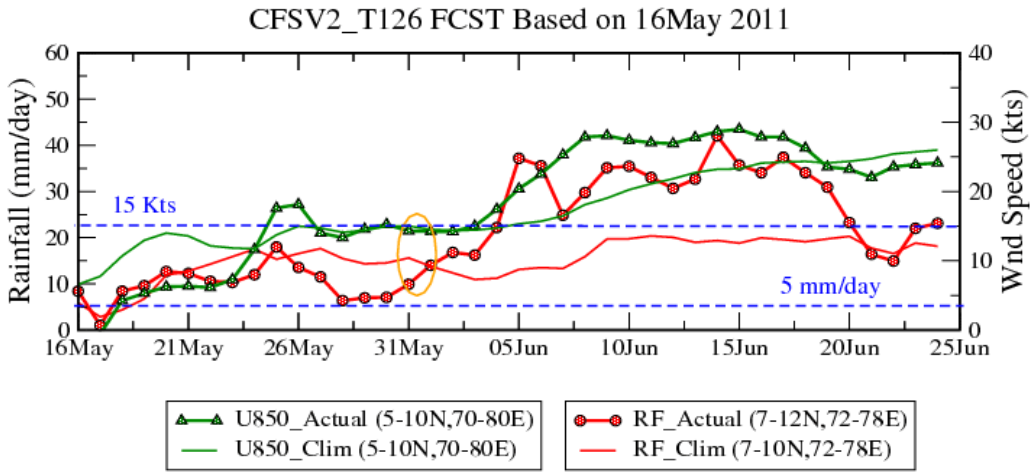


Figure 5: Evolution of area averaged zonal wind at 850 hPa and rainfall.

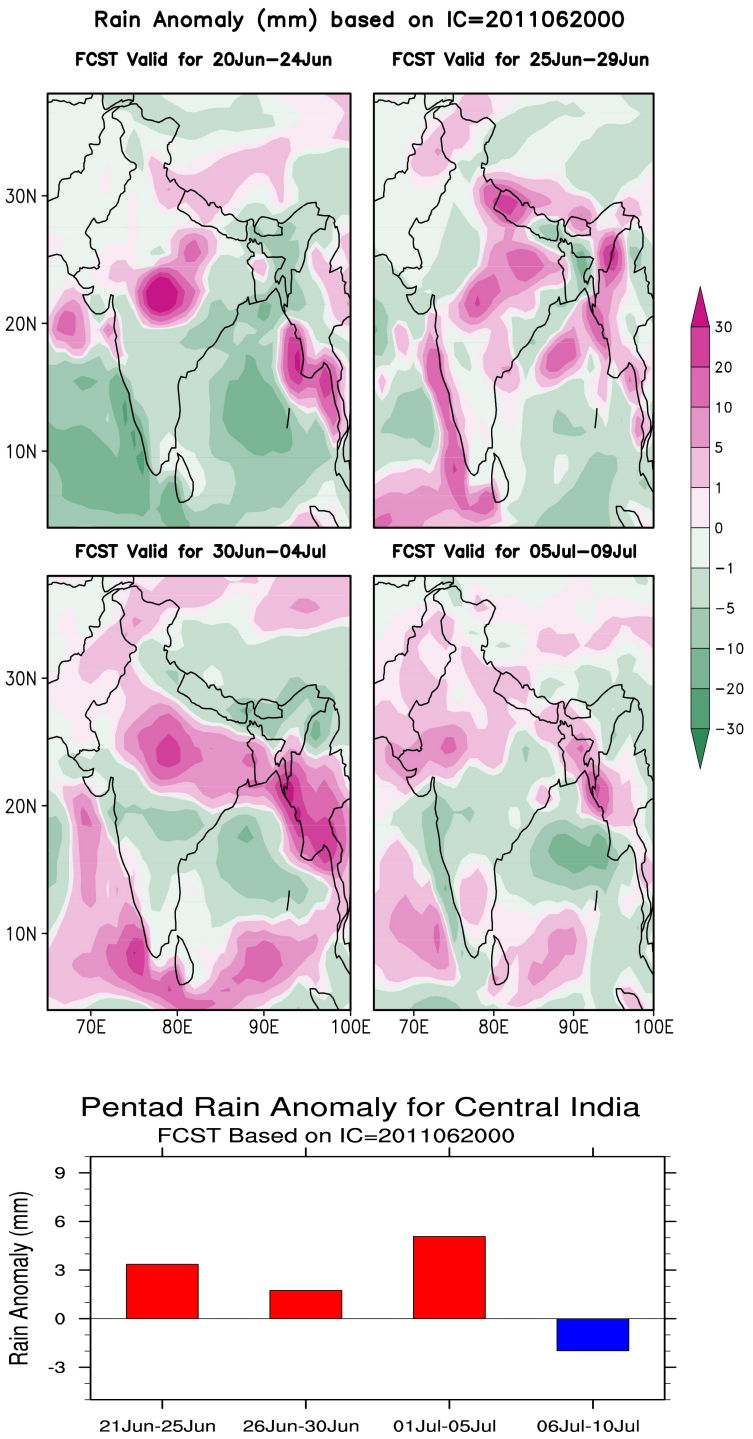


Figure 6: Pentad rain anomaly forecast from 20th June initial condition for up to 4th pentad lead (top panels), and corresponding area averaged rainfall forecast over MZI.

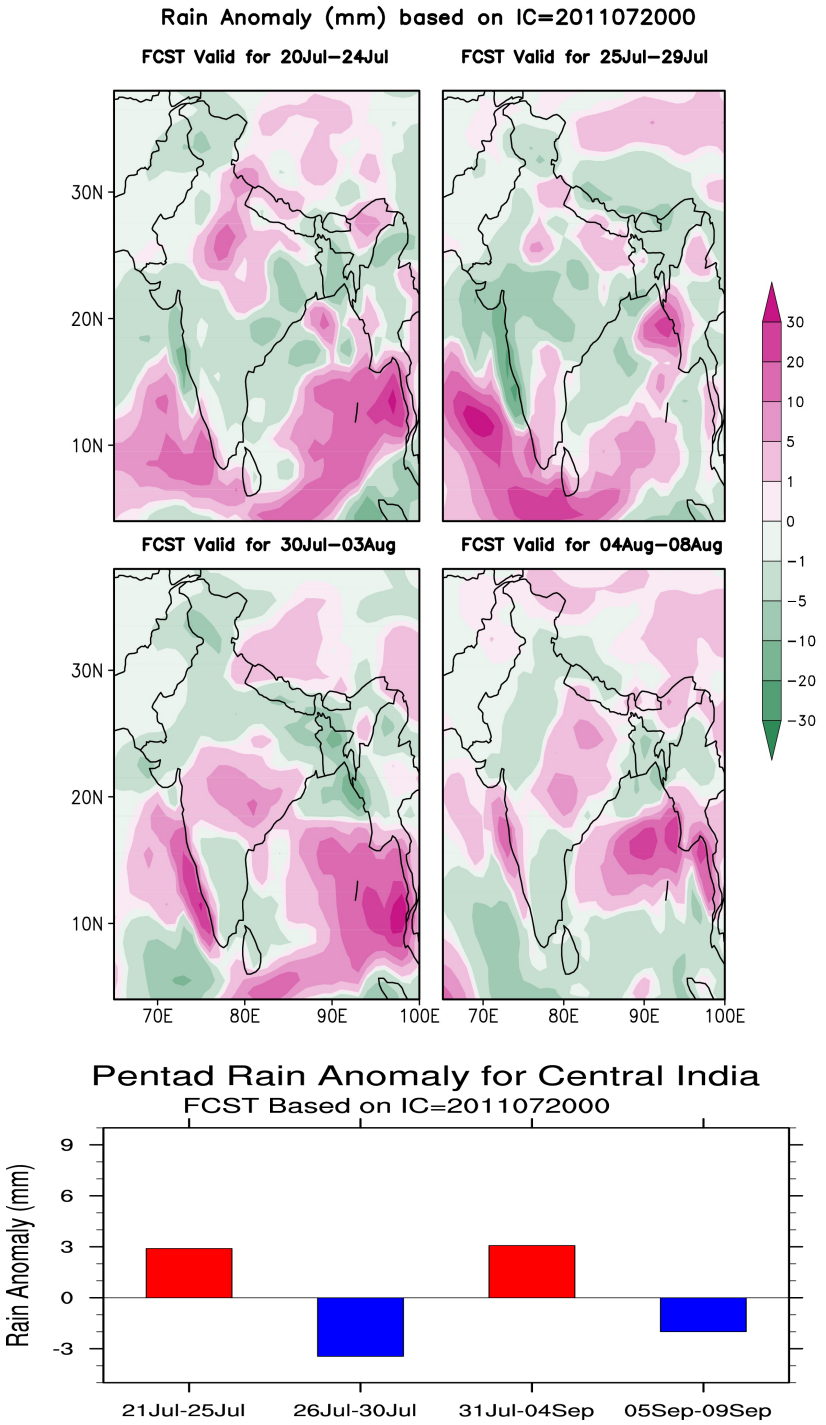


Figure 7: Same as figure 4 but for 20th July Initial condition.

