NMRF/VR/MAY/2021

VERIFICATION REPORT

NCUM Global Model Verification: Monsoon (JJAS) 2020

S. Karunasagar, Mohana S. Thota, K. Niranjan Kumar, Abhishek Lodh, Anumeha Dube and Raghavendra Ashrit

May 2021

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences, Government of India
A-50, Sector-62, NOIDA-201 309, INDIA

Data Control Sheet

1	Name of the Institute	National Center for Medium range weather Forecasting
2	Document Number	NMRF/VR/MAY/2021
3	Date of Publication	May 2021
4	Title of the document	NCUM Global Model Monthly Verification: Monsoon 2020
5	Type of the document	Verification Report
6	Number of pages, figures and Tables	46 pages, 38 figures, 2 Tables
7	Authors	S. Karunasagar, Mohana S. Thota, K. Niranjan Kumar, Abhishek Lodh, Anumeha Dube and Raghavendra Ashrit
8	Originating Unit	National Centre for Medium Range Weather Forecasting (NCMRWF), A-50, Sector-62, NOIDA201 309, India
9	Abstract	This report documents performance of the NCMRWF model forecasts during monsoon season (JJAS) 2020. The verification results are presented to address (a) forecasters and (b) model developers. The information on biases in the forecast winds, temperature humidity, rainfall, etc., are crucial for the forecasters to interpret the model guidance for forecasting. Additionally, information on recent improvements in the model skill adds to confidence in the model forecasts.
10	References	
11	Security classification	Unrestricted
12	Distribution	General

Table of Contents

1. Introduction

This report documents performance of the NCMRWF model forecasts during monsoon season (JJAS) 2020. Forecast verification is carried out against the model analysis and observations. The results are summarized for the season as a whole to understand the average biases and forecast performances. The report is oriented towards both (a) forecasters and (b) model developers. Section 3 to 5 of the report discusses the systematic biases in forecast large scale upper fields, namely, wind, temperature humidity, and rainfall, etc., which are expected to be useful for the forecasters to interpret the model forecasts. Additionally, recent improvements in forecast skills are also presented for the benefit of forecasters and in general for monitoring the performance of the forecasting system. Section 6-7 of the report presents diagnostic analysis and focuses on verification of intra-seasonal, features like active/break spells, the onset of monsoon, synoptic scale variability, along with investigation of some of the model biases and understanding of the sources to provide feedback to modelers. Section 2 describes the recent development in the NCUM-G model and data assimilation system at NCMRWF. Observational data sets used in this study are also briefly discussed. Section 8 gives a summary of the results.

2. NCMRWF Unified Modelling System& Observed Data

2.1 Model Description

The NCMRWF Unified model (NCUM) was implemented in 2012 (Figure 1) with a grid resolution of 25km (NCUM-G:V1) which was upgraded to 17km (NCUM-G:V3) in 2015, 12km (NCUM-G:V5) in 2018. The present version (NCUM-G:V6) of NCUM-G has a horizontal grid resolution of ~12 km with 70 levels in the vertical reaching 80 km height. It uses "ENDGame" dynamical core, which provides improved accuracy of the solution of primitive model equations and reduced damping. This helps in producing finer details in the simulations of synoptic features such as cyclones, fronts, troughs, and jet stream winds. ENDGame also increases variability in the tropics, which leads to an improved representation of tropical cyclones and other tropical phenomena (Walters et al., 2017). An advanced data assimilation method of Hybrid 4-Dimensional Variational (4D-Var) is used for the creation of NCUM global analysis. The advantage of the Hybrid 4D-Var is that it uses a blended background error, a blend of "climatological" background error, and day-to-day varying flow dependent background error (derived from the 22–member ensemble forecasts). The hybrid approach is scientifically attractive because it elegantly combines the benefits of ensemble data assimilation (flow-dependent co-variances) with the known benefits of 4D-Var within a single data assimilation system (Barker, 2011). A brief description of the NCUM Hybrid 4D-Var system is given in Kumar et al. (2020 & 2018). More details about the DA system are presented in section 2.2.

NCMRWF Unified Model (NCUM) Global Data Assimilation System

Implementation Date	NCUM System (Version)	Data Assimilation (Atmosphere)	Land Data Assimilation	Atmos. Model Resolution
April-2012	NCUM-G:V1	4D-Var	UKMO Surface files	25 km (UM : N512 L70)
Dec-2012	NCUM-G:V2	4D-Var	Soil Moisture nudging	25 km (UM : N512 L70)
Nov-2015	NCUM-G:V3	4D-Var	EKF (Soil Moisture)	17 km (UM : N768 L70)
Oct-2016	NCUM-G:V4	Hybrid 4D-Var (Ensemble forecasts from ETKF system)	EKF (Soil Moisture)	17 km (UM : N768 L70)
May-2018	NCUM-G:V5	Hybrid 4D-Var (Use of High resolution ensemble forecasts)	EKF (Soil Moisture, LST)	12 km (UM : N1024L70)
June-2020	**NCUM-G:V6**	**Hybrid 4D-Var (Use of High resolution ensemble forecasts)**	**EKF (Soil Moisture, LST)**	**12 km (UM 11.2 : N1024L70)**

4D-Var (Atmosphere)

Started Surface Analysis

Improved Land Data Assimilation-EKF

Hybrid 4D-Var (Atmosphere)

BE from high Resolution Ensemble. forecast

Cloud Affected Radiance Assimilation (AMSU-A)

Figure 1. Recent developments and upgrades in the NCUM-G Modelling and Data Assimilation system at NCMRWF

