MAUSAM, **66**, 3 (July 2015), 433-444

551.509.5 : 551.577 : 551.509.313.6

Diurnal cycle of rainfall as predicted by WRF model : Verification using model evaluation tools software

SOMA SEN ROY, SUBHENDU BRATA SAHA, ANANDA KUMAR DAS,
S. K. ROY BHOWMIK and P. K. KUNDU*

India Meteorological Department, Lodi Road, New Delhi – 110 003, India

**Jadavpur University, Department of Mathematics, West Bengal, India*

e mail : senroys@gmail.com

सार – मौसम अनुसंधान एवं पूर्वानुमान मॉडल से दिए गए उप-दैनिक वर्षा पूर्वानुमान की गुणवत्ता का सत्यापन तीन-तीन घंटे के अंतराल पर 72 घंटों के पूर्वानुमानों के लिए उष्णकटिबंधीय वर्षा मापन मिशन (TRMM) बहुउपग्रहीय वर्षा विश्लेषण (TMPA) डेटा सेट से किया गया है। इसका वैधीकरण दो चरणों में किया गया है:- प्रथम चरण में दैनिक चक्र के तीन-तीन घंटों में हुई ऋतुनिष्ठ वर्षा की कुल मात्रा का सत्यापन वर्षा डेटासेट (TMPA) की तुलना डब्ल्यू आर एफ मॉडल से तीन-तीन घंटों के अंतराल में 72 घंटों में होने वाली ऋतुनिष्ठ कुल वर्षा के दिए गए पूर्वानुमान से की गई है। इस मॉडल में स्थानिक एवं कालिक अधिकतम वर्षा में विसंगति को वर्षा के दैनिक चक्र को टी एम पी ए डेटासेट के माध्यम से प्राप्त किया गया है और उसके बाद उसकी मात्रा की जाँच मॉडल इवैल्यूएशन टूल्स (MET) सॉफ्टवेयर के अवयव MODE (ऑब्जेक्ट बेस्ड डायग्नोस्टिक इवैल्यूएशन विधि) का प्रयोग प्रत्येक पूर्वानुमान सत्यापन डेटासेट युग्म के लिए किया गया है। इस विश्लेषण से यह पता चला है कि इस मॉडल में वर्षा और संवहन क्रिया सतह संवेदी गर्मी से काफी प्रभावित होती है। इससे दैनिक संवहन काफी बढ़ जाता है और समूचे उपमहाद्वीप में दैनिक तापमान बढ़ने से रूक-रूक कर वर्षा होती है। इसलिए दैनिक चक्र त्रुटियाँ उन क्षेत्रों में सबसे अधिक होती है जहाँ अन्य प्रभावकारी कारक वास्तविक रूप से हावी होते हैं, विशेष तौर पर समुद्र तटीय क्षेत्रों और हिमालय के तराई वाले क्षेत्रों में। डब्ल्यू आर एफ मॉडल में संवेद्य उष्मा संवहन क्षेत्रों को बढ़ाती है जबकि कई क्षेत्र संवहन को निर्मित करते हैं, इससे पता चलता है कि इनका प्रभाव बादल की संरचना में अधिक है न कि डब्ल्यू आर एफ मॉडल में नए संवहन के निर्मित करने में। इससे स्थान विशेष का पूर्वानुमान देने में अधिक त्रुटियाँ हो सकती हैं। इस मॉडल पूर्वानुमान में वस्तु की आकृति में वृद्धि होती है जो दैनिक महत्तम तापमान के करीब होती है, इसके साथ ही वर्षा की तीव्रता सीमा अधिक हो जाती है। इस मॉडल में वर्षा के परिणाम ज्ञात करने में प्राचलीकृत संवहन का प्रभाव देखा गया है।

ABSTRACT. The quality of sub-daily rainfall forecast from the Weather Research and Forecasting model during June to September, 2013 is verified at three hourly intervals for forecasts up to 72 hours against the corresponding data from the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) dataset. The validation is done in two stages, in the first stage, the diurnal cycle of the 3 hourly accumulated seasonal totals of the verification rainfall dataset (TMPA) is compared with the corresponding 3 hourly accumulated seasonal totals of the rainfall forecasts from WRF model for forecast up to 72 hours. The discrepancy in the spatial and temporal maximum of model derived diurnal cycle of rainfall with respect to the TMPA dataset is then investigated quantitatively for every forecast-verification dataset pair using the MODE (Method for Object-based Diagnostic Evaluation) component of the Model Evaluation Tools (MET) software. The analysis demonstrates that the rainfall and convection is strongly forced by surface sensible heating in the model. This causes the diurnal peak of convection and rainfall over the entire subcontinent to be in phase with the diurnal peak of temperature. Hence, the diurnal cycle errors are highest over regions, where other forcing factors actually pre-dominate; especially along the coasts and the foothills of the Himalayas. The sensible heat forcing tends to increase the areas of convection in the WRF model, rather than the number of zones of genesis of convection, indicating that its influence is more towards cloud organization rather than genesis of new convection in the WRF model. There is a strong component of displacement error in the WRF model forecast, which may cause large errors in location specific forecasts. The increase in the object size in the model forecast close to the diurnal maximum of temperature, accompanied by a peak of the rainfall intensity range, suggests the dominance of parameterized convection in the model rainfall output.

Key words – Rainfall, WRF model, Forecast verification.

1. Introduction

The main rainfall period over most of the Indian subcontinent occurs during summer monsoon season (June to September), in association with the monsoon circulations. Since most of agriculture and repletion of fresh water resources over the region occurs from these rainfall episodes, accurate prediction of the rainfall from numerical weather prediction models is of great importance. More recent focus has also been on the prediction of heavy rainfall episodes and their potential for causing devastation. These episodes are marked by short, intense rainfall spells within the main episode, over a localized region (Kotal *et al.*, 2014). Hence the thrust of numerical weather prediction models has moved from daily scale to a sub-daily scale of prediction of location specific rainfall episodes.

India Meteorological Department (IMD) is tasked with providing real-time weather forecasts for the Indian region to the general public, on various scales. For short range forecasts of rainfall and other parameters, on a sub-daily time scale and location specific nature, IMD operationally runs the Weather Research and Forecasting - Advanced Research (WRF-ARW) Numerical Weather Prediction (NWP) model (developed by National Center for Atmospheric Research NCAR, USA). The increasing operational demand for WRF model output on such short spatial and temporal scales indicates that rigorous validation is necessary, to bring to focus any model inconsistencies in the sub-daily scale. The focus of this study is on the most significant output of the model, *viz*., rainfall, whose validation on a sub-daily scale will focus on the errors in the simulation of the observed diurnal cycle, which are smoothened out while accumulating to daily scale. This study addresses the important issue by validating the sub-daily scale (3 hourly intervals) rainfall forecast for forecast up to 72 hours from the WRF model over the Indian region with respect to the TMPA dataset, which is of similar temporal and spatial resolution.

Many previous studies have validated the rainfall forecast from NWP models on the daily scale for the monsoon season over the Indian subcontinent (Ranade *et al.*, 2014; Roy Bhowmik and Prasad, 2001; Mandal *et al.*, 2007). All these validation studies have noted that the forecast by NWP models has greater skill for low intensity rainfall events as compared to high intensity events. That has, in fact, been the justification for moving towards a multimodel ensemble forecast approach, which apparently provides better results (Roy Bhowmik and Durai, 2010).

One of the few studies that analyzed the diurnal cycle of model predicted rainfall over the Indian subcontinent, noted that the model forecasts of the Global Forecast System of National Centre for Environmental Prediction (NCEP) weaken the mesoscale effects on precipitation forecast (Basu, 2007). The study also noted the small thermal inertia in rainfall initiation over the land regions, as indicated by a shorter lag in the diurnal temperature and rainfall maxima in the model forecast. The frequency and amount of precipitation was observed to increase with the forecast length; although the duration of maximum precipitation remains almost the same. One of the important studies to verify the diurnal cycle of rainfall by the WRF model was carried out for the Continental United States during July-August 2001 (Davis *et al.*, 2006). The study noted that the WRF produces too many large rain areas, and the spatial and temporal distribution of the rain areas reveals regional underestimates of the diurnal cycle in rain-area occurrence frequency. They also noted that WRF model rain errors exhibit no large biases in location, but do suffer from a positive size bias that maximizes during the later afternoon. This coincides with an excessive narrowing of the rainfall intensity range, which is consistent with the dominance of parameterized convection.