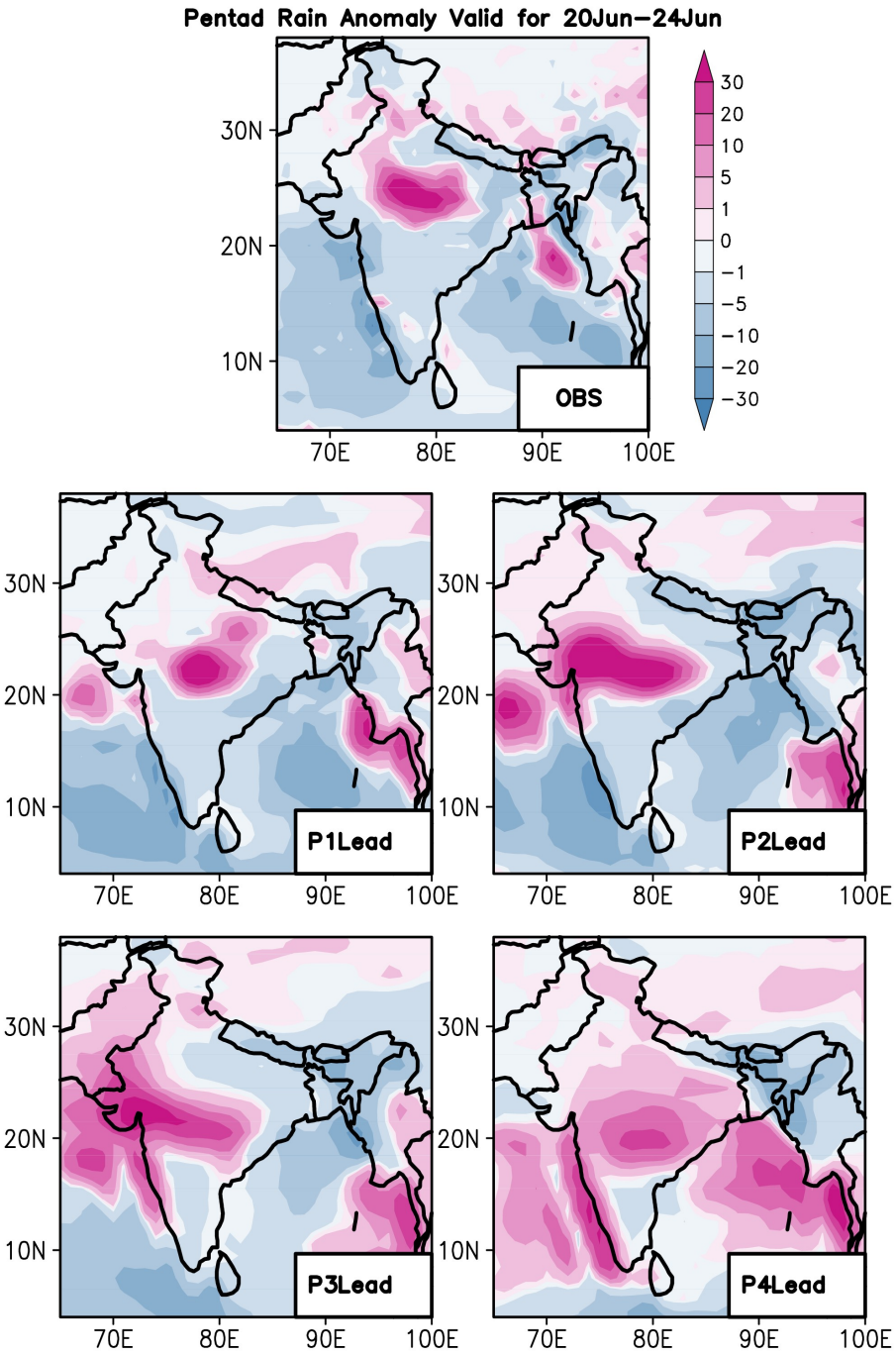


Figure 8: Pentad lead forecast of rain anomaly valid for 20-24 June for up to pentad 4 lead and corresponding observations from IMD_MRG.

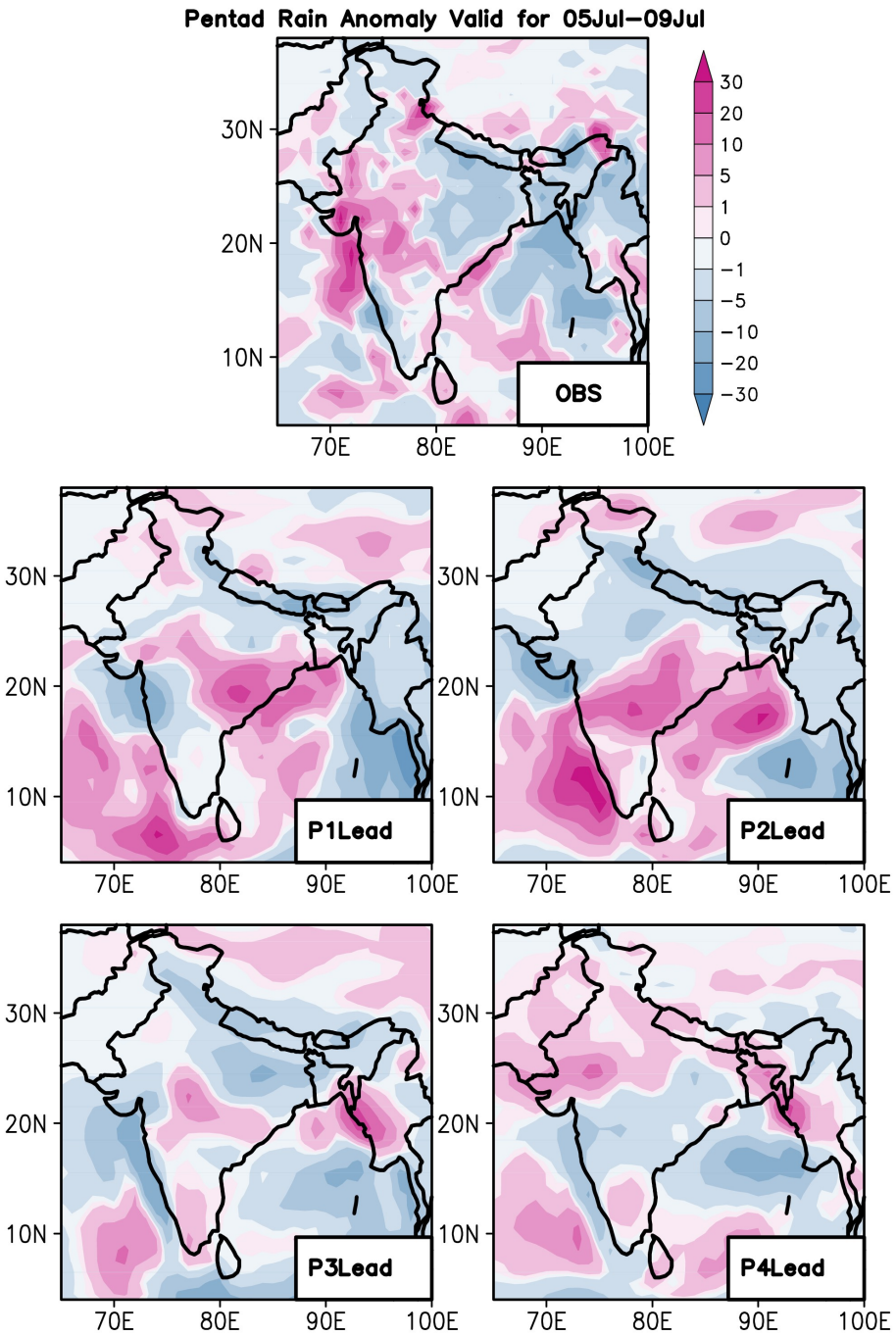


Figure 9: Same as figure 6 but for 5th July Initial conditions.

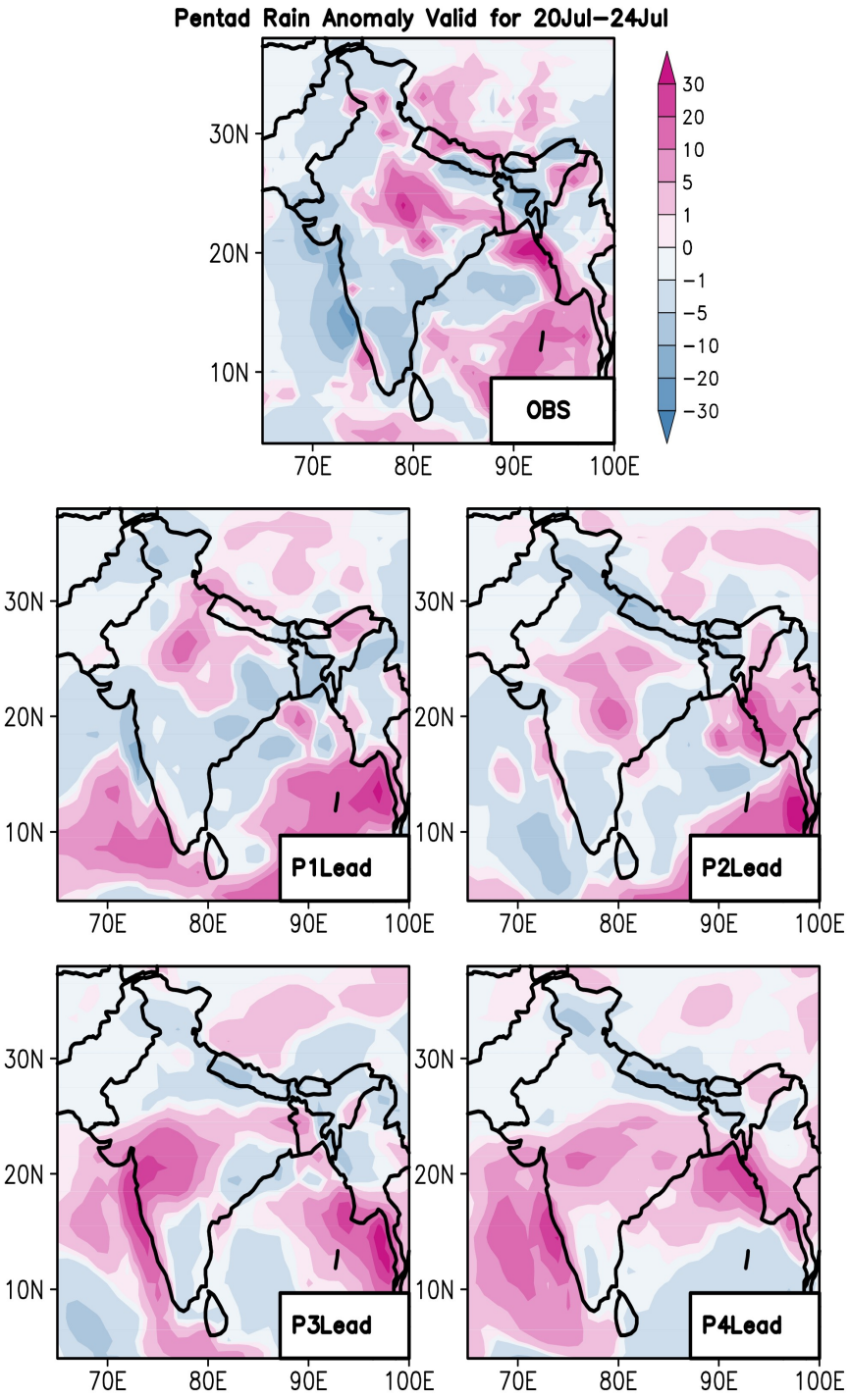


Figure 10: Same as figure 6 but for 20th July Initial conditions.