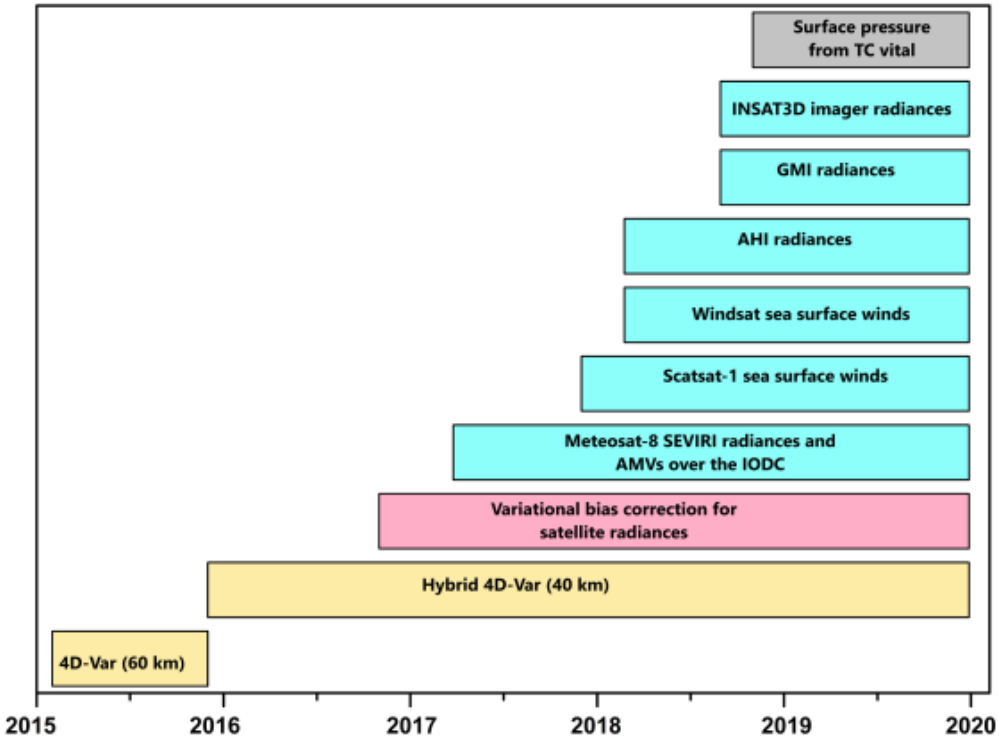

Figure 2. Evolution of Data Assimilation (DA) system at NCMRWF since 2015

2.2 Data Assimilation

The basic data assimilation technique adapted in the NCUM-G is the 4D-Var data assimilation method based on Rawlins et al., (2007) for the generation of 6 hourly atmospheric analyses since its operational implementation at NCMRWF in 2012.4D-Var allows more effective use of observations through the consistent use of observation operators in the model equations (Rabier et al., 1998, 2007). The NCUM-G DA has undergone many modifications during the study period, May 2015-2019. A pictoral summary is given in Figure 2 describing the timeline of the changes in the NCUM-G data assimilation system and the new observations introduced during this period. During the early part of 2015, NCUM-G (NCUM-G:V2) 4D-Var assimilation was carried out at 60 km resolution with the baseline observations shown in Table 1, while the forecast model resolution was ~25 km (Rajagopal et al., 2012). The data assimilation resolution was increased to 40 km during late 2015 (November) while increasing the forecast model resolution to 17 km (George et al., 2016). A weakness of the basic 4D-Var method is it uses a fixed "climatological" model of the error covariance in the background forecast, which lacks the flow dependent error of the day (Lorenc 2003). To address this, the NCUM-G 4D-Var data assimilation system has been upgraded to Hybrid 4D-Var following Clayton et al. (2013) and implemented operationally in October 2016 (Kumar et al., 2018). The term "hybrid" in the Hybrid 4D-Var method refers to the combination of a climatological covariance model with covariances calculated from an ensemble of forecasts, which are designed to sample the current uncertainty. NCMRWF runs an ensemble prediction system based on NCUM-G, the NCMRWF Ensemble Prediction System, NEPS (Sarkar et al., 2016, Mamgain et al., 2018). The NCUM-G Hybrid 4D-Var assimilation system combines the flow dependent errors calculated from the Ensemble Transform Kalman Filter (ETKF) based NEPS forecasts with the climatological background errors. The hybrid approach is scientifically attractive as it elegantly combines the benefits of ensemble data assimilation with the benefits of 4D-Var within a single data assimilation system (Barker, 2011).

Satellite instruments do not measure temperature, wind, or humidity, which are the main prognostic variables in the numerical model governing equations but measure the radiance that reaches the satellites at a given frequency. Before assimilating the radiances (as brightness temperatures) in the NWP models, satellite measured radiances have to be bias-corrected. The bias correction method can be static or adaptive irrespective of the data assimilation system. A static bias correction scheme based on Harris and Kelly (2001) was used in the NCUM-G assimilation system till October 2016, using 1000 to 300 hPa thickness, 200 to 50 hPa thickness, and total column water vapor as predictors. Since October 2016, the bias correction scheme has been changed from static to adaptive. The adaptive bias correction, commonly known as Variational bias correction (VarBC) is based on Cameron and Bell (2018) and it closely follows the incremental formulation based on Auligne et al.,

(2007), while keeping the same set of predictors as that of the static method. In VarBC, the bias correction is applied through the bias control variable in the variational analysis equation, and the data assimilation system keeps track of the bias parameters for radiances from all available satellites/instruments.

In addition to the above changes in the data assimilation method and bias correction, several observations were also introduced in the operational NCUM-G data assimilation system, particularly in the post-2017 period. They include, the radiances and atmospheric motion vectors from Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard Meteosat-8 relocated to the Indian Ocean Data Coverage (IODC) in February 2017, scatterometer sea surface winds from Scatsat-1 operated by Indian Space Research Organization (ISRO) since November 2017, Advanced Himawari Imager (AHI) radiances from Himawari-8 satellite and sea surface winds from Windsat since May 2018, Global Precipitation Mission (GPM) Microwave Imager (GMI) radiances and INSAT-3D imager radiances (Rani et al., 2019) since August 2018. For the accurate prediction of the location and track of the tropical cyclones, the location and the magnitude of the low pressure in the model analysis should be as realistic as possible. In the NCUM-G, the location and the minimum low pressure associated with the tropical cyclone are adjusted by assimilating the estimated surface pressure information from the Tropical Cyclone Vital (TC Vital) reports since October 2018. Figure 2 shows the timeline of the changes in the data assimilation system including the new observations during the period January 2015 to December 2019.

2.3 Observed Data

The seasonal mean analysis and anomalies are studied using the ERA-5 (Hershbach et al., 2020) climatology (1979-2018). The high resolution (12km) NCUM-G analysis data is interpolated to ERA-5 grid resolution (0.25 x0.25). For verification of the forecasts, the NCUM-G model analysis is used. All systematic errors are computed at a native grid resolution of (12km).

Detailed quantitative rainfall forecast verification is based on the IMD-NCMRWF daily high resolution (0.25°) rainfall analysis (Mitra et al. 2009, 2013). The rainfall analysis objectively analyses IMD daily rain gauge observations onto a 0.25° grid using a successive corrections technique with the GPM Satellite rainfall providing the first guess estimates. The model forecasts are gridded to the 0.25° observed rainfall grids over Indian land regions for 122 days from 1st June to 30th September 2020. As noted by Mitra et al. (2009), the merged analysis at 0.25° is appropriate for capturing the large scale rain features associated with the monsoon. The merging of the IMD gauge data into GPM estimates not only corrects the mean biases in the satellite estimates but also improves the large-scale spatial patterns in the satellite field, which is affected by temporal sampling errors (Mitra et al. 2009).