The present study is divided into the following sections: Section 2 discusses the datasets used in the study and methodology of analysis. Section 3 presents the results of comparison of WRF model output with TMPA estimates of rainfall, while in Section 4 the conclusions of the study are summarized.

2. Data and methodology

2.1. *Data sources*

The 3 hourly WRF model forecast of rainfall, which is run with the initial condition at 0000 UTC everyday for forecasts up to 72 hours during the south-west monsoon season (June to September) of 2013 is validated with respect to the TMPA dataset for the corresponding time. The regional mesoscale analysis system WRF (ARW) is operational on the High Performance Computing (HPC) system at IMD, New Delhi with its essential comprehensive components of preprocessing program (WPS and REAL), assimilation program (WRFDA), boundary condition update (update_bc) and forecasting model (WRF-ARW). The model domain, at 27 km resolution, covers the Indian subcontinent and surrounding regions and is displayed in Fig. 1 (23.204° S to 46.154° N and 39.565° E to 120.435° E). The observations from different sources (including satellites) are assimilated in WRFDA system to improve the first guess GFS analysis. Assimilation is done with 27 km horizontal resolution and 38 vertical eta levels. Usable cold-start mode of assimilation is presently adopted for WRFDA system. WRF-ARW model is then integrated for 75 hours with a nested configuration (27 km mother and 9 km child domain). The forecast model is configured with full physics (including cloud microphysics, cumulus, planetary boundary layer and surface layer parameterization) as well. The RRTM long-wave and Goddard short-wave radiations physics schemes have been selected in the study. WRF single moment 5-class cloud microphysics and Grell 3 dimensional ensemble cumulus physics scheme have been selected while the

Mellor-Yamada-Janjic planetary boundary layer scheme. Model configuration includes Eta and Noah Land Surface Model for surface physics. These schemes have been selected by taking into account the various optimization studies carried out for tropical regions. Further details can be found in Das *et al.* (2015). The model rainfall forecasts are available at both 27 and 9 km for the Indian sub-continent.

While daily scale rainfall data is available from multiple sources [*e.g*., GPCP V2 (Adler *et al*., 2003), IMD0.25 (Pai *et al.*, 2014) and APHRODITE (Yatagai *et al.*, 2009)], very few reliable sub-daily rainfall estimates are available over the Indian region. Rain gauge data, at 3 hourly intervals is not very dense over the Indian region. Hence it cannot be used. Uniform sub-daily rainfall values over the Indian region is available from satellites (*e.g*., TMPA and CMORPH). The non-realtime version of the CMORPH and TMPA are both available at 0.25 deg. grid resolution. However, the CMORPH dataset (Joyce *et al.*, 2004) has not been sufficiently validated over the Indian region. On the other hand, the strengths and weaknesses of the TMPA dataset (3-hourly, 0.25 deg. product) (Huffman *et al*., 2007) over the Indian region are well documented (Rahman *et al*., 2009; Nair *et al*., 2009; Durai *et al*., 2010). The TMPA dataset has been validated on the daily scale with the 1 deg. gridded rainfall data of IMD and reliably depicts the pattern and intensity of heavy rainfall from individual monsoon low-pressure systems and depressions (Rahman *et al*., 2009). The errors are also less for non-hill regions (within about 15%) and more so over the hilly terrain. It is unable to resolve the heavy orographic rainfall on the windward side of the Western Ghats and overestimates the rainfall on the immediate leeward side of mountains. TMPA estimates over the Western Ghats were found to be most accurate over regions of moderate rainfall and mainly inaccurate in regions of sharp rainfall gradient (Nair *et al*., 2009). From the use of three years data (2006-2008), Durai *et al*. (2010) concluded that TMPA dataset distinctly captures characteristic features of summer monsoon rainfall. Hence we decided to use this TMPA (TRMM 3B42 V7) dataset for validating the WRF model forecast rainfall. This dataset has also been reliably used in other sub-daily rainfall validation studies also (Sen Roy *et al*., 2012 for example). Since the validation dataset was at 0.25 deg. resolution, the rainfall forecast at the closest resolution (27 km) was selected for validation.

2.2. *Methodology of validation*

A smaller box within the main domain (5.125° N to 40.125° N and 65.125° E to 100.125° E) enclosing the Indian subcontinent is chosen for validation (box outlined

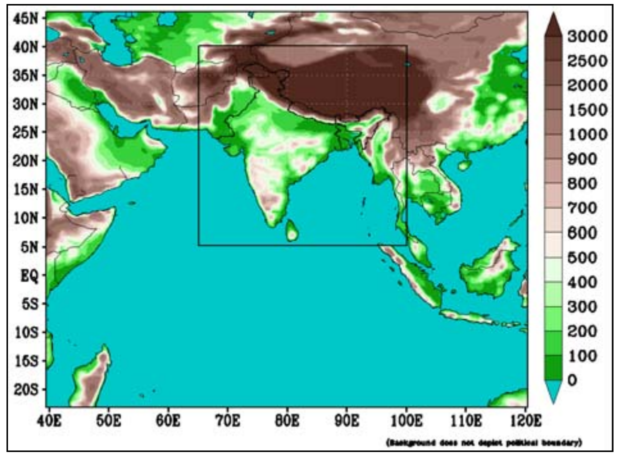
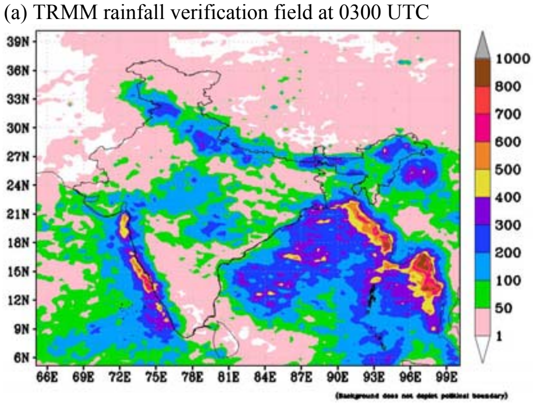


Fig. 1. Domain of the WRF model at 27 km resolution. The inner box delineates the domain for which the rainfall data has been validated

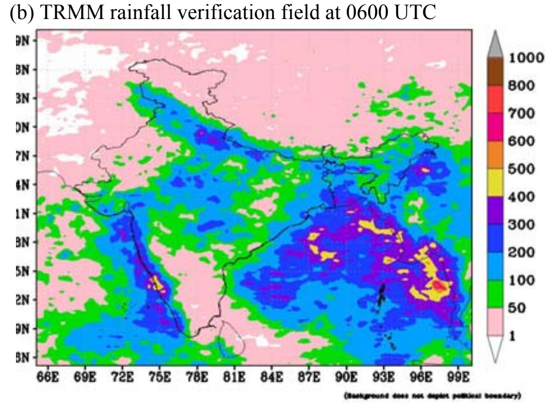
in Fig. 1). The study is done in two steps. Initially, the broad, climatological scale spatial variation of the diurnal cycle of monsoon rainfall is analyzed from the seasonal totals of the three hourly accumulations of the TMPA rainfall. This is compared to the model forecast of 3 hourly rainfall totals accumulated for the entire monsoon season. Hence, for example, the 0600 UTC seasonal rainfall accumulation of three hourly rainfall from TMPA, is compared separately to the 06 hr, 30 hr and 54 hr 3 hourly accumulated rainfall from the WRF model accumulated for the entire season to obtain the gross composite characteristics of day 1, day 2 and day 3 forecast behaviour at 0600 UTC of the day for the entire monsoon season. These three hourly spatial composite maps for TMPA are displayed in Figs. 2(a-h) and for WRF (from 03 hour to 72 hours) in Figs. 3-5(a-h). Subsequently, the individual WRF forecasts are validated against the corresponding TMPA rainfall. The three hourly rainfall forecast of the model, at 27 km resolution, are interpolated to a 0.25 deg. grid size employing mass conservation (*copygb utility*) to bring the model forecast to the same resolution as the verification data of TMPA**.** The object based area identification and verification tool MODE (Method for Object-based Diagnostic Evaluation), which is part of the forecast verification MET (Model Evaluation Tools) package developed by the National Centre for Atmospheric Research (Brown *et al.* 2007; Davis *et al.* (2006), is used to compare the rainfall forecast valid for a particular time, with the corresponding data of TMPA.