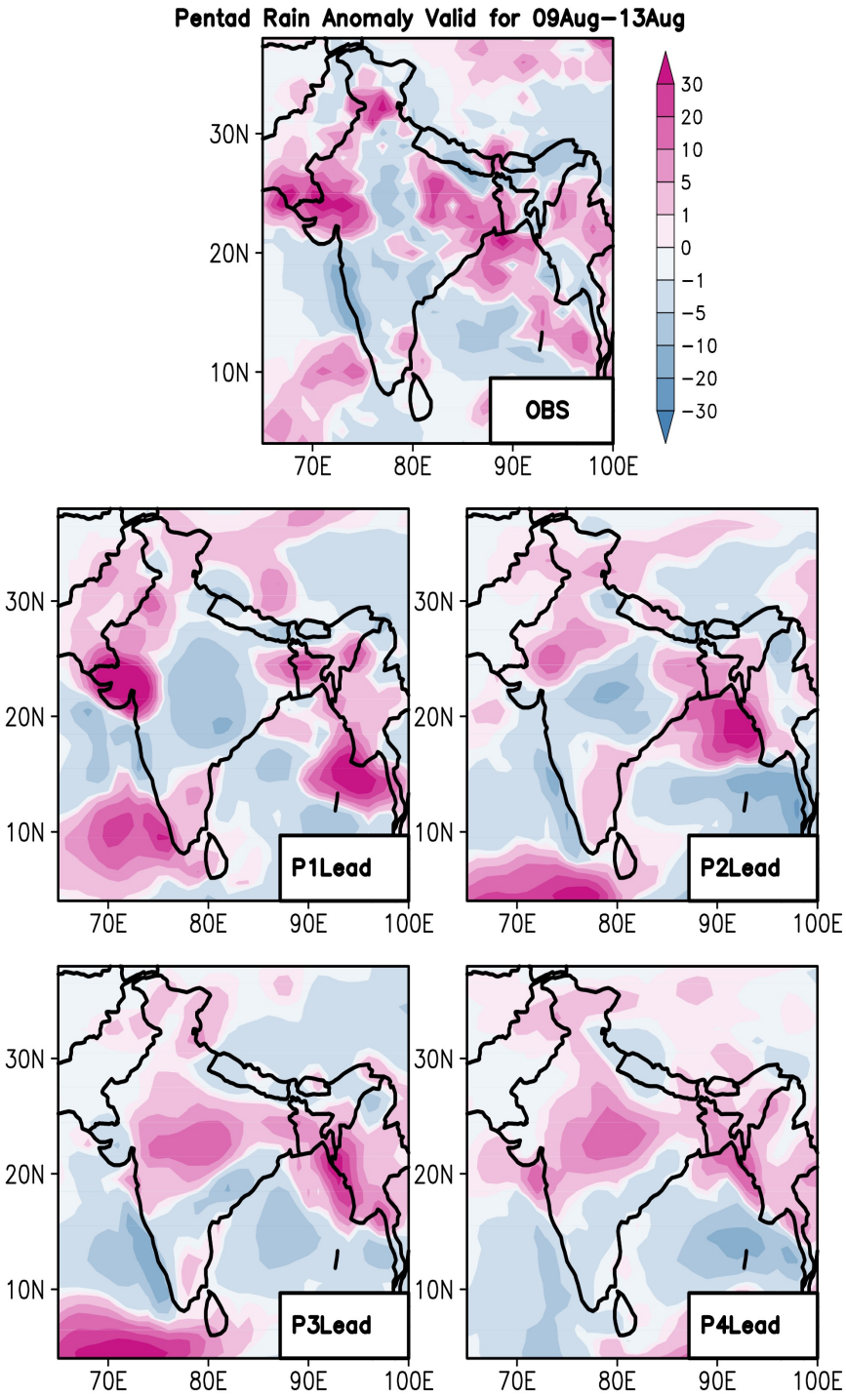


Figure 11: Same as figure 6 but for 9th August Initial conditions.

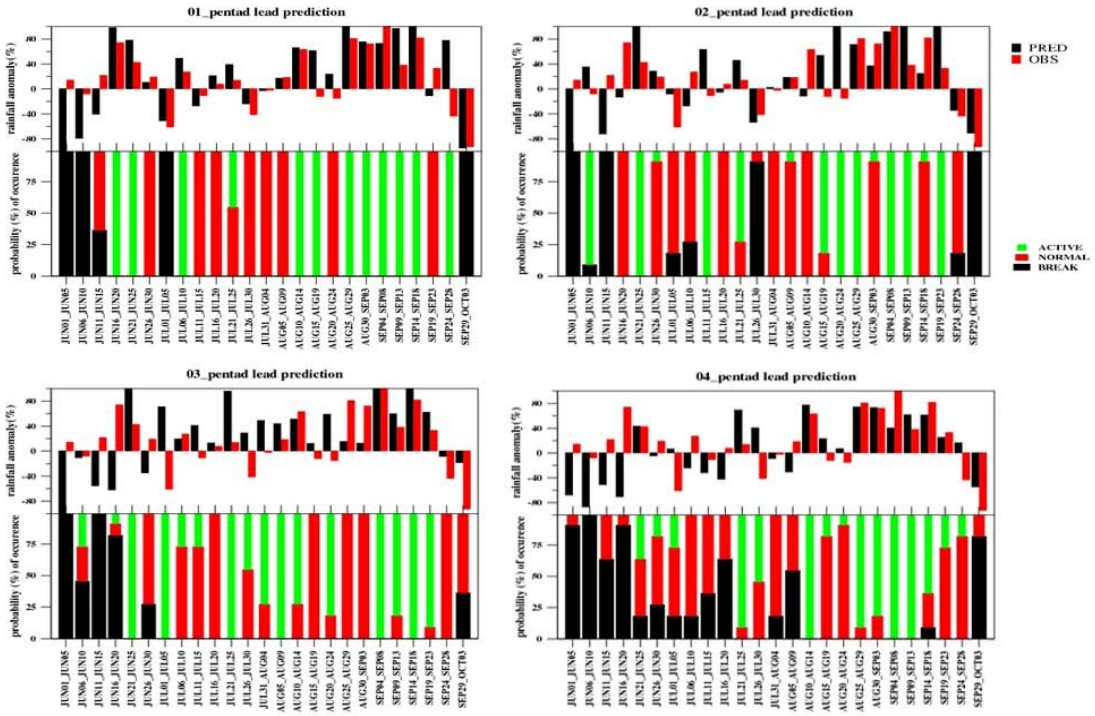


Figure 12: Observed and forecasted rain anomalies and probabilities for three categories of rainfall for pentad 1 to 4 lead.

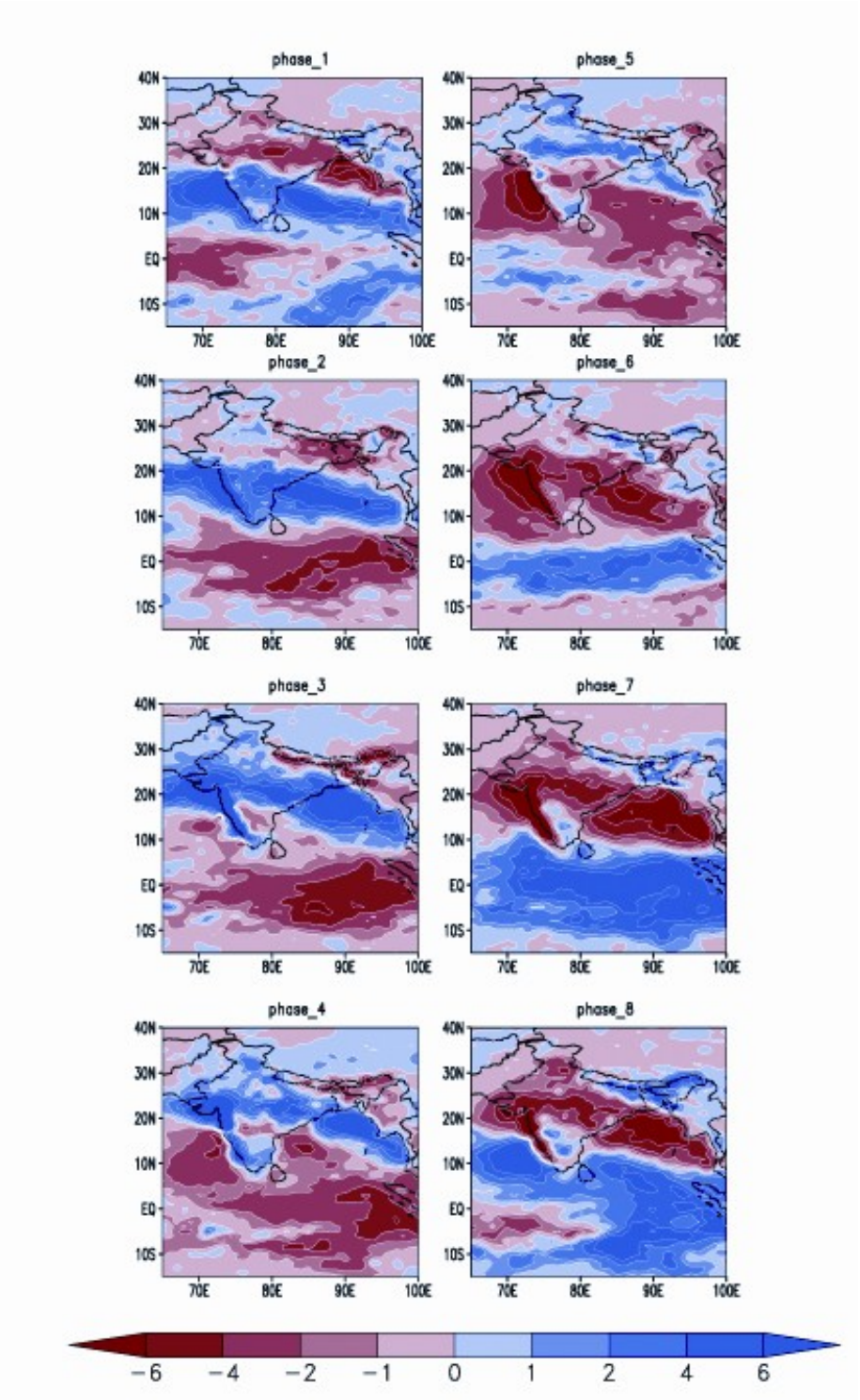


Figure 13: Composite plot for the days clustered in the 8 phases defined by MISO indices.