Table 1. List of Satellite based observation types and variables assimilated in the NCUM-G data assimilation system

Observation Type	Observation Description	Assimilated Variables
ABI	GOES Advanced Baseline Imager (ABI) radiances	*Brightness Temperature (*T_b)
AHI	Advanced Himawari Imager radiances from Himawari-8	T_b
Aircraft	Upper-air wind and temperature from aircraft	u, v, T
AIRS	Atmospheric Infrared Sounder of MODIS	T_b
AMSR	Radiances from AMSR-2 onboard GCOM-W1 satellite	T_b
ATOVS	AMSU-A (including cloud affected radiances), AMSU-B/MHS from NOAA-18 &19, MetOp-A&B	T_b
ATMS	Advanced Technology Microwave Sounder in NPP& NOAA20 satellites	T_b
CrIS	Cross-track Infrared Sensor observations in NPP & NOAA20 satellite	T_b
FY3C	MWHS radiances from FY3C	T_b
GMI	Global Precipitation Measurement (GPM) Microwave Imager (GMI) instrument	T_b
GPSRO	Global Positioning System Radio Occultation observations from various satellites	Bending Angle
GroundGPS	Ground based GPS observations from various locations	Zenith Total Delay
IASI	Infrared Atmospheric Sounding Interferometer from MetOp-A&B	T_b
IN3DImgr	INSAT-3D Imager Radiances	T_b
SAPHIR	SAPHIR microwave radiances from Megha-Tropiques	T_b
Satwind	Atmospheric Motion Vectors from various geostationary and polar orbiting satellites (including INSAT-3D& INSAT-3DR)	u, v
Scatwind	Advanced Scatterometer in MetOp-A & B, ScatSat-1, WindSat	u, v
SEVIRI	Cloud clear observations from SEVIRI of METEOSAT 8 & 11	T_b
Sonde	Radiosonde (TAC & BUFR),Pilot balloons, Wind profiles &Radar VAD winds	u, v, T, q
Surface	Surface observations over Land and Ocean (TAC & BUFR), TC bogus *(*Surface Pressure)	u, v, T, q, P_s
SSMIS	SSMIS Radiances	T_b
Radar VAD Winds (Indian DWR)	VAD Winds	u,v

3. NCUM-G Analysis Mean and Anomalies during JJAS 2020

The NCUM-G mean analysis fields and anomalies relative to climatology are assessed in this section during JJAS 2020. The discussion is presented for Winds, Temperature, and Relative Humidity for four standard levels of 850, 700, 500, and 200 hPa levels. The anomalies are computed against the ERA5 climatology (1979-2018).

The mean winds and anomalies at 850& 700hPa levels from NCUM-G analysis are shown in Figure 3a-d. The 850 hPa winds feature (a) cross-equatorial flow (b) strong south-westerlies over the Arabian Sea with core winds exceeding 18 m/s (c) south-westerlies over Bay of Bengal (d) westerlies/south-westerlies over peninsular India, central India, and easterlies/south easterlies over IG plains and foothills of Himalayas with monsoon trough embedded in between and (e) south-westerlies and southerlies widespread over the SE Asia. The anomaly winds in Figure 3c show (a) strong cross-equatorial flow (b) stronger westerlies over the Arabian Sea near the region of strong monsoon flow, as indicated by the positive anomaly in blue-shaded regions. The negative anomaly indicated in red shade suggests (c) weaker south-westerlies over the Bay of Bengal (d) weaker westerlies/south-westerlies over peninsular India and central India (e) widespread weaker southwesterlies and southerlies over SE Asia. At 700 hPa level also (Figure 3b) the mean winds show features very similar to flow at 850 hPa. Northeasterly flow is prominent over the northern Arabian Sea which was not seen at 850 hPa. The wind anomalies at 850 and 700 hPa show prominently widespread cyclonic anomaly circulation over the Arabian Sea indicating a stronger monsoon current over the Arabian Sea. Similarly, anticyclonic anomaly circulation over the head Bay of Bengal is seen at both levels, suggesting weaker SW winds over the Bay of Bengal. Similarly, over the Indian land region, the negative anomalies over peninsular India and western India indicate weaker winds relative to ERA-5 climatology. A similar analysis is presented for 500 and 200 hPa in Figure 4a-d. At 500 hPa also the SW flow over the Arabian Sea (*Bay of Bengal & peninsula*) is stronger (*weaker*) as indicated by a positive (*negative*) anomaly in the blue (*red*) shaded region. Weaker south-westerlies indicated by negative anomalies over the Bay of Bengal and peninsular India over all three levels discussed extend up to western parts of India and into adjoining Arabian Sea.

At 200 hPa (Figure 4b), the mean circulation prominently features the (a) strong westerlies over the north of India (b) Tibetan anticyclone (b) widespread easterlies/northeasterly over the peninsula and the Indian Ocean. The anomaly winds (Figure 4d) indicate stronger than normal (a) westerlies over the north of India (b) easterlies and northeasterly over western Indian Ocean, Arabian Sea, Peninsular India, northern Bay of Bengal, and parts of SE Asia, as indicated by a positive anomaly in blue shade. (c) Over large parts of the eastern Indian Ocean, the southern Bay of Bengal, and the South China Sea the analysis features weaker easterlies and northeasterly (a negative anomaly in the red shade). (d) The anticyclonic anomaly flow over near 100°E indicates that the analysis features a stronger Tibetan anticyclone with an eastward shift relative to climatology.