For example, the 0600 UTC three hourly rainfall accumulation from TMPA on a particular day, is compared separately to the 06 hr rainfall forecast from the model run of same day, 30 hr rainfall forecast from the model run of the previous day and 54 hr rainfall forecast

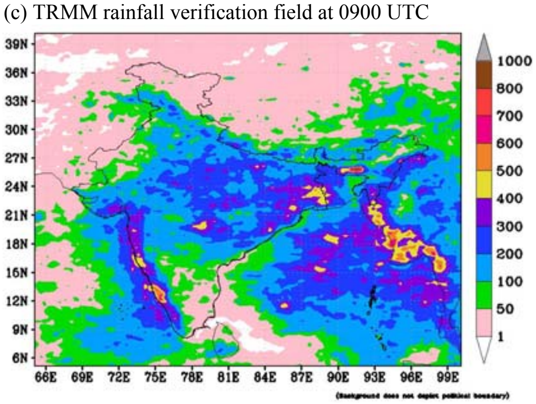
(a) TRMM rainfall verification field at 0300 UTC



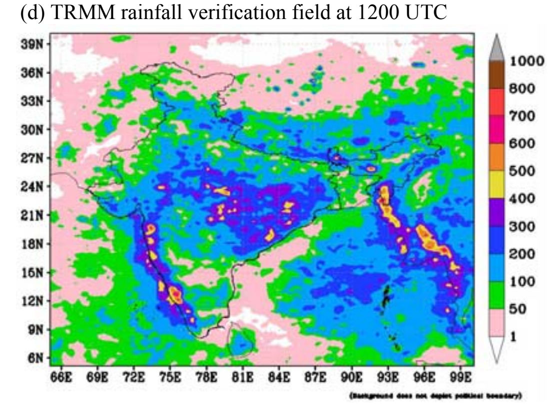
(b) TRMM rainfall verification field at 0600 UTC



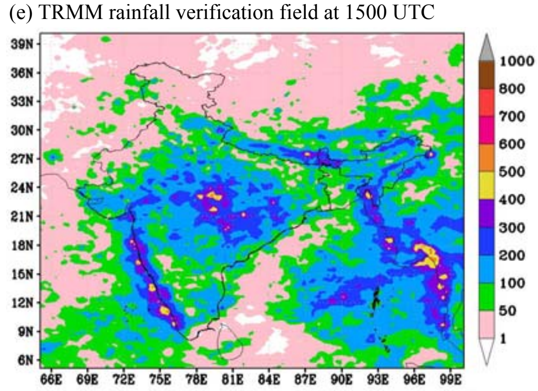
(c) TRMM rainfall verification field at 0900 UTC



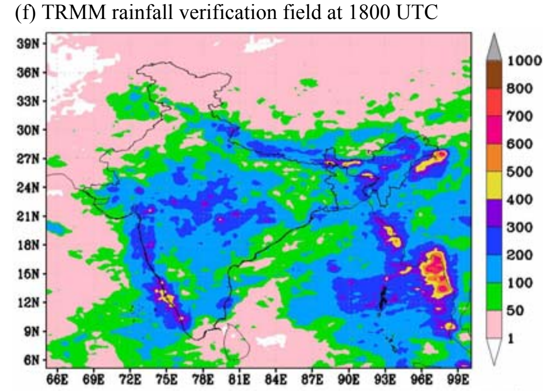
(d) TRMM rainfall verification field at 1200 UTC



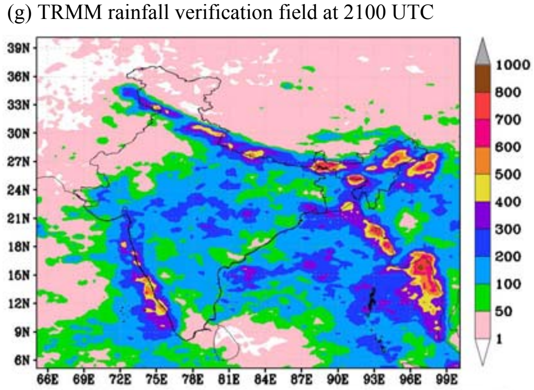
(e) TRMM rainfall verification field at 1500 UTC



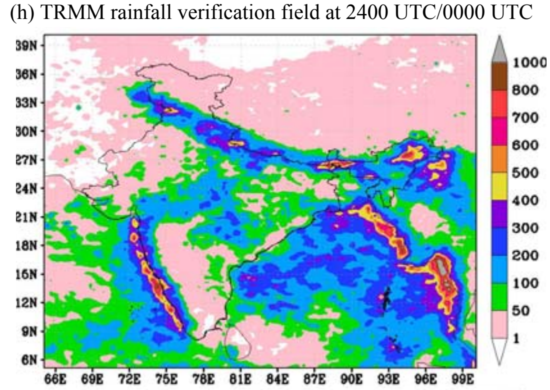
(f) TRMM rainfall verification field at 1800 UTC



(g) TRMM rainfall verification field at 2100 UTC



(h) TRMM rainfall verification field at 2400 UTC/0000 UTC



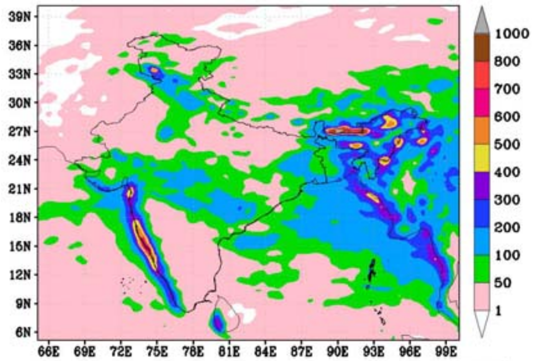
Figs. 2(a-h). Three hourly accumulation of rainfall (mm) computed by the TMPA data accumulated for the monsoon season (June to September) of 2013 from the model run of two days ago. This gives us three sets of MODE parameters which quantify the model forecast behaviour at 0600 UTC of the particular day, based on day 1, day 2 and day 3 forecasts.

The MODE tool identifies areas of rainfall as objects and computes a wide variety of object attributes. As a first step, both the analysis and forecast rainfall data fields above a minimum raw threshold (in the present case 0 mm) is considered for analysis. The rainfall amount and spatial distribution during the monsoon season, is not uniform throughout the Indian region. However, low intensity stratiform precipitation has a substantial contribution to the monsoon rainfall over India (Saha *et al.*, 2014; Sen Roy *et al.*, 2014a; Schumacher and Houze, 2003). While it is true that models have a pre-disposition towards production of spurious drizzle as a result of anomalous super-saturations at cloud edges (Stevens *et al.*, 1996), its exact contribution in the rainfall estimates of WRF with the present configuration over the Indian region, is not very well documented. In the absence of such documentation, in order to avoid biasing the validation scheme against low rainfall events and to objectively assess all rainfall estimates from the model, the minimum raw threshold in the MODE configuration file has been chosen to be 0 mm. The precipitation data is then convolved to replace the precipitation value at a point with its average over the area within a disk whose centroid is located at that point. However, a minimum threshold of 5 mm/3 hours has been chosen as the convolution threshold for defining the objects and to eliminate spatial regions of uniformly low intensity rainfall values. The convolution radius of both forecast and analysis fields has been kept small (60 km by 60 km) to retain the non-uniformity of the rainfall field. The convolved precipitation field results in local, contiguous patches of positive precipitation surrounded by regions of zero values. Matching precipitation areas (objects) in the forecast and observed data field are compared and verification metrics are derived. More details are available in the online User's Guide for MET version 3.0 at Davis *et al.* (2006) and http://www.dtcenter.org/met/users/docs/users_guide/MET_Users_Guide_v3.0_rev2.pdf.