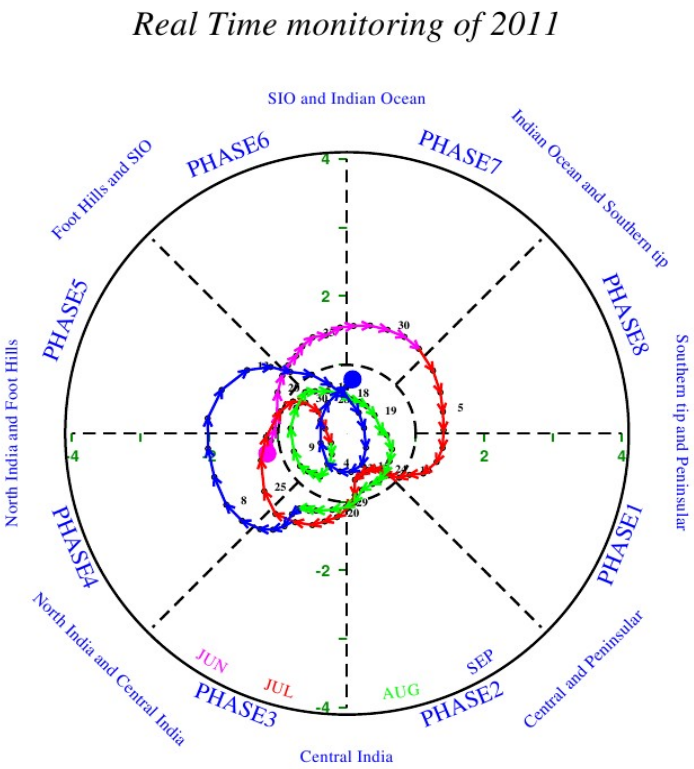


Figure 14: Phase plot of the MISO index amplitude for the year 2011 monsoon season in the phase space defined by the two MISO indices (see text).

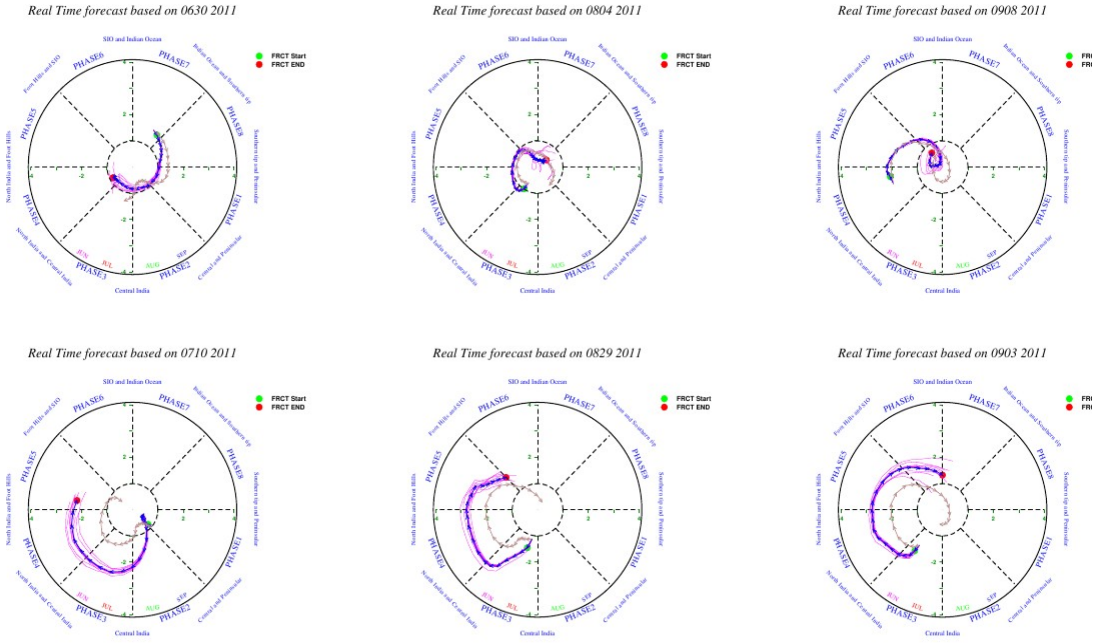


Figure 15: (top) showing the forecasted (pink) and observed (brown) MISO amplitude in the Phase space for the three "best" cases. (Bottom) showing the same for the "worst" cases.

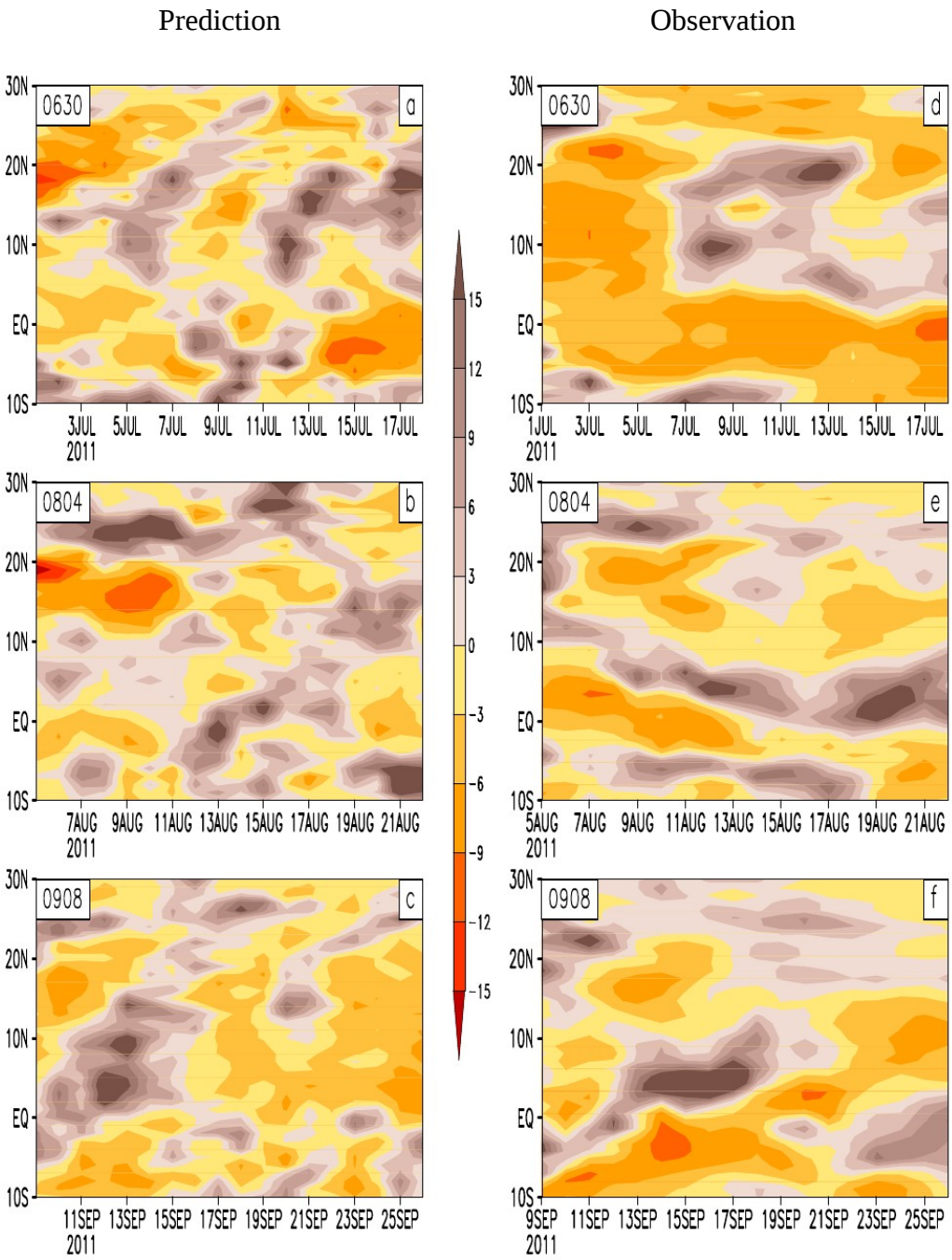


Figure 16: Hovmuller plot showing the propagation for the "best" cases.