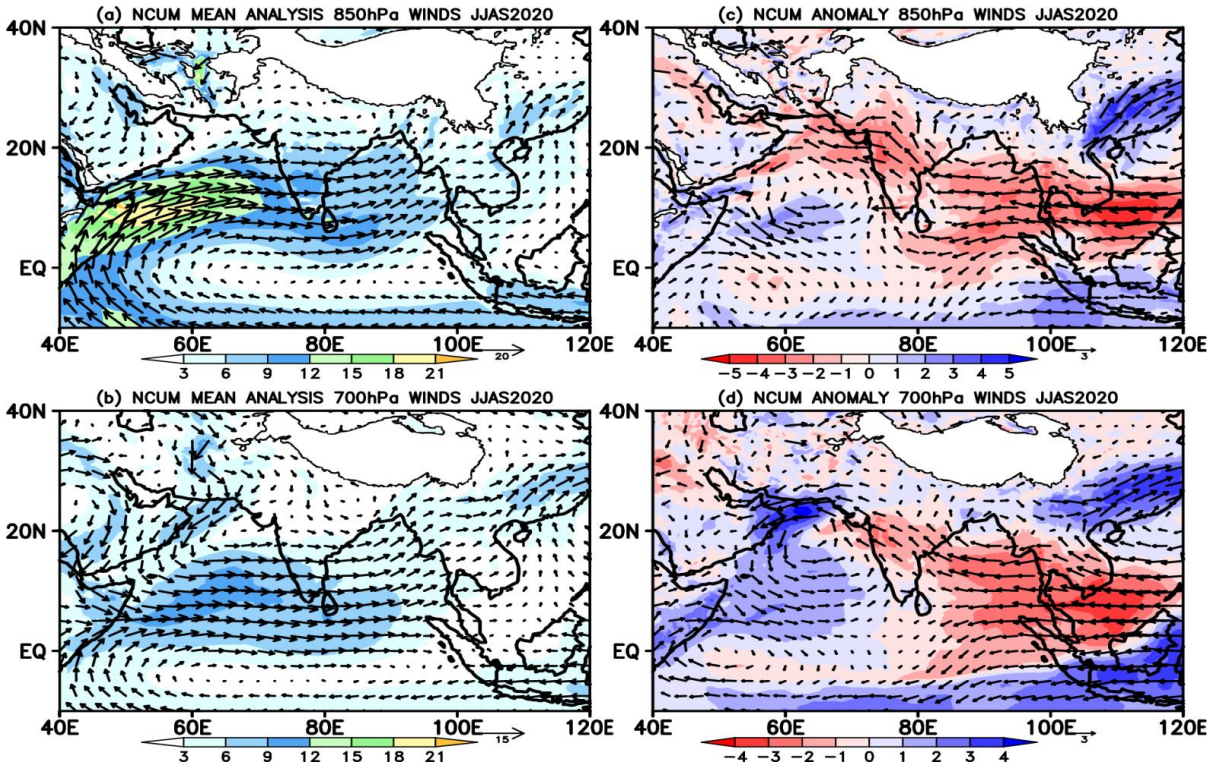

Figure 3. Analysis mean wind at(a) 850 hPa(b) 700 hPa. Wind anomaly at (c) 850 hPa (d) 700 hPa levels.

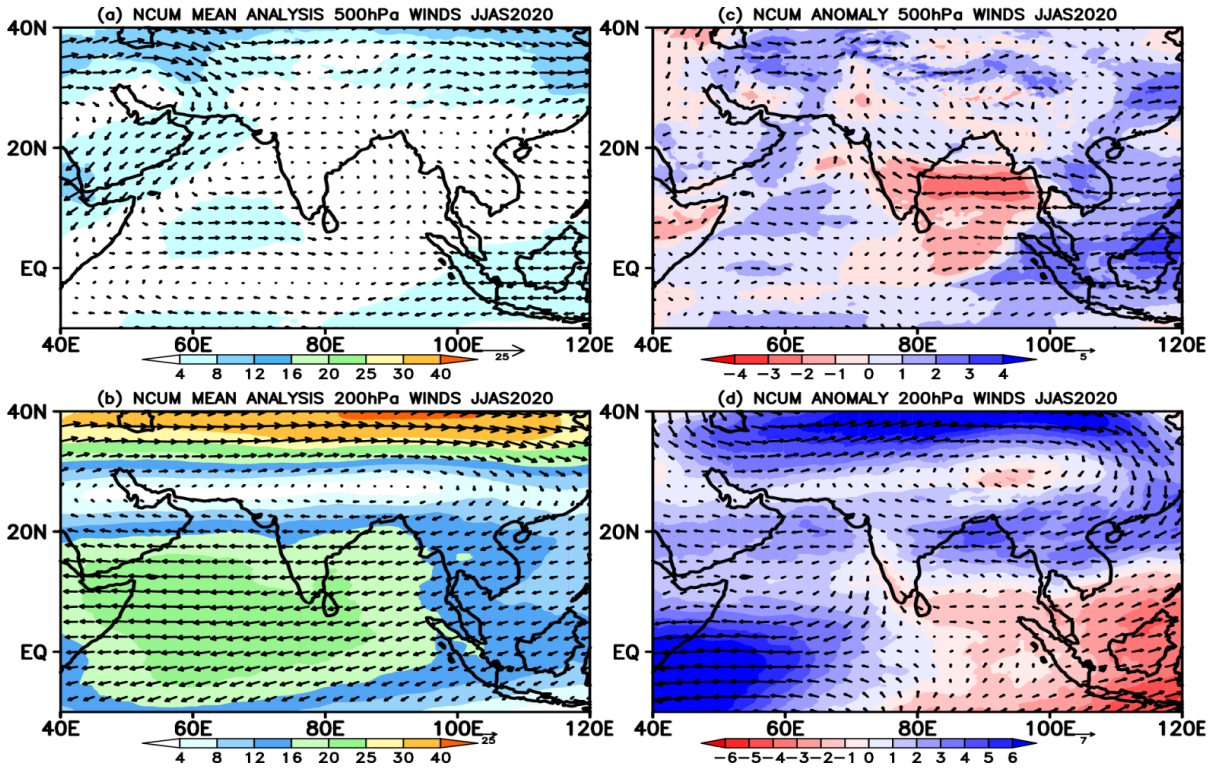

Figure 4. Analysis mean wind at (a) 500 hPa(b) 200 hPa. Wind anomaly at (c) 500 hPa (d) 200 hPa levels.

A similar analysis for mean temperature and anomalies at 850 and 700 hPa levels are shown in Figure 5a-d. The positive anomalies (1-2°C; red) are widespread, except over the northern Arabian Sea. The negative anomalies