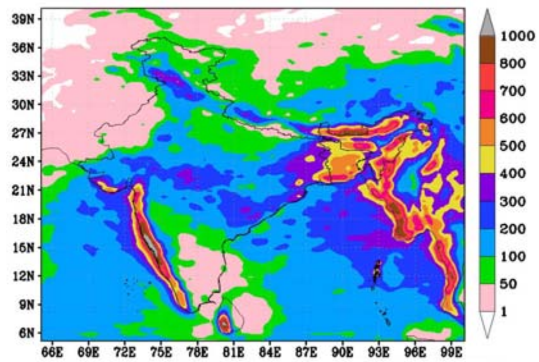
Six verification metrics of the MODE output are selected for detailed study. The first two metrics are (a) the ratio of the number of simple objects and (b) the ratio of the total area of the objects identified in the forecast and analysis fields. Both parameters compare gross nature of the spread of precipitation and the areal distribution of rainfall. The third parameter selected for detailed study is the centroid distance between the matched object pairs in the forecast and observed data fields. Since the centroid position can vary due to (*i*) displacement of the object in the forecast field with respect to the observation, or (*ii*) change in the shape and areal spread of the two matched objects (precipitation areas), this parameter has to be seen in conjunction with the second parameter. This is further confirmed by the fourth MODE parameter which is studied in detail, namely; Critical Success Index (CSI). This parameter is the ratio of the total number of overlapping grid points of the forecast and observed objects, to the total number of grid points of the matching objects in the forecast and observed data fields. A large centroid displacement, accompanied by low CSI score indicates that the centroid displacement error is primarily due to the displacement of the matched objects. On the other hand, a large centroid displacement, accompanied by high object area ratio indicates that the centroid displacement is primarily due to a shape and size error of the matched objects. An absence of high frequency bias score would indicate a displacement error. In addition to the geometric aspect of the objects, one also needs to compare the intensity values in the objects matched. MODE has the ability to identify the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of intensity of the data field within the identified object. The ratio of the values for a pair of matching objects in the WRF forecast and TMPA fields should be close to 1 to indicate similar assessment of rainfall intensity. As already stated in the introduction, previous studies indicate that the model estimates of rainfall tend to be more accurate for low intensity rainfall, and less so for heavy rainfall patches and mesoscale variations of rainfall intensity. The 50^{th} percentile intensity of the matched objects in the forecast and verification fields was selected for comparison, in order to highlight the average values within an object while the 90^{th} percentile was selected to highlight the more intense values in the objects identified.

Each of the six MODE parameters for each of the forecast-observation matched datasets are then accumulated together for the entire season (120 days) corresponding to each forecast hour (120 such pairs corresponding to each 3 hourly forecast up to 72 hours), to obtain the histogram distribution of the MODE parameter for that forecast hour. The contour plot of these histogram distributions is then displayed through 72 hours at three hourly intervals to demonstrate the diurnal variation of the MODE parameter. Thus we obtain six such contour plots demonstrating the verification of the 72 hours of model forecast, which are displayed in Figs. 6(a-f). This methodology of verification of the diurnal cycle characteristics for the entire season, rather than on individual day-by-day basis has been adopted in view of the short duration of most rainfall episodes over India which makes it difficult to discuss the diurnal cycle characteristics on a individual day-by-day basis.

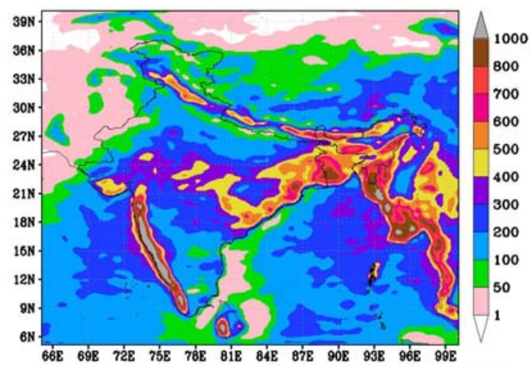
(a) WRF 3-hrly accumulated rainfall forecast field at 03 hrs



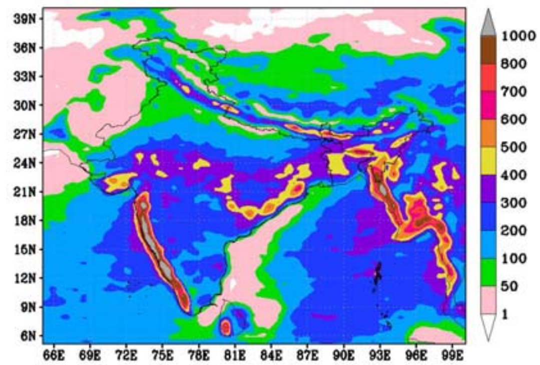
(b) WRF 3-hrly accumulated rainfall forecast field at 06 hrs



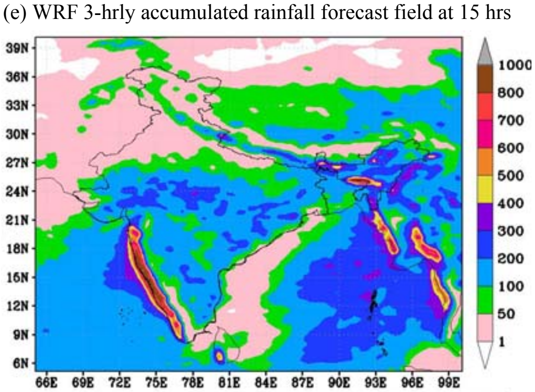
(c) WRF 3-hrly accumulated rainfall forecast field at 09 hrs



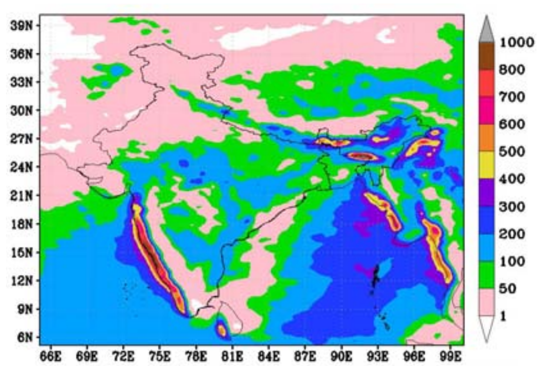
(d) WRF 3-hrly accumulated rainfall forecast field at 12 hrs



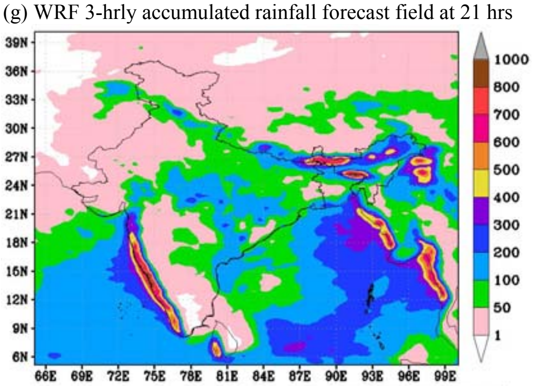
(e) WRF 3-hrly accumulated rainfall forecast field at 15 hrs



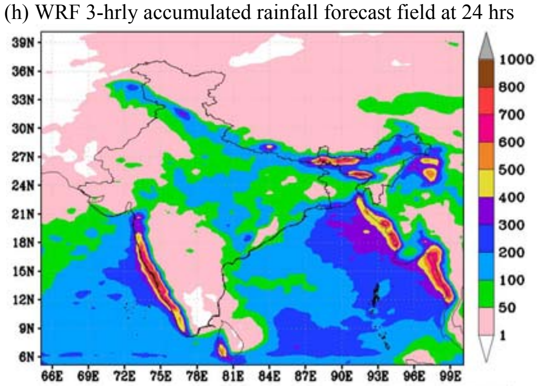
(f) WRF 3-hrly accumulated rainfall forecast field at 18 hrs