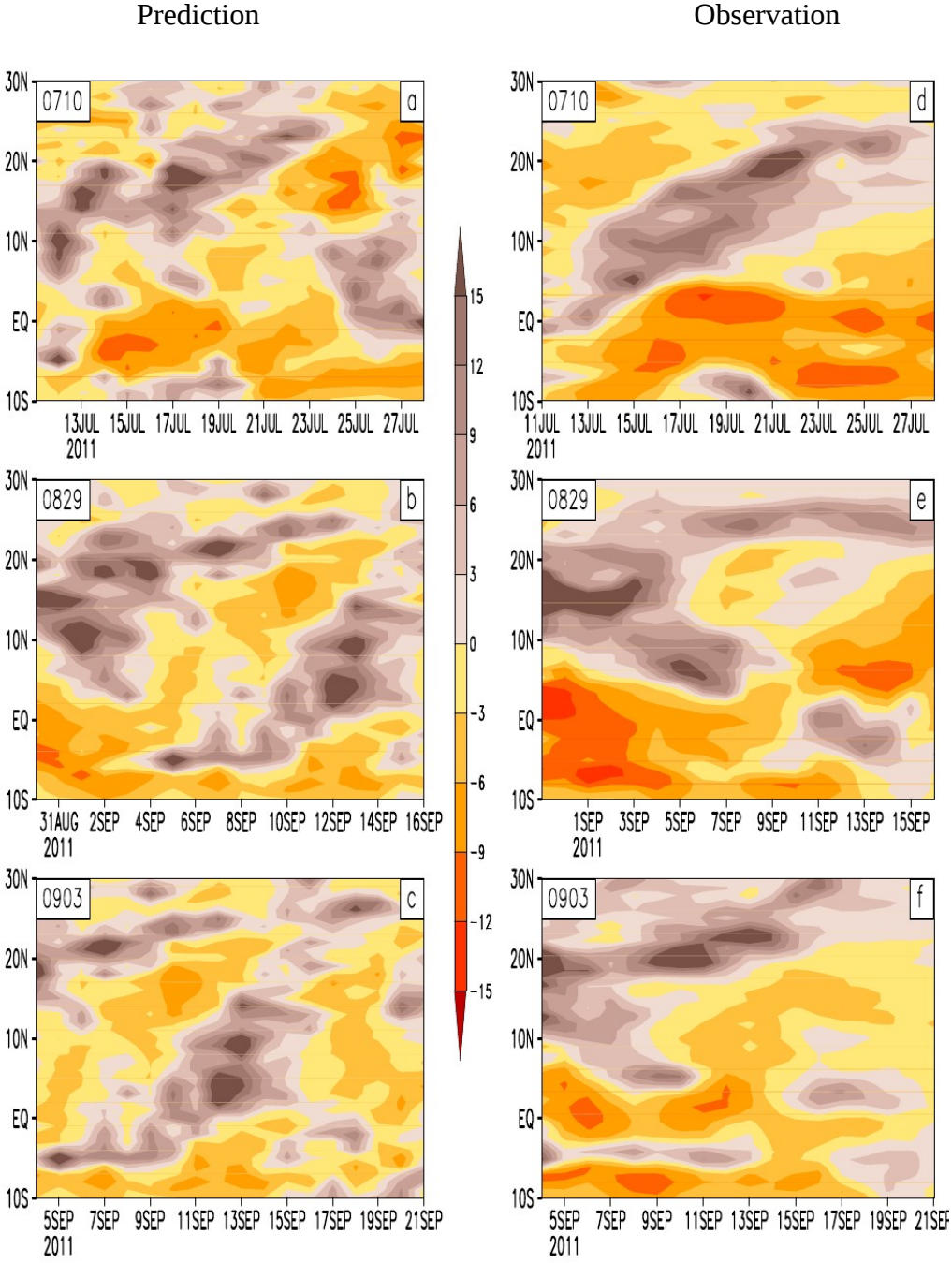


Figure 17: Same as figure 14 but for the cases defined as "worst" cases.

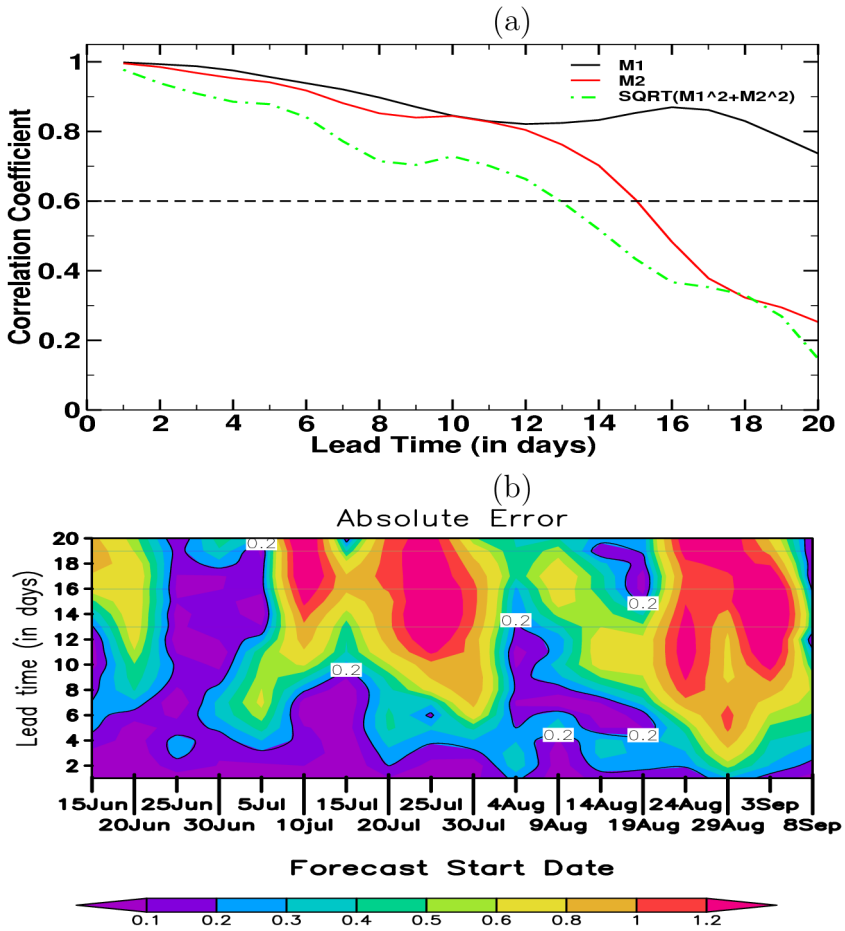


Figure 18: The correlation coefficient for the 18 cases of forecasted MISO indices as a function of lead time (top) and the absolute error as a function of model integration initial date (bottom) are shown.

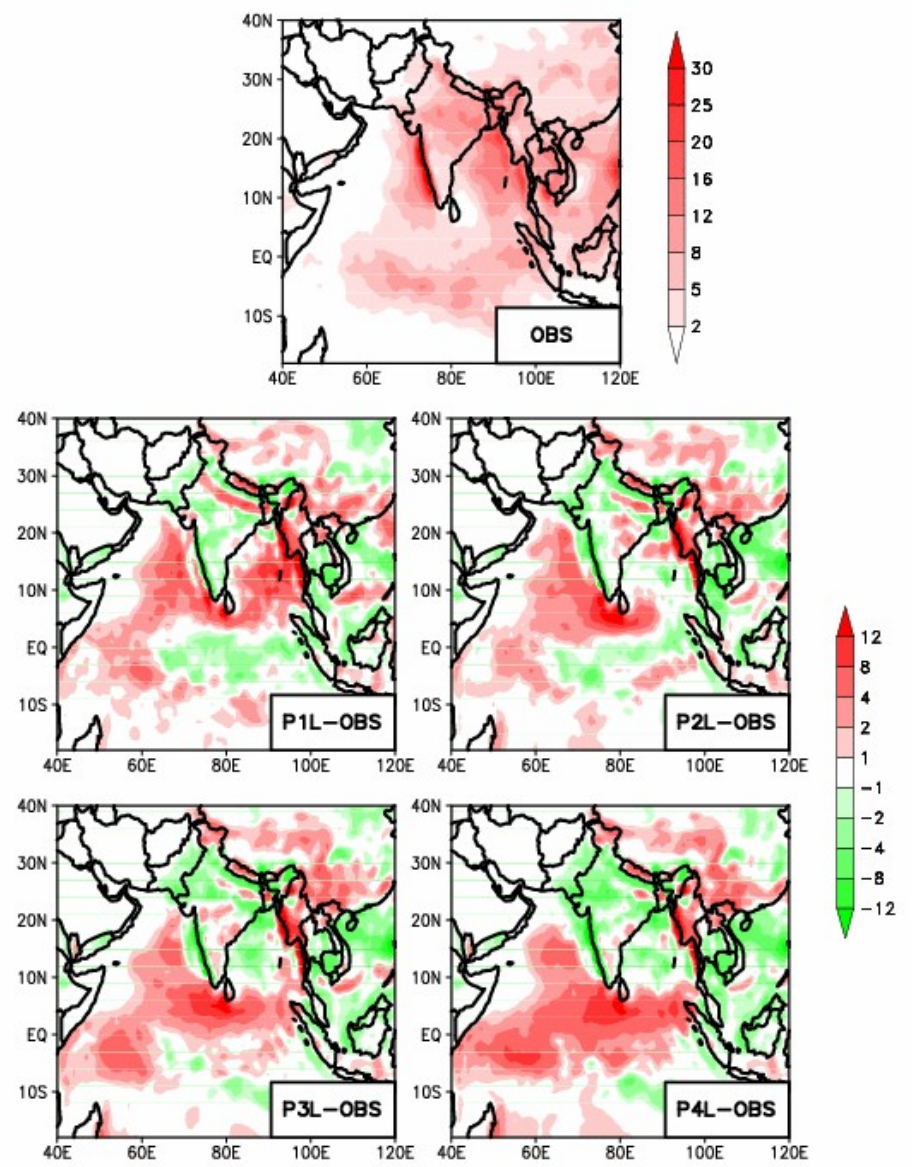


Figure 19: Seasonal mean rainfall during JJAS season from IMD_MRG (top panel) and difference between pentad lead forecast and observations.