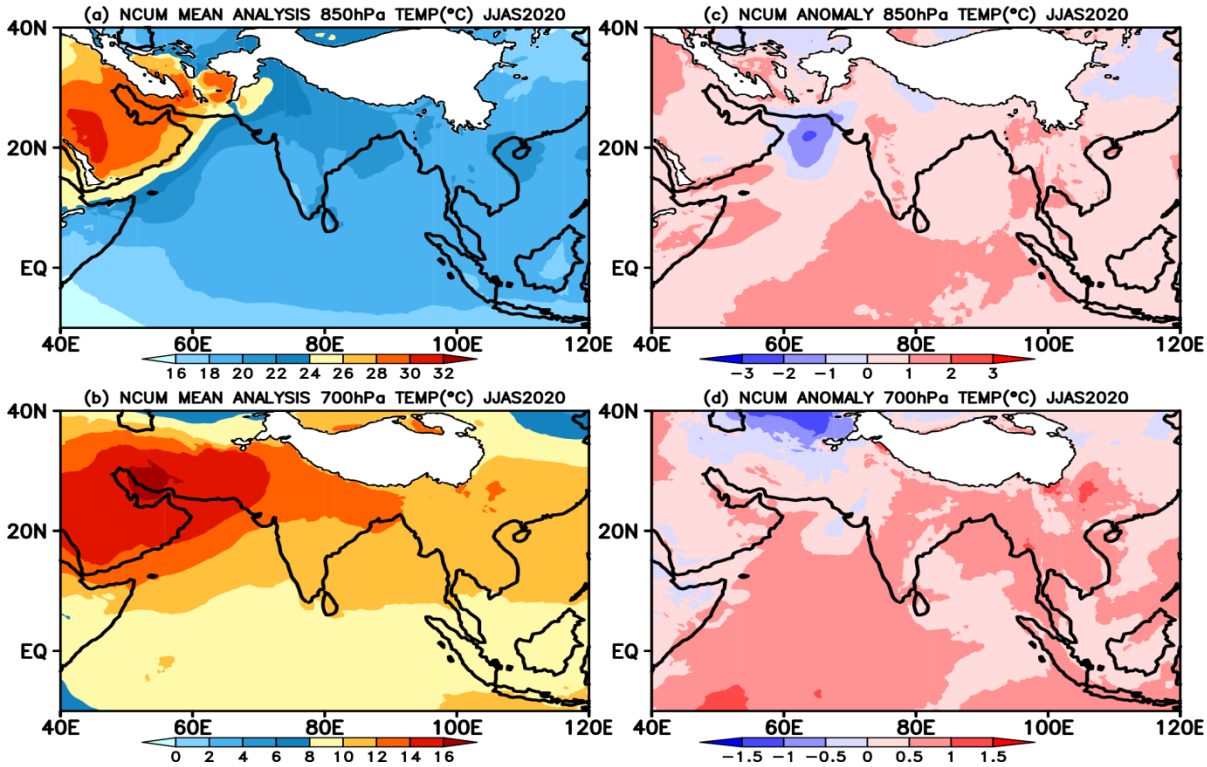

Figure 5. Analysis mean temperature at (a) 850 hPa(b) 700 hPa. Temperature anomaly at (c) 850 hPa (d) 700 hPa levels.

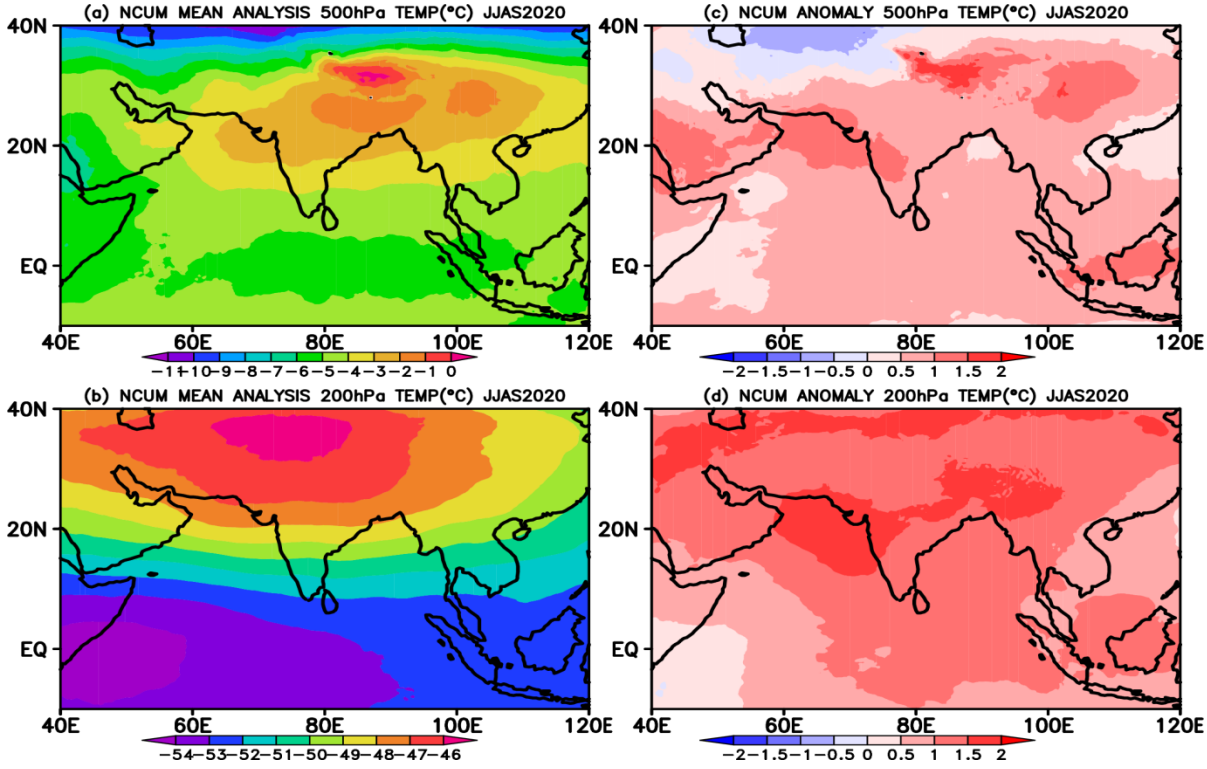

Figure 6. Analysis mean temperature at (a) 500 hPa(b) 200 hPa. Temperature anomaly at (c) 500 hPa (d) 200 hPa levels.

(-2°C at 850 hPa) over the northern Arabian Sea could be due to weak westerlies at 850 hPa (Figure 3c) and strong northeasterly flow from land, particularly at 700 hPa (Figure 3d).

Mean temperature and anomalies at 500 and 200 hPa levels are shown in Figure 6a-d. The positive anomalies (>1°C at 500 and >2°C at 200 hPa level) are seen widespread over the entire domain.Negative anomalies in temperature at 700 and 500 hPa over central Asia (30-40°N 50-70°E) could be speculated as due to the enhanced flow from mid-latitude westerlies in the analysis over the same region at 700 hPa(Figure 3d) and more prominent at 500 hPa level (Figure 4c).