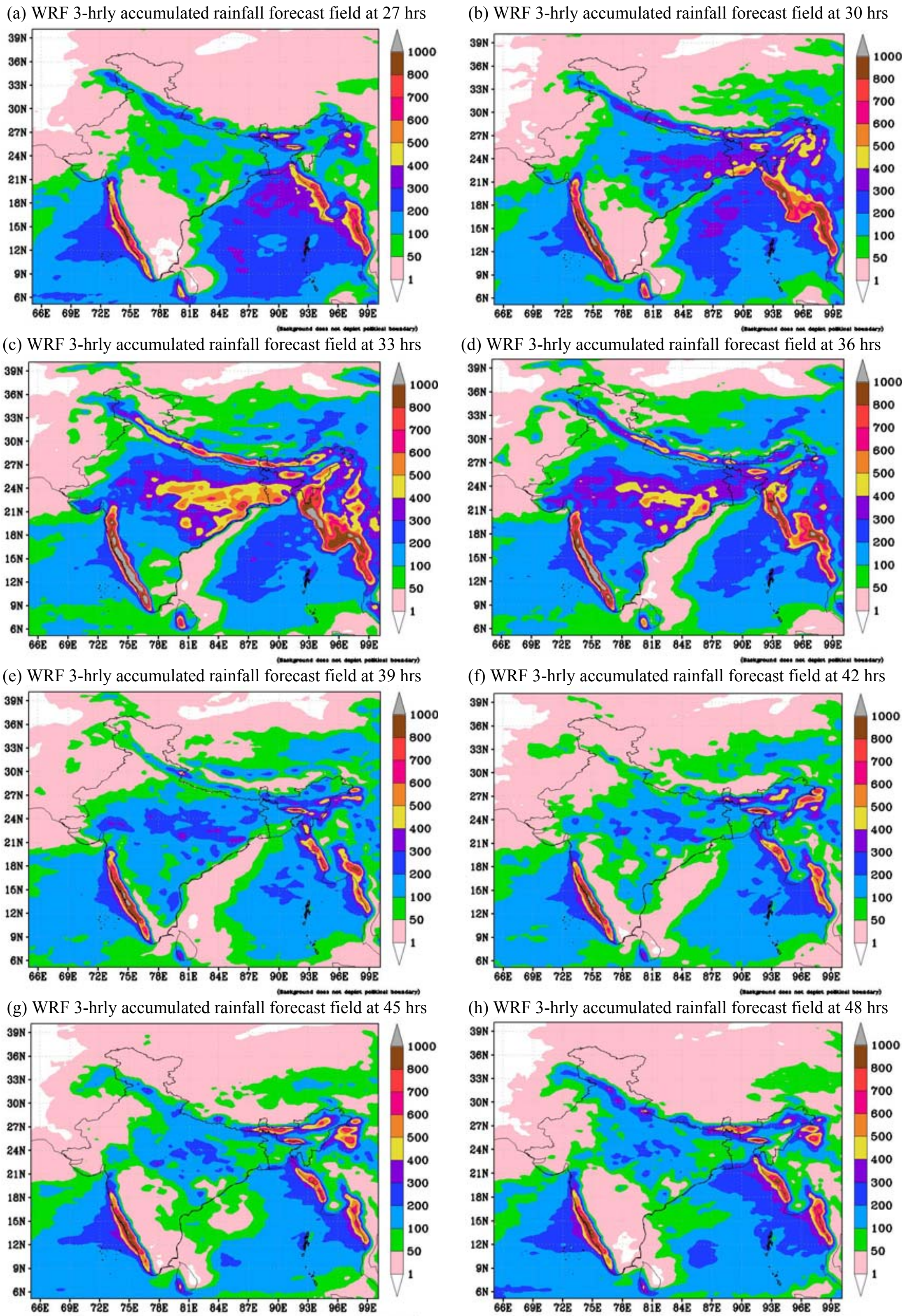
(g) WRF 3-hrly accumulated rainfall forecast field at 21 hrs



(h) WRF 3-hrly accumulated rainfall forecast field at 24 hrs

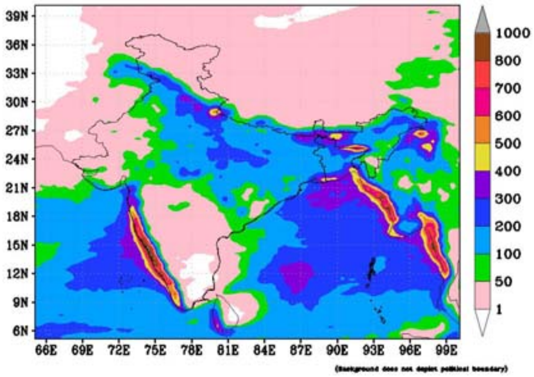


Figs. 3(a-h). Three hourly accumulated rainfall (mm) for 03-24 hour forecast from WRF model initialized at 0000 UTC, accumulated for the monsoon season (June to September) of 2013

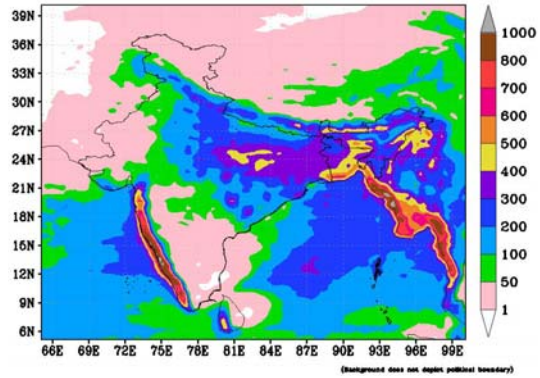


Figs. 4(a-h). Three hourly accumulated rainfall (mm) for 27-48 hour forecast from WRF model initialized at 0000 UTC, accumulated for the monsoon season (June to September) of 2013

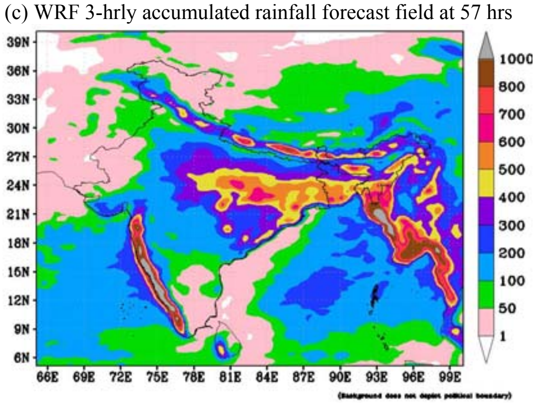
(a) WRF 3-hrly accumulated rainfall forecast field at 51 hrs



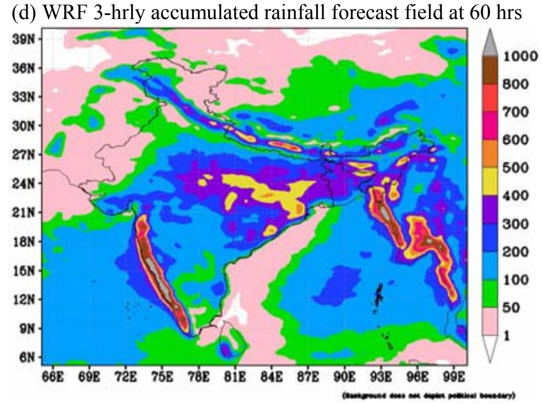
(b) WRF 3-hrly accumulated rainfall forecast field at 54 hrs