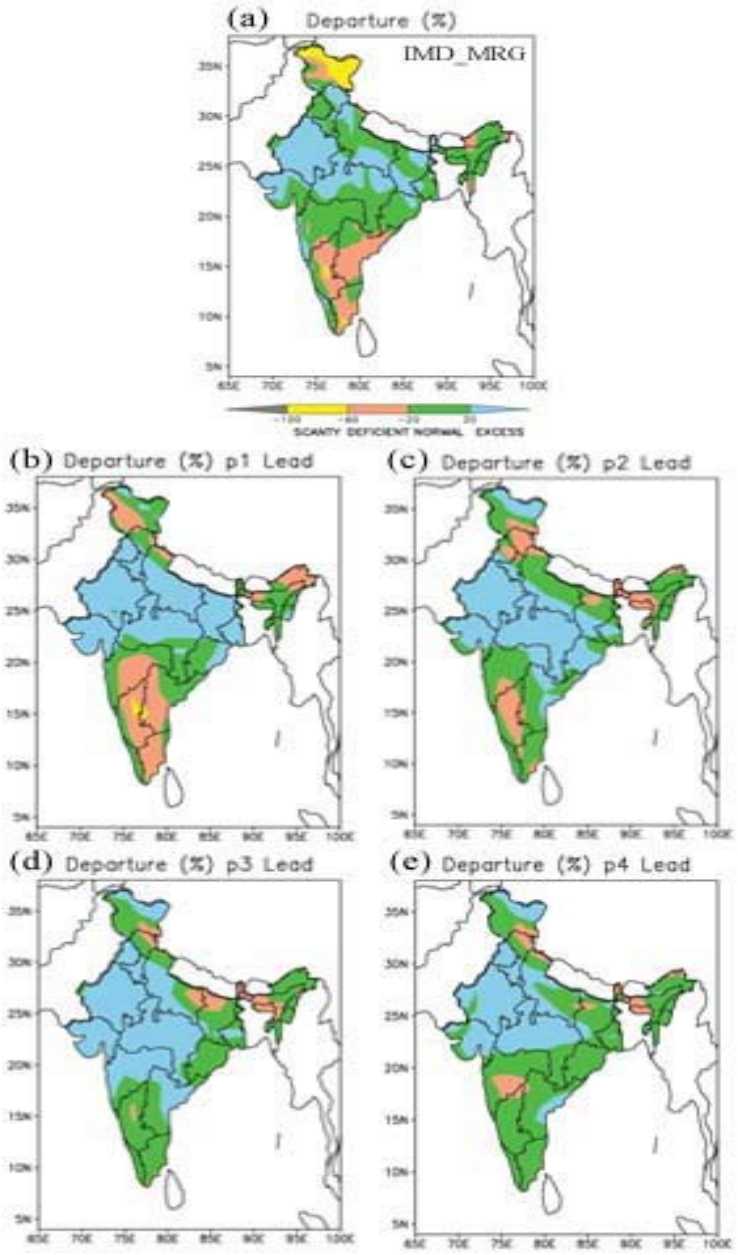


Figure 20: Rainfall departure during JJAS season from IMD_MRG (top panel) and from 1-4 pentad lead forecast.

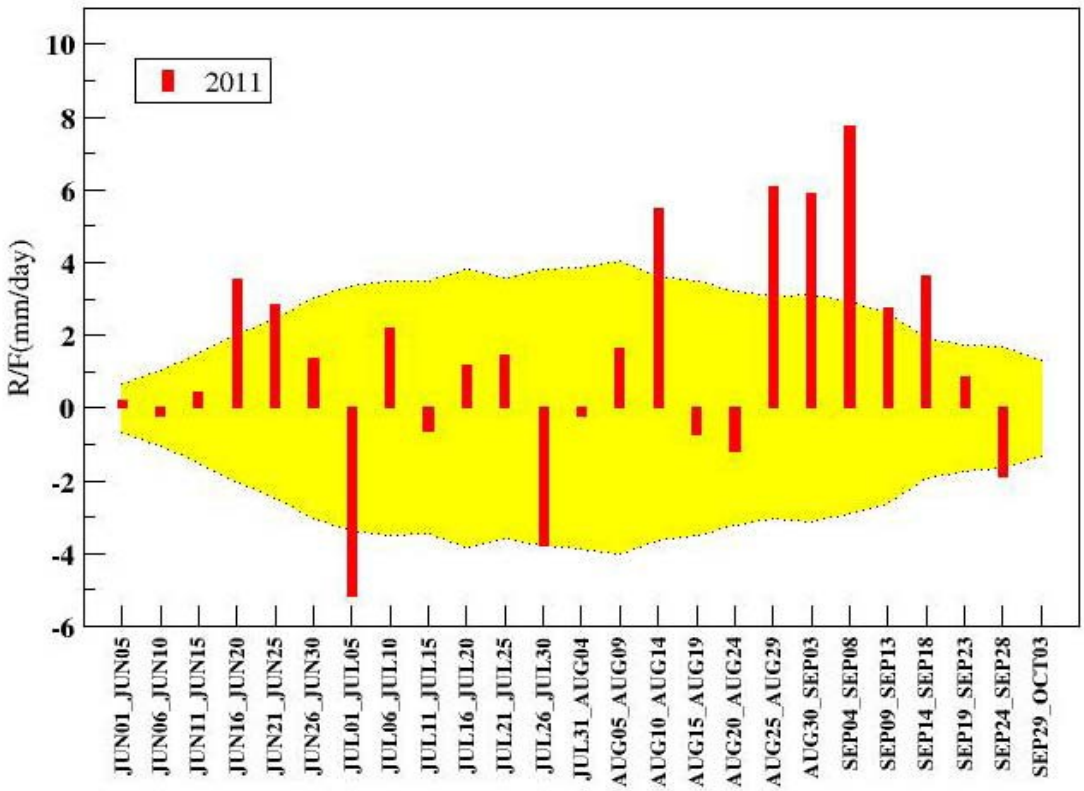


Figure 21: Active and break spells identified during the monsoon season of 2011.

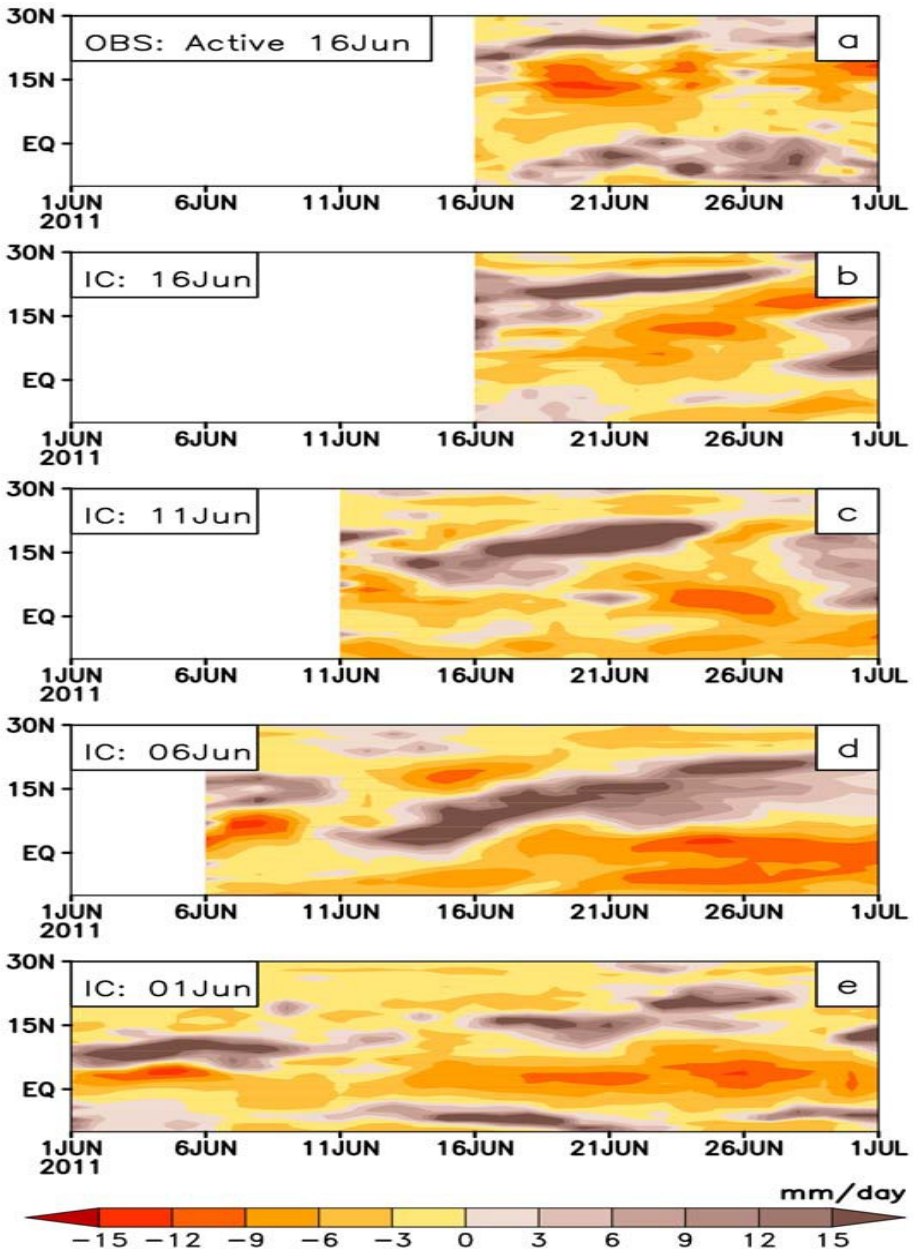


Figure 22: Time-latitude plot of rainfall anomalies (in mm/day) averaged over 70°-90° E corresponding to the active spell starting from 16 June 2011, for (a) observation and (b) forecast using concurrent initial condition. Sub-plots (c) - (e) corresponds to the forecasts using initial conditions from one to three pentads earlier.