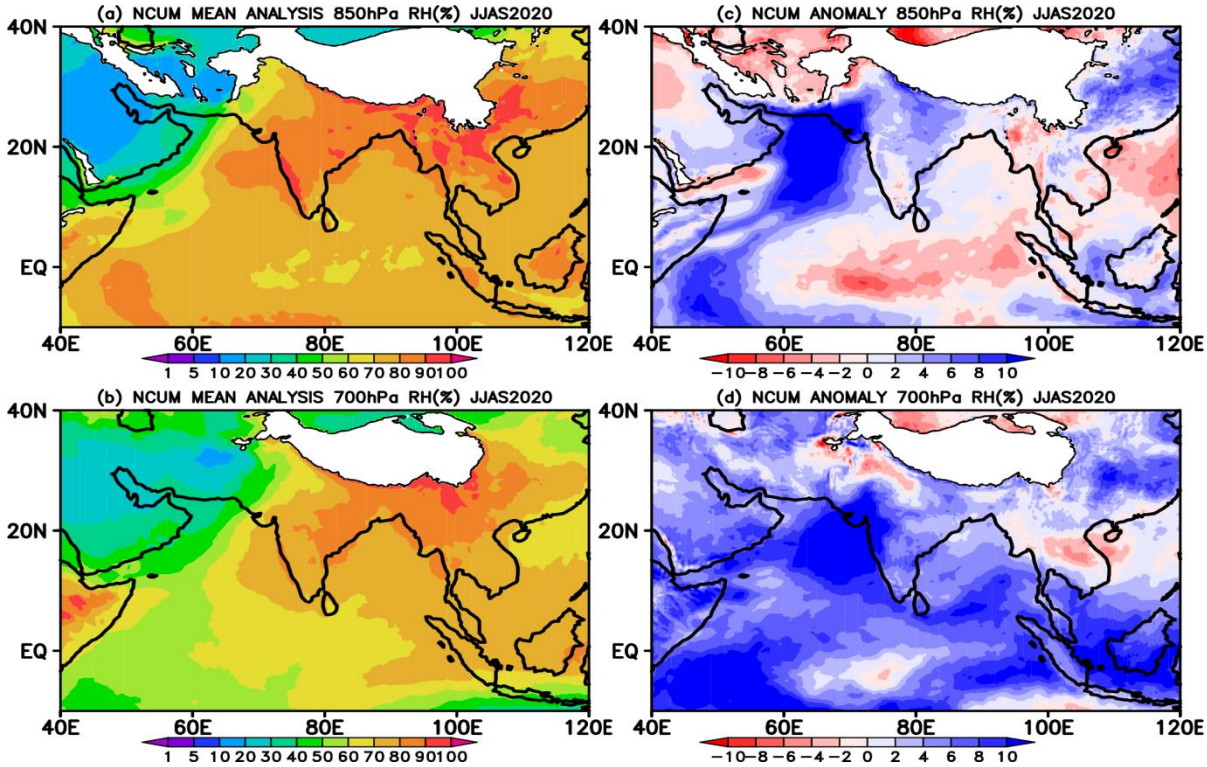

Figure 7. Analysis mean RH at (a) 850 hPa(b) 700 hPa. RH anomaly at (c) 850 hPa (d) 700 hPa levels.

The lower tropospheric mean and anomaly Relative Humidity at 850 and 700 hPa levels is shown in Figure 7a-d. The mean RH is high (>80%) at 850 and 700 hPa (Figure 7a,b) over the Indian land region and neighboring seas typically indicating the humid monsoon climate. The mean RH values exceed 90% at 850 hPa over the west coast of India and parts of Northeast India. The RH anomaly at 850 hPa level (Figure 7c) is prominently positive (blue) indicating a wet anomaly in the analysis over the entire Arabian Sea and parts of central and northern India. However, the RH anomaly is negative (red) indicating a dry anomaly over the Bay of Bengal and large parts of the equatorial Indian Ocean (central and eastern). At 850 hPa level, the dry RH anomaly could be partly explained as due to weaker southwesterly monsoon flow (Figure 3c) over a large area covering the Bay of Bengal and large parts of the equatorial Indian Ocean. However, at 700 hPa level, the RH anomaly is positive (blue) over the entire domain, except for isolated regions of NW India and SE Asia, which may not be explained fully by the wind anomaly shown in Figure 3d.

4. Systematic Errors in NCUM-G Forecasts

This section discusses systematic errors in the Day-1, Day-3, and Day-5 forecasts relative to model analysis for Winds, Temperature, and Relative Humidity for 850, 700, 500, and 200 hPa levels (Figure 8-13).

4.1 Winds at 850, 700, 500, and 200 hPa levels

At 850 hPa level (Figure 8a-d), the systematic errors in the forecast indicate four important features. (a) A westerly & southwesterly bias over the Indian region and neighboring seas suggesting stronger monsoon flow in forecasts. The magnitude of the positive bias is increasing from Day-1 to Day-5 (Figure 8b-d) and extends over a large area covering the Somali coast over the Arabian Sea and parts of SE Asia. (b) An easterly bias over the southern Bay of Bengal indicating a weakened westerly/southwesterly flow in the forecasts. The pattern of biases (positive and negative) over the Bay of Bengal and the anticyclonic anomaly in Figure 8c,d, suggests a slight northwards shift in the monsoon flow over the Bay of Bengal. (c) The westerly bias along the west coast of India can be seen shifting northwestwards in Day-5 with the strongest bias over the northern Arabian Sea, western India in Day-3 and Day-5. (d) The negative bias (red) over isolated parts of the western Arabian Sea suggests weaker monsoon flow in the forecasts at 850 hPa.

At 700 hPa level (Figure 9a-d), the westerly bias over the Indian peninsula and Bay of Bengal is prominent extending over SE Asia. Over the Arabian Sea, the pattern of bias is reversed compared to 850 hPa (Figure8). Negative bias (red) over the northern and western Arabian Sea indicates weakened monsoon flow, particularly in the Day-3 and Day-5 forecasts. The cross-equatorial flow can also be seen weakened of the African coast. Additionally, the negative bias over the southern Bay of Bengal, prominent at 850 hPa is also pronounced at 700 hPa level. Positive bias over northwest India shows an anticyclonic anomaly flow while over eastern IG plains, forecasts show positive bias with cyclonic anomaly flow. At 500 hPa level (Figure 10a-d) the pattern and sign of biases over the western and eastern equatorial Indian Ocean are more or less similar. Over the northern Arabian Sea, positive bias (blue) indicates strong northerly flow from NW India where anomalous anticyclonic flow is evident in Day-3 and Day-5. At 200 hPa (Figure 11a-d), The positive easterly bias (blue) extends over the Indian peninsula and neighboring Arabian Sea and Bay of Bengal. The negative westerly bias (red) occupies the entire equatorial region of the domain suggesting weakened tropical easterlies.