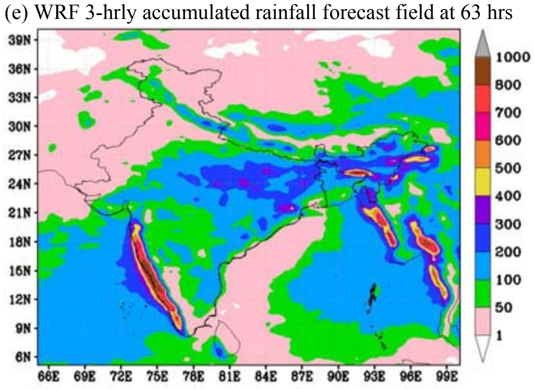
(c) WRF 3-hrly accumulated rainfall forecast field at 57 hrs



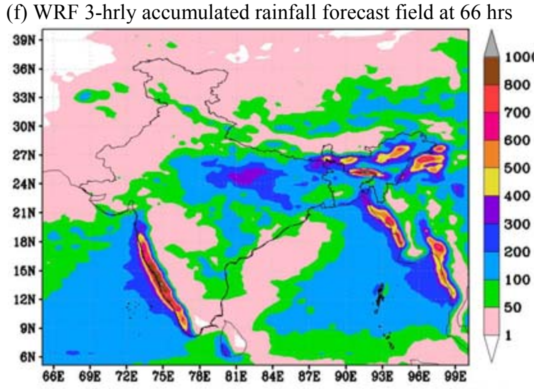
(d) WRF 3-hrly accumulated rainfall forecast field at 60 hrs



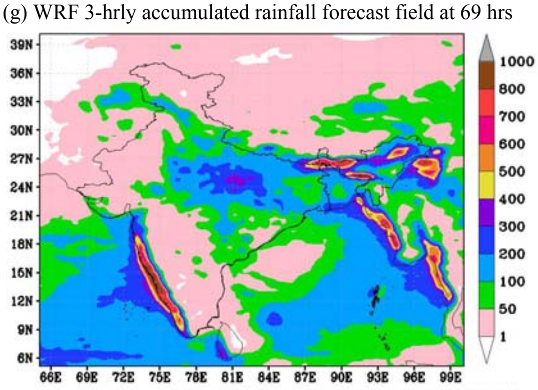
(e) WRF 3-hrly accumulated rainfall forecast field at 63 hrs



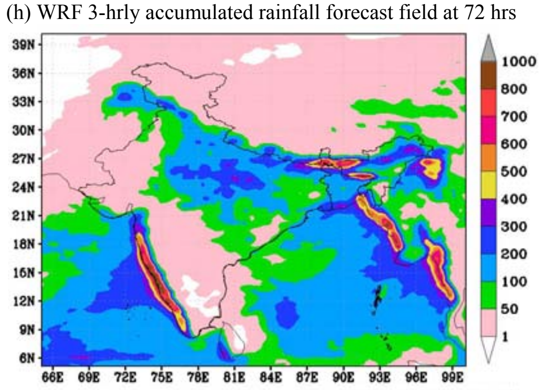
(f) WRF 3-hrly accumulated rainfall forecast field at 66 hrs



(g) WRF 3-hrly accumulated rainfall forecast field at 69 hrs



(h) WRF 3-hrly accumulated rainfall forecast field at 72 hrs



Figs. 5(a-h). Three hourly accumulated rainfall (mm) for 51-72 hour forecast from WRF model initialized at 0000 UTC, accumulated for the monsoon season (June to September) of 2013

3. Results and discussion

Figs. 2(a-h) display the three hourly rainfall observed over the Indian subcontinent during the monsoon season of 2013. There is a clear early morning maxima at 0000 UTC (0530 LT) along the west coast of India and the west coast of Myanmar [Fig. 2(h)] and a minima in rainfall at 1500 UTC [Fig. 2(e)]. Similarly, there is a late night maximum along the foothills of the Himalayas, from the western to the eastern end at around 2100 UTC and a minimum in rainfall amount around 1500 UTC. On the other hand, the core monsoon region over central India, has a maximum value at 0900 UTC to 1200 UTC. These qualitative results are well in line with the quantitative values of rainfall maxima discussed in previous studies (Sen Roy and Sen Roy, 2011; Sen Roy and Balling 2007; Sen Roy and Sen Roy, 2014b etc.). While the mid night rainfall maxima along the Himalayan foothills, is due to the rainfall triggered by the down slope katabatic wind flow from the Himalayas (Sen Roy and Sen Roy, 2014 b), the early morning maxima along the west coasts is due to the diurnal maximum in the monsoon flow inland, in association with land-sea breeze phenomenon (Sen Roy and Balling, 2007). The convection maximum after the local noon over Central India is triggered by the local land heating over the plains (Sen Roy and Sen Roy, 2014b).

The most obvious difference between the observed and forecast rainfall [Figs. 3(a-h)] for 03 to 24 hour forecast, Figs. 4 (a-h) for 27 to 48 hour forecast, Figs. 5 (a-h) for 51 to 72 hour rainfall forecast of WRF model) is that the diurnal maximum in forecast rainfall over the entire validation domain is at around 0900 UTC, close to the local maximum of temperature. This is true for day 1 [Fig. 3(c)], day 2 [Fig. 4(c)] and day 3 [Fig. 5(c)] forecasts. Hence, a broad inference from this result is that the present configuration of the WRF model tends to force convection primarily due to solar heating. The effect of orography and land-sea contrast has less effect in modulating the diurnal cycle of convection in the WRF model. This is similar to the observations of other studies with different models (Basu, 2007), indicating that it may be an inherent weakness of NWP models. However, this requires more investigation. The amount of rainfall and its spatial distribution at the time of the local maximum temperature also shows a day to day variation with the increase in the forecast lead time. For example, the spatial spread of the forecast rainfall maximum at 0900 UTC, is more in the 24 hour forecast [Fig. 3(c)]. The forecast tend to dry up over the Indian region west of 78° E and over the oceanic regions and increases the inland rainfall along the west coasts of India and Myanmar in the 48 hour forecast [Fig. 4(c)] and more so in the 72 hour forecast [Fig. 5(c)]. While these observations all point out discrepancies in the model forecast with respect to the observations in a qualitative sense, a quantitative source of error estimation is necessary to bring the source of error with clearer focus.

As discussed in section 2.2, the MODE software is used for object based quantitative validation of the rainfall forecast and the results are displayed in Figs. 6(a-f). The first parameter to be discussed in detail is the object count ratio of the objects in the forecast to the observed rainfall field [Fig. 6(a)]. While in the initial forecast hours, the modal value of the histograms is centered at 1, indicating perfect match of the object number in the observed and forecast fields, the ratio generally decays with forecast lead time, indicating gradual decrease in number of forecast objects with respect to the number of observed objects with increase in forecast lead time. There is also a strong semi daily scale periodicity in the ratio of total number of objects in the forecast and observed fields with the values clustering about the modal value at 24, 48, 60 and 72 hours of forecast. The areal cut through the three hourly histogram plot of the ratio of the object areas in the forecast and the observed rainfall fields is given in Fig. 6(b). While the ratio is generally symmetric around the mean value of 1, the modal value of the histogram curves for individual hours shifts to larger values at the 9^{th}, 33^{rd} and 57^{th} hour resulting in a diurnal cycle of variation in the contour plot of the histogram curves in Fig. 6(b). This is similar to the findings of Davis *et al.* (2006), who too noted that positive size biases in the model forecast which coincided with the temperature maximum in the afternoon. When these results are seen in conjunction with each other against qualitative results of Figs. 1-4, it reveals that without significant change in the number of objects, the relative sizes of the forecast areas of precipitation increases with respect to the observed areas at the 9^{th}, 33^{rd} and 57^{th} hour forecast, thereby increasing the overall precipitation received at these hours. This may be the reason that the diurnal peak of rainfall occurs at these hours. The peak is sharpest at the 9^{th} hour, slightly less in the 33^{rd} hour and broadest at the 57^{th} hour. This, coupled with the gradual decay in the number of objects in the forecast field indicates a decrease in the diurnal rainfall maxima with increase in the forecast lead time. Davis *et al.* (2006) noted that the WRF model over-predicts precipitation areas at a length scale of 80-120 km and those exceeding 350 km. It also under-predicts rain areas between 150 and 250 km roughly. Our results indicate that the number of predicted rainfall objects generally decrease with increase in the forecast length; although their size changes throughout the day according to the diurnal cycle of temperature.