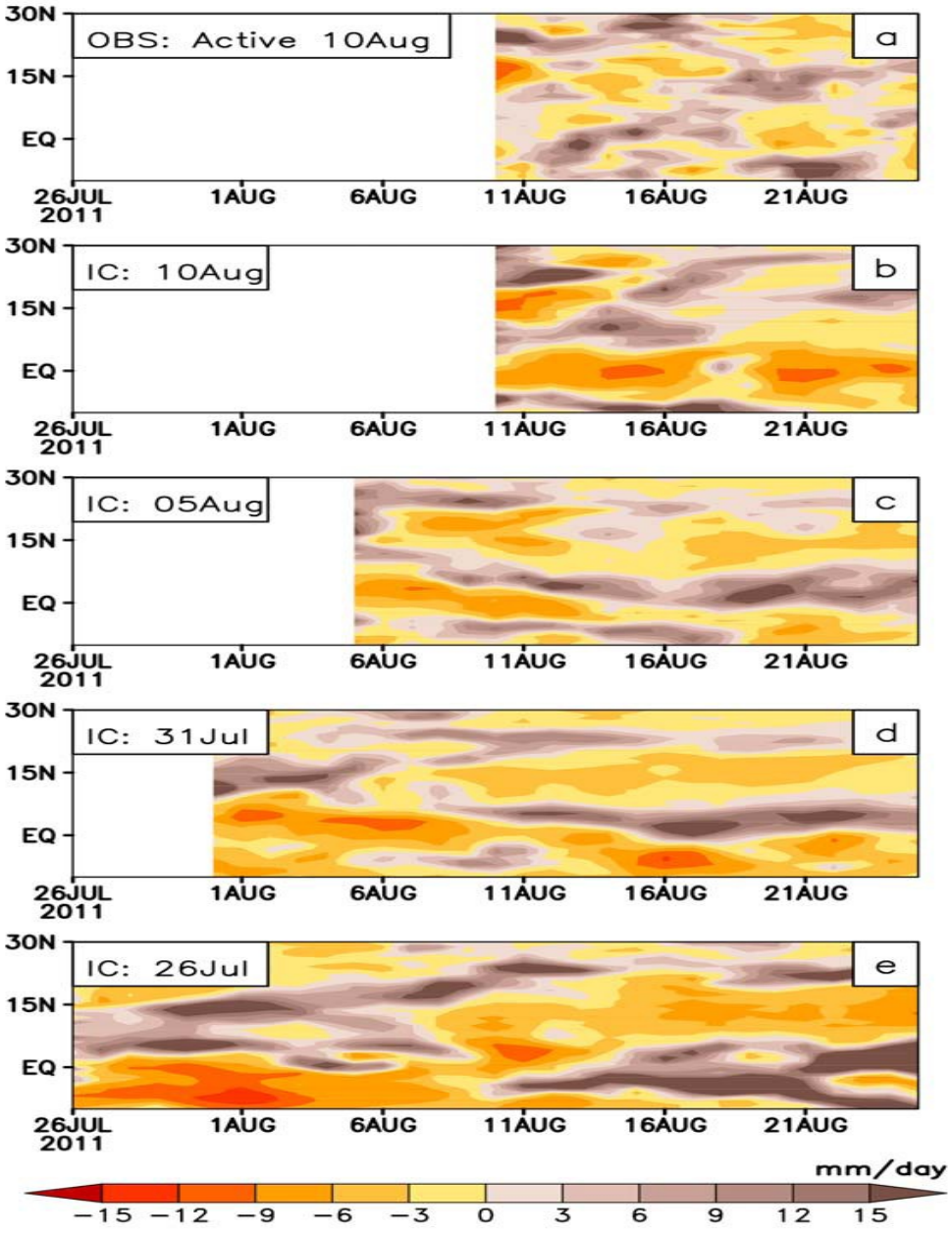


Figure 23: Same as Fig.20, but for the active spell starting from 10 August 2011.

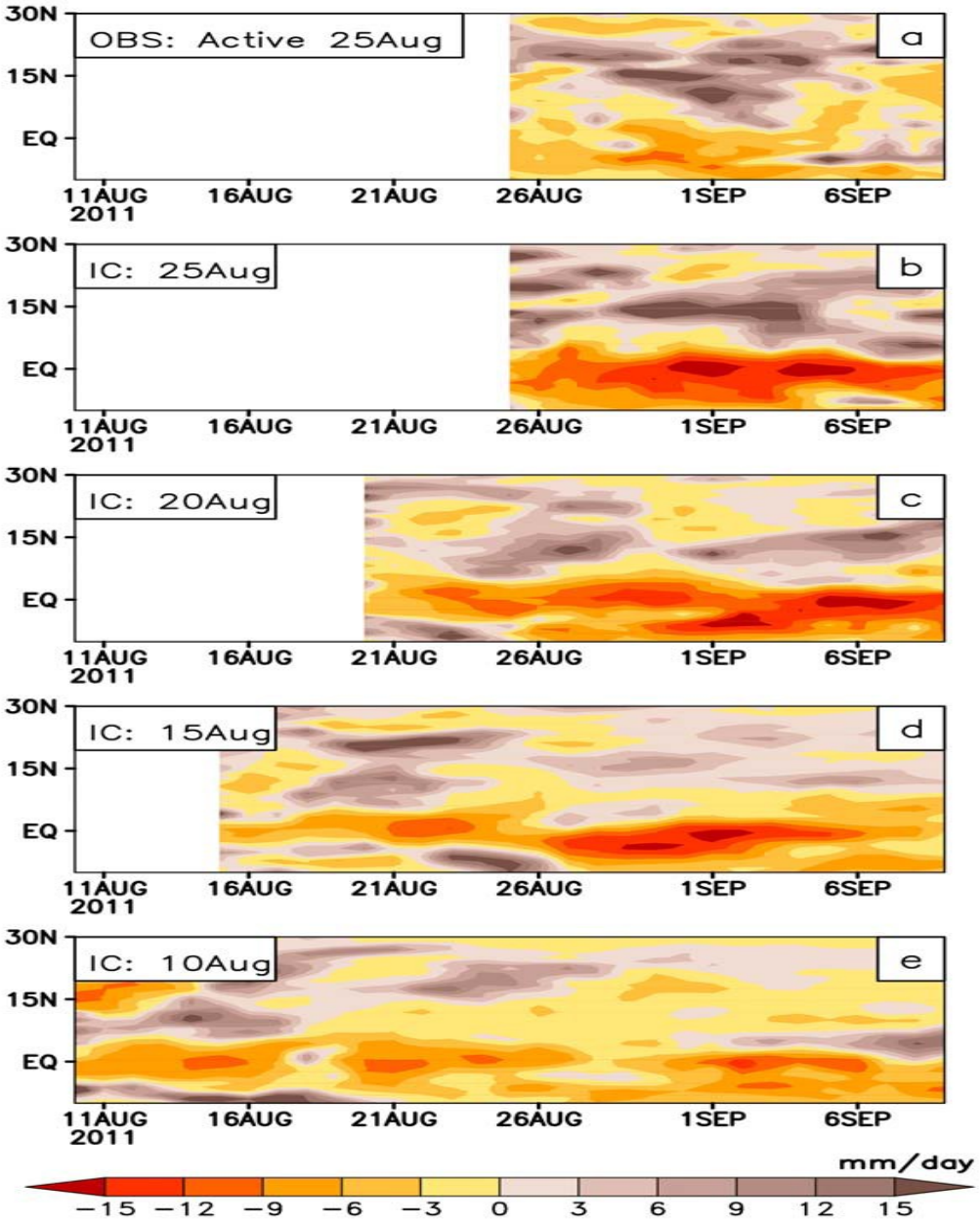


Figure 24: Same as Fig.20, but for the active spell starting from 25 August 2011.

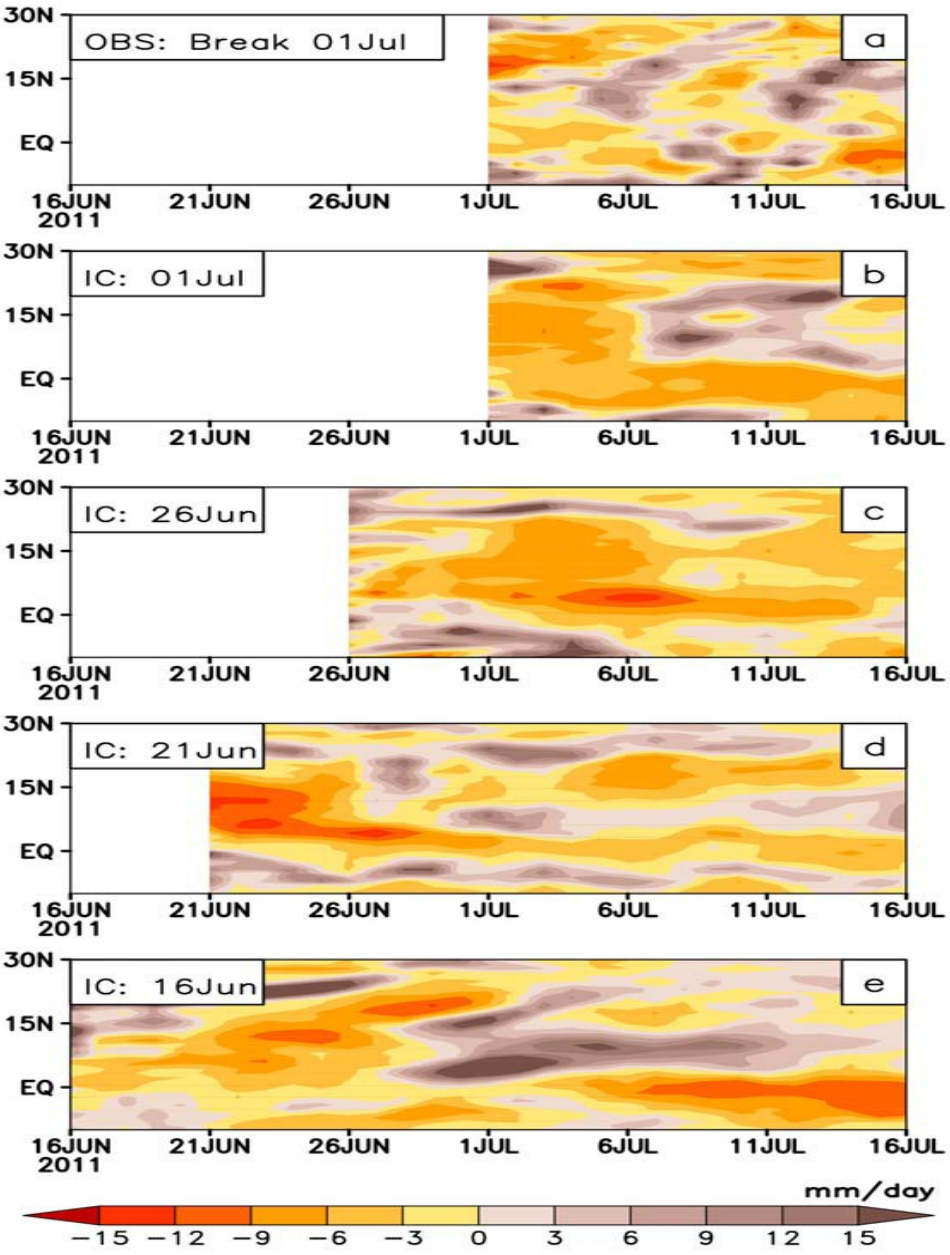


Figure 25: Same as Fig.20, but for the break spell starting from 01 July 2011.