4.2Temperatures& Relative Humidity

At 850 hPa level, (Figure 12a-d) systematic errors indicate (a) a warm bias (>1°C) over the Indian land region and the neighboring Arabian Sea, Bay of Bengal, and the equatorial Indian Ocean. (b) Strong warm bias (>2°C) is prominent in Day-3 and Day-5 forecasts over NW India. (c) Mild cold bias (about -0.5°C) over western India along the coast.

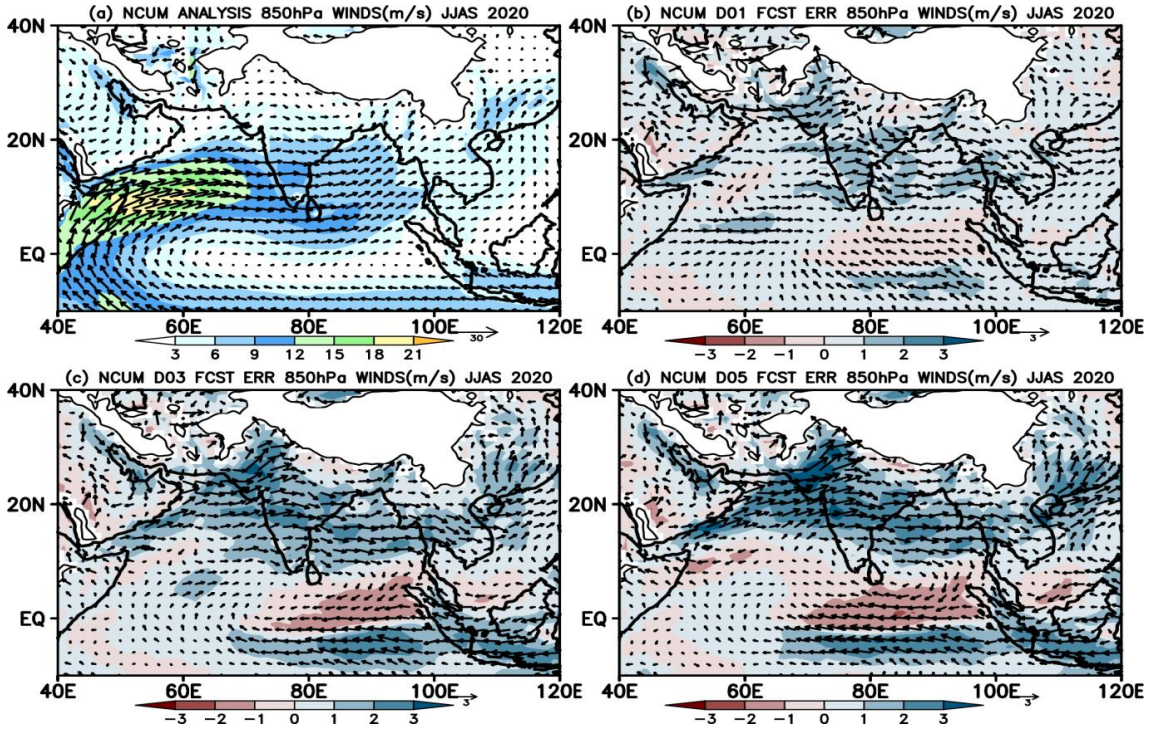

Figure 8. Mean 850 hPa analysis winds and systematic errors in the (b)Day-1 (b) Day-3 and (d) Day-5 forecasts.

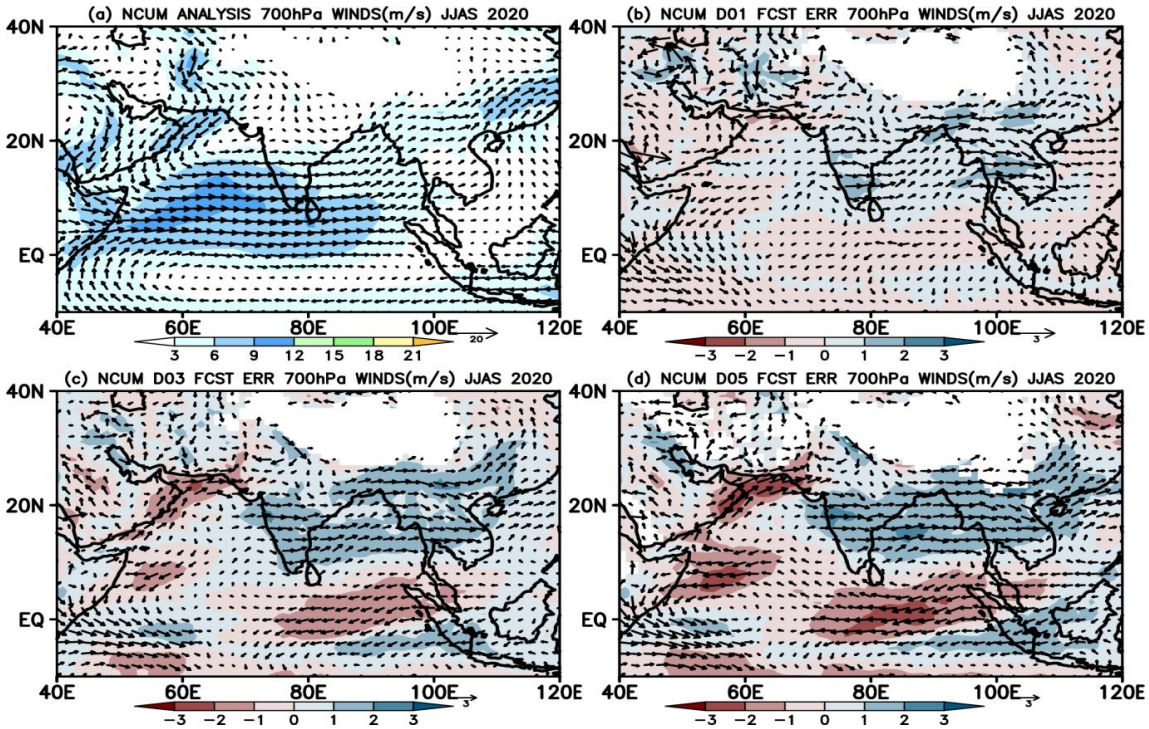

Figure 9. Mean 200 hPa analysis winds and systematic errors in the (b) Day-1 (b) Day-3 and (d) Day-5 forecasts.