The third parameter analyzed is the Centroid distance between the matched object pairs of the forecast and observed objects [Fig. 6(c)]. Surprisingly, this

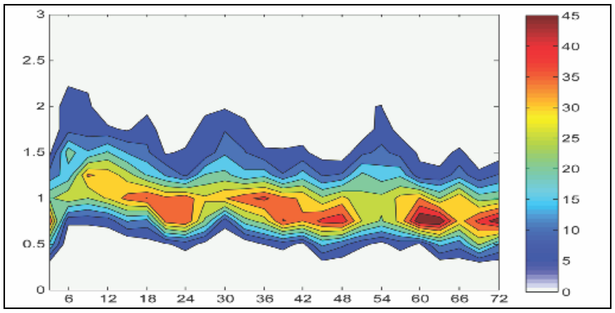


Fig. 6(a). Contour plot of the three hourly Histogram curve of MODE parameter - Ratio of Object Counts (in percentage value) of forecast and observed precipitation field. X-axis values are forecast Hours and Y-axis values are Forecast/observed object count ratio

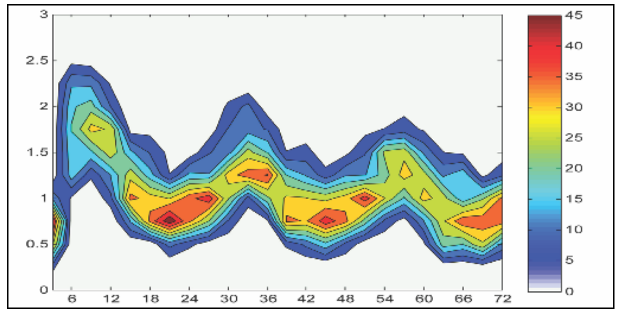


Fig. 6(b). Same as Fig. 6(a) for MODE parameter - Ratio of Object Areas of forecast and observed precipitation field. X-axis values are forecast Hours and Y-axis values are Forecast/observed object area ratio

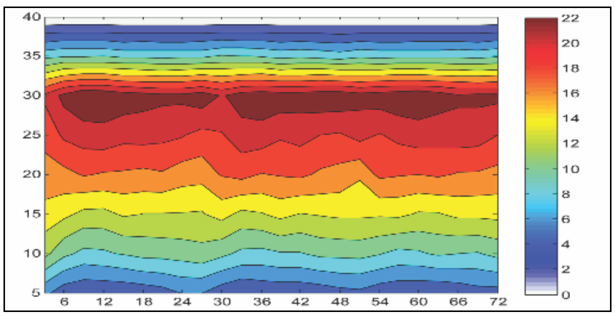


Fig. 6(c). Same as Fig. 6(a) for MODE parameter – Object centroid distance histogram (in percentage value). X-axis values are forecast Hours and Y-axis values are centroid distance in grid units

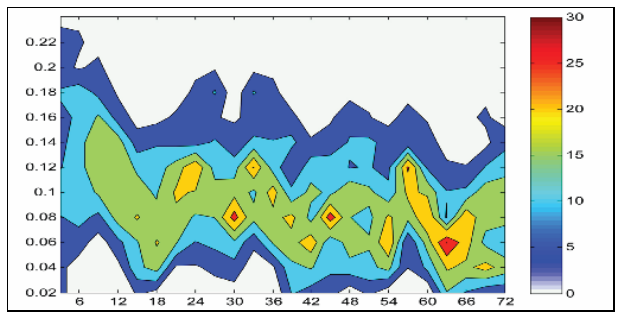


Fig. 6(d). Same as Fig. 6(a) for MODE parameter – Critical Success Index (CSI) histogram (in percentage values). X-axis values are forecast Hours and Y-axis values are error index

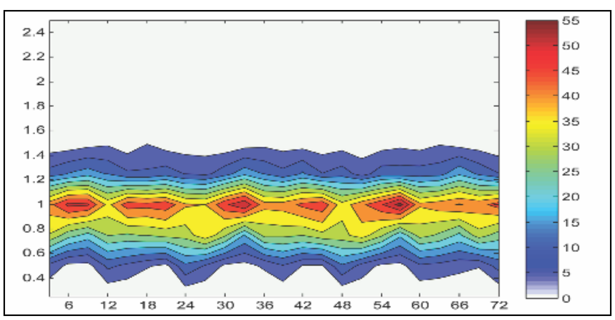


Fig. 6(e). Same as Fig. 6(a) for MODE parameter - 50th Intensity percentile Histogram (in percentage value). X-axis values are forecast Hours and Y-axis values are Forecast/observed ratio of 50th percentile intensity values

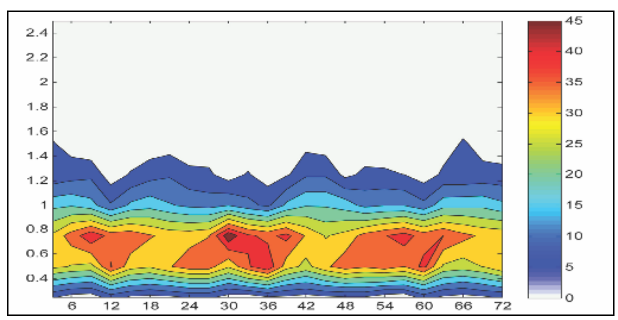


Fig. 6(f). Same as Fig. 6(a) for MODE parameter - 90th Intensity percentile Histogram (in percentage value). X-axis values are forecast Hours and Y-axis values are Forecast/observed ratio of 90th percentile intensity values

parameter has no strong diurnal scale of variation, and tends to be fixed at around 30 grid points at all hours. The fourth parameter considered is the CSI [Fig. 6(d)]. The CSI values also tend to be very low (centered at about 0.1 but decreasing with time). When the above results are seen in conjunction with each other [Figs. 6(b-d)], it is noted that the area errors are not very high in the forecast field at all times, and show a strong diurnal cycle. On the other hand the CSI values are low and show no diurnal cycle. Hence the centroid distance error between the matched objects in the forecast and the verification fields (which has no diurnal cycle) appears to be mostly due to the displacement of the forecast objects with respect to the observations. This matches very well with the findings of other objects based verification studies (Das *et al.*, 2014b).

In addition to the shape and location attributes discussed above, the rainfall forecast by the WRF model is also investigated in terms of accuracy with which it predicts the intensity of the rainfall areas. The ratio of the 50th percentile intensity values of the matched objects in the forecast and observed rainfall data fields compares the central median values of rainfall in a similar patch. The ratio of the 90th percentile values, on the other hand, compares the maximum values of the matching areas in the observed and forecast rainfall fields. By comparing these values for matched objects, one can infer the strength with which the forcing factors in the model are affecting the convection. These are displayed in Figs. 6(e&f). The 50th percentile intensity ratio of the matched forecast and observation objects indicates that the median precipitation values of the forecast and observation areas match very well. While the values were more closely clustered about the mean at specific forecast lead time, namely 9th, 18th, 33rd, 45th and 57th hour forecast, the mean value of the histograms was mostly around 1. It indicates that the median rainfall over the Indian region is captured more accurately at these hours. However, the 90th percentile intensity ratio is generally below 1 for all forecast lead times, with stronger clustering about the mean value at the 9th, 33rd and the 57th hour. It indicates that, in general, the model under-predicts the rainfall intensity of the heavy rainfall zones. This is similar to the findings of Davis *et al.* (2006) who noted that the increase in the object sizes in the model forecast at specific hours, is accompanied by a peaking of the rainfall intensity range. They inferred that this pattern is indicative of the dominance of parameterized convection.