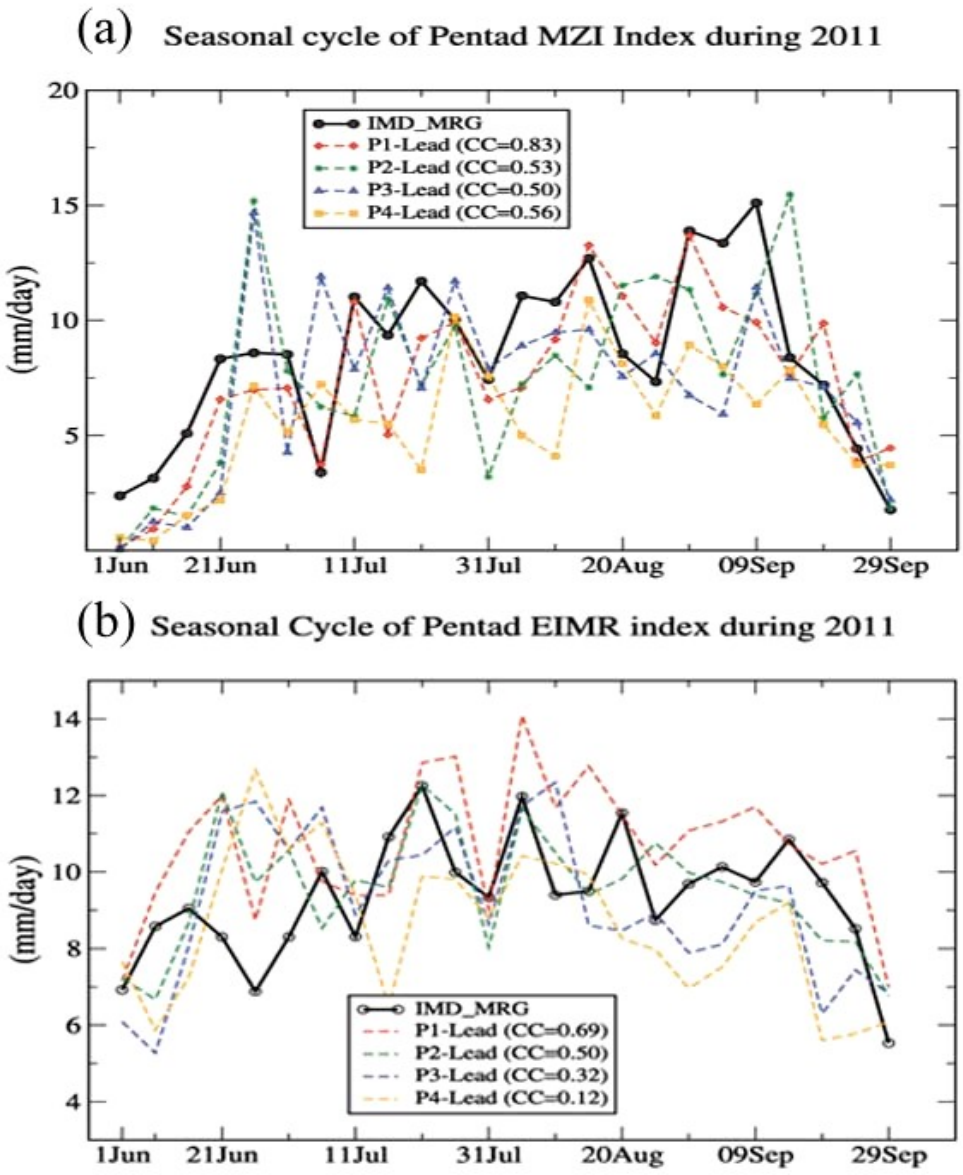


Figure 26: Seasonal cycle of pentad (a) MZI rainfall index and (b) EIMR rainfall index during JJAS monsoon season.

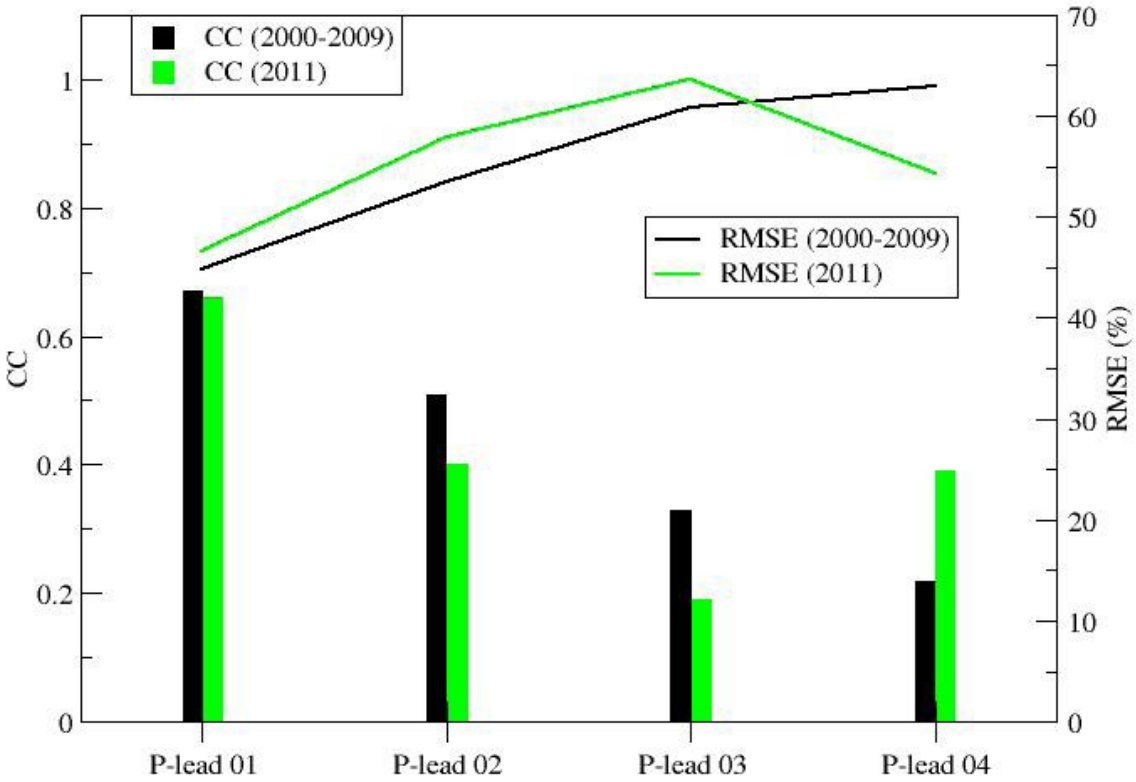


Figure 27: Correlation coefficients and RMSE calculated against IMD observations during 2011 monsoon season.

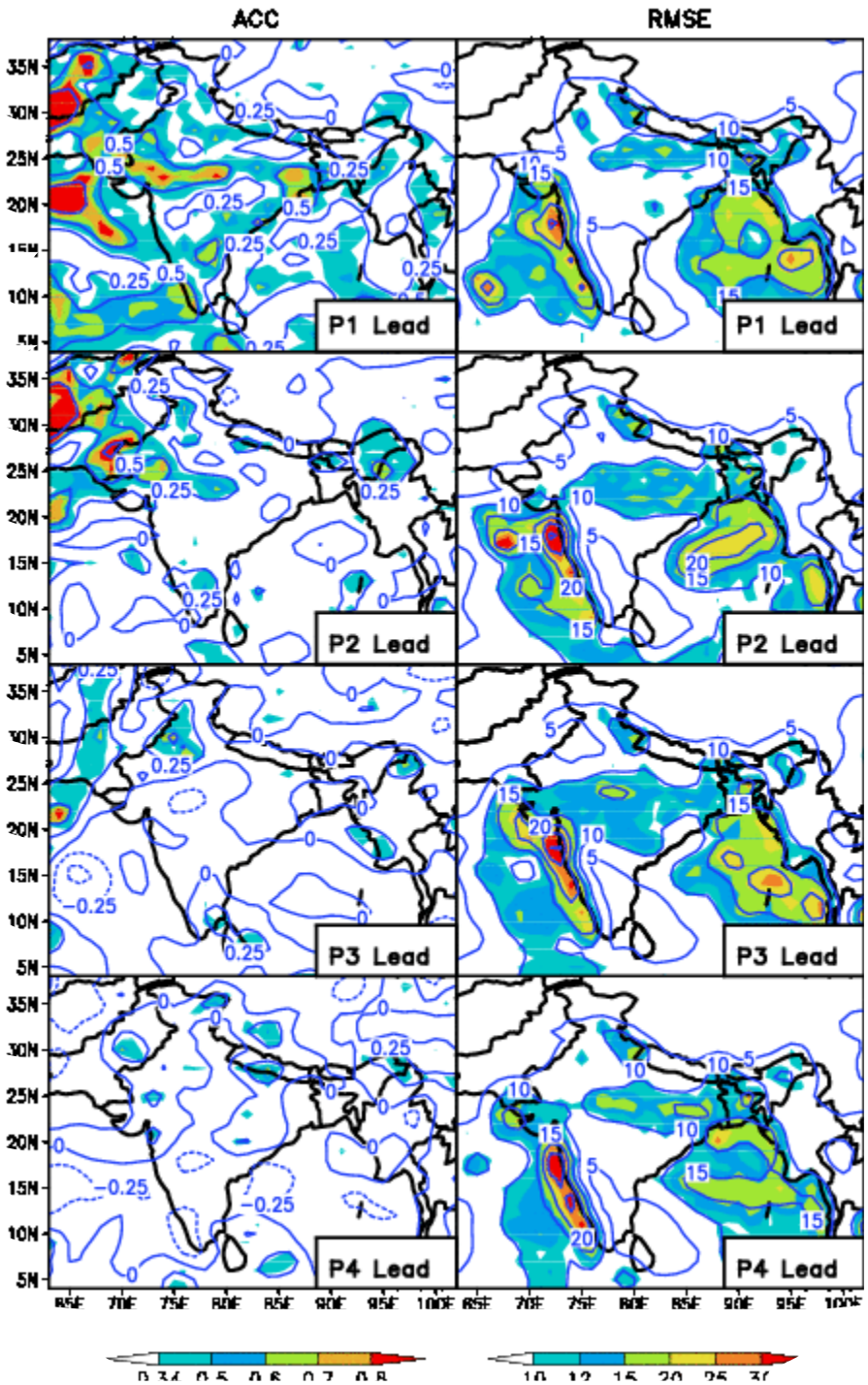


Figure 28: Anomaly pattern correlation and RMSE for different pentad lead forecast. Correlation coefficient significant at 95% are shown in shades.