4. Conclusions

The verification of the qualitative and quantitative aspects of the rainfall prediction by the operational WRF model brings to light some interesting aspects of model errors on the sub-daily scale of rainfall prediction. The study reveals that there is a clear climatological early morning maximum during the monsoon season along the west coast of India and west coast of Myanmar and a minimum in rainfall in the late evening over the same region. There is a late night maximum along the foothills of the Himalayas, from the western to the eastern end and a minimum in rainfall amount around late evening. On the other hand, the core monsoon region over central India, has a maximum value at 0900 UTC to 1200 UTC. While the midnight rainfall maximum is due to the rainfall triggered by the down slope katabatic wind flow from the Himalayas (Sen Roy and Sen Roy, 2014b), the early morning maxima along the west coasts is due to the diurnal maximum in the monsoon flow inland in association with land-sea breeze phenomenon (Sen Roy and Balling, 2007). The convection maximum after the local noon over Central India is triggered by the local land heating over the plains (Sen Roy and Sen Roy, 2014b).

The analysis highlights the model bias towards convection initiated by surface sensible heating. This causes the diurnal maximum in rainfall to be nearly in phase with the diurnal maximum in temperature. Orography and land-sea contrast has less effect in modulating the diurnal cycle of convection in the WRF model. The diurnal maximum in rainfall in the model appears to be due to the increase in the predicted area of rainfall at 0900 UTC, rather than increase in the number of rainfall elements. This indicates that the forcing due to sensible heat in the present configuration of the WRF model may be increasing the cloud organization rather than generating fresh convection zones to produce the diurnal rainfall maximum in the forecast field. In fact, the number of rainfall areas in the model decrease with increase in forecast lead time. The diurnal cycle errors are highest over regions, where other forcing factors are predominate; especially along the coasts and the foothills of the Himalayas. There is a strong component of displacement error (about 30 grid units) in the WRF model forecast. This may cause large errors in location specific forecasts, where a displacement of a grid, can give widely varying forecast values. This has to be analyzed in greater detail to find out whether the displacement arises due to errors in the initial analysis field due to the absence of proper dense and good quality observations. The forecast of mean climatological rainfall is generally accurate in the model (the 50th percentile or the median value). However, the higher rainfall patches within the same precipitation area tend to be underestimated. One may infer that the model forecast is generally smoothened, and does not adequately pick up the mesoscale features of a rainfall area.

Acknowledgements

Authors are grateful to Dr. L. S. Rathore, Director General of Meteorology, India Meteorological Department, New Delhi for his encouragement and providing all the facilities to carry out this research work. We are thankful to NASA Tropical Rainfall Measuring Mission's Goddard Earth Sciences Data and Information Services Center (GESDISC) for providing TRMM data for this project.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E., 2003, "The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-Present)", *J. Hydrometeorol.*, **4**, 1147-1167.

Basu, B. K., 2007, "Diurnal Variation in Precipitation over India during the Summer Monsoon Season: Observed and Model Predicted", *Mon. Wea. Rev.*, **135**, 2155-2167.

Brown, B. G., Bullock, R., Halley, G. J., Ahijevych, D., Davis, C., Gilleland, E. and Holland, L., 2007, "Application of the MODE object-based verification tool for the evaluation of model precipitation fields", AMS 22nd Conference on Weather Analysis and Forecasting and 18th Conference on Numerical Weather Prediction, 25-29 June, Park City, Utah, American Meteorological Society (Boston), available at http://ams.confex.com/ams/ pdfpapers/124856.pdf.

Das, A. K., Bhowmik, M., Kundu, P. K. and Roy Bhowmik, S. K., 2014, "Verification of WRF Rainfall Forecasts over India during Monsoon 2010: CRA Method", *Mausam*, **66**, 3, 385-396.

Davis, C. A., Brown, B. G. and Bullock, R. G., 2006, "Object-based verification of precipitation forecasts, Part I: Methodology and application to mesoscale rain areas" *Mon. Wea. Rev.*, **134**, 1772-1784.

Durai, V. R., Roy Bhowmik, S. K. and Mukhopadhayay, B., 2010, "Evaluation of Indian Summer Monsoon rainfall features using TRMM and KALPANA-I derived precipitation and raingauge observation", *Mausam*, **61**, 3, 317-336

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P. and Stocker, E. F., 2007, "The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined- sensor precipitation estimates at fine scales", *J. Hydrometeorol.*, **8**, 38-55.

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.

Kotal, S. D., Sen Roy, S. and Roy Bhowmik, S. K., 2014, "Catastrophic heavy rainfall episode over Uttarakhand during 16-18 June, 2013- observational aspects", *Current Science*, **107**, 2, p234.

Mandal, V., De, U. K. and Basu, B. K., 2007, "Precipitation Forecast Verification of the Indian Summer Monsoon with Intercomparison of Three Diverse Regions", *Wea. Forecasting*, **22**, 428-443.

Nair, S., Srinivasan, G. and Nemani, R., 2009, "Evaluation of Multi-Satellite TRMM Derived Rainfall Estimates over a Western State of India", *Journal of the Meteorological Society of Japan,* **87**, 6, 927-939

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S. and Mukhopadhyay, B., 2014, "Development of a new high spatial resolution (0.25 degrees × 0.25 degrees) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region", *Mausam*, **65**, 1, 1-18.

Ranade, A., Mitra, A. K., Singh, N. and Basu, S., 2014, "A verification of spatio-temporal monsoon rainfall variability across Indian region using NWP model output," *Meteorology and Atmospheric Physics*, **125**, 1-2, 43-61.

Rahman, S. H., Sengupta, D. and Ravichandran, M., 2009, "Variability of Indian summer monsoon rainfall in daily data from gauge and satellite", *J. Geophys. Res.*, **114**, D17113, doi:10.1029/2008JD011694.

Roy Bhowmik, S. K. and Durai, V. R., 2010, "Application of multimodel ensemble techniques for real time district level rainfall forecasts in short range time scale over Indian region", *Meteorology and Atmospheric Physics*, **106**, 1-2, 19-35.

Roy Bhowmik, S. K. and Prasad, K., 2001, "Some characteristics of limited-area model-precipitation forecast of Indian monsoon and evaluation of associated flow features", *Meteorology and Atmospheric Physics*, **76**, 3-4, 223-236.

Schumacher, C. and Houze Jr, R. A., 2003, "Stratiform rain in the tropics as seen by the TRMM precipitation radar", *Journal of Climate*, **16**, 11, 1739-1756.

Saha, S. B., Sen Roy, S., Bhowmik, S. K. and Kundu, P. K., 2014, "Intra-seasonal variability of cloud amount over the Indian subcontinent during the monsoon season as observed by TRMM precipitation radar", *Geofizika*, **31**, 1, 29-53.

Sen Roy, S., Saha, S. B., Roy Bhowmik, S. K. and Kundu, P. K., 2014 a, "Analysis of monthly cloud climatology of the Indian subcontinent as observed by TRMM precipitation radar", *International Journal of Climatology*, DOI: 10.1002/ joc.4108.

Sen Roy, S. and Sen Roy, S., 2014 b, "Diurnal variation in the initiation of rainfall over the Indian subcontinent during two different monsoon seasons of 2008 and 2009", *Theoretical and Applied climatology*, **117**, 1-2, 277-291.

Sen Roy, S., Saha, S. B., Fatima, H., Roy Bhowmik, S. K. and Kundu, P. K., 2012, "Evaluation of short-period rainfall estimates from Kalpana-1 satellite using MET software", *Journal of Earth System Science*, **121**, 5, 1113-1123.

Sen Roy, S. and Sen Roy, S., 2011, "Regional variability of convection over northern India during the pre-monsoon season", *Theoretical and Applied Climatology,* **103**, 1-2, 145-158

Sen Roy, S. and Balling, R. C., 2007, "Diurnal variations in summer season precipitation in India", *International Journal of Climatology*, **27**, **7**, 969-976.

Stevens, B., Feingold, G., Cotton, W. R. and Walko, R. L., 1996. "Elements of the microphysical structure of numerically simulated nonprecipitating stratocumulus", *Journal of Atmospheric Sciences*, **53**, 980-1006.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I. and Hamada, A., 2009, "A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges", *Sola*, **5**, 137-140.