At 700 and 500 hPa (see appendix) the warm (*cold*) bias over the land (*sea*) regions persists by and large with reduced magnitude. At 200 hPa level, the land (*sea*) regions partly feature cold (*warm*) bias in all the forecasts with magnitudes of about 0.5°C.

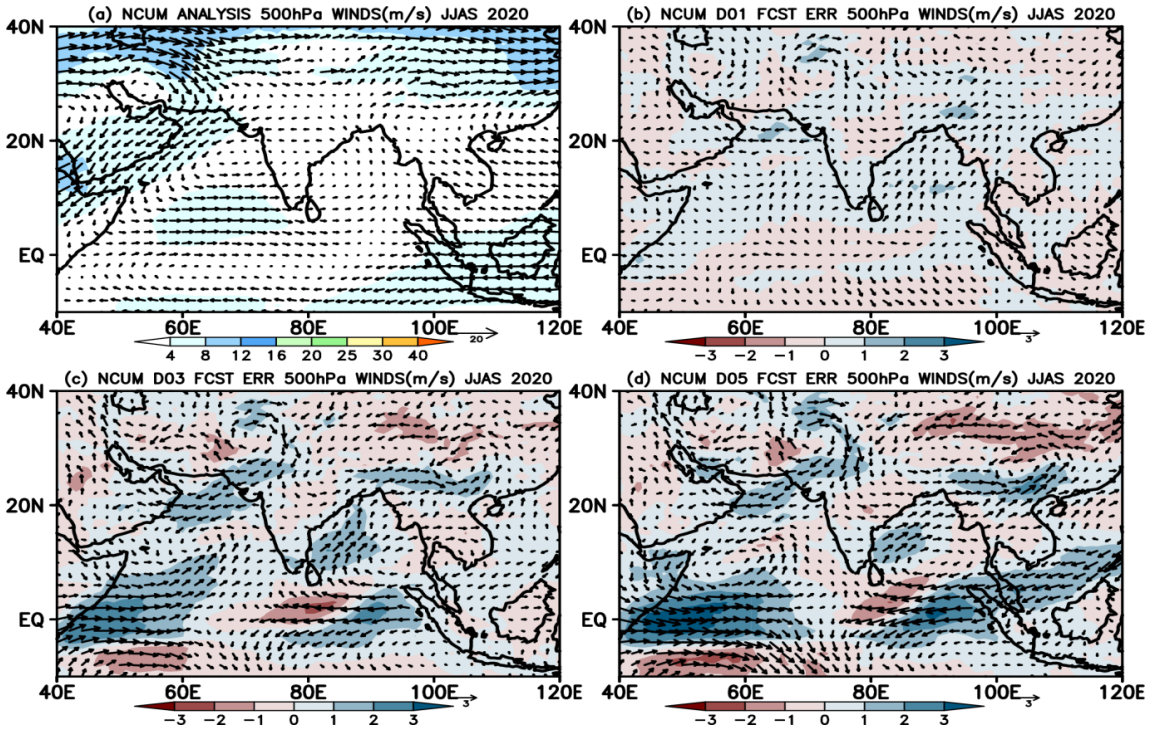

Figure 10. Mean 850 hPa analysis winds and systematic errors in the (b) Day-1 (b) Day-3 and (d) Day-5 forecasts.

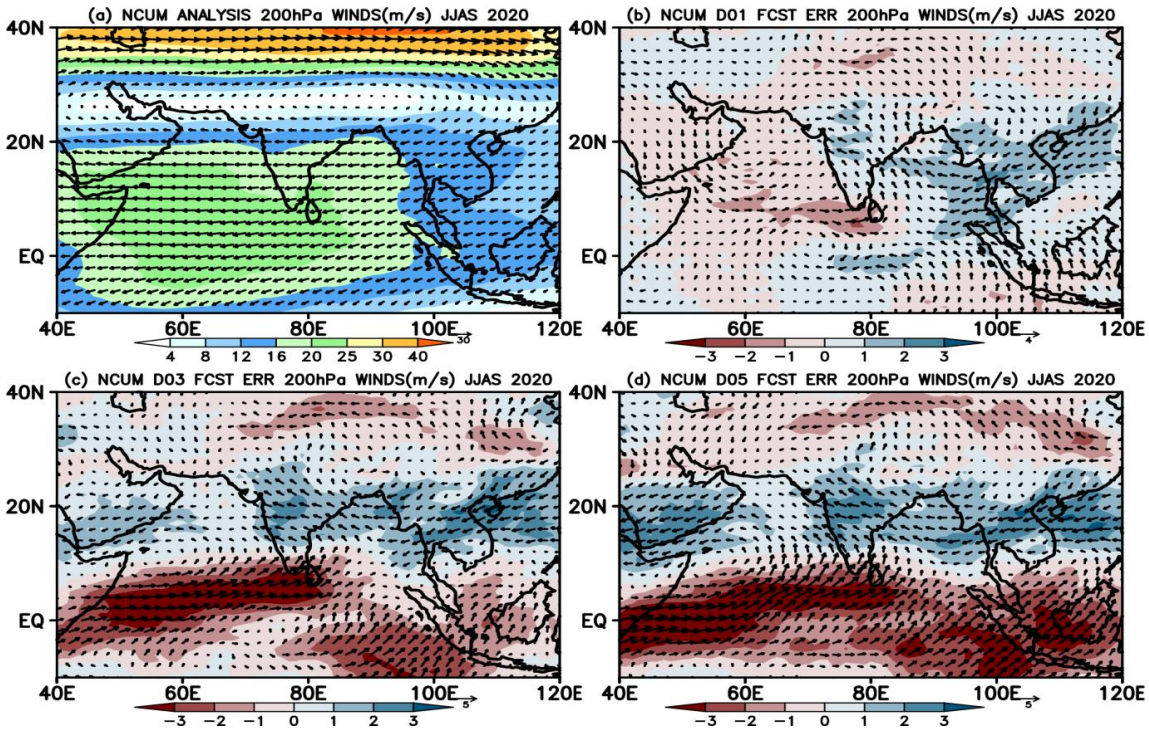

Figure 11. Mean 850 hPa analysis winds and systematic errors in the (b) Day-1 (b) Day-3 and (d) Day-5 forecasts.

Figure 13 shows the systematic errors in RH at the 850 hPa level. The forecasts prominently show negative (red) dry bias over the Arabian Sea, NW India, and IG plains. It is rather difficult to corroborate all the dry biases based on the wind biases at 850 hPadiscussed in Figure 3. Similarly, the 700, 500, and 200 hPa level RH systematic errors (see Appendix) also show strong dry bias over land regions and mild wet bias over Sea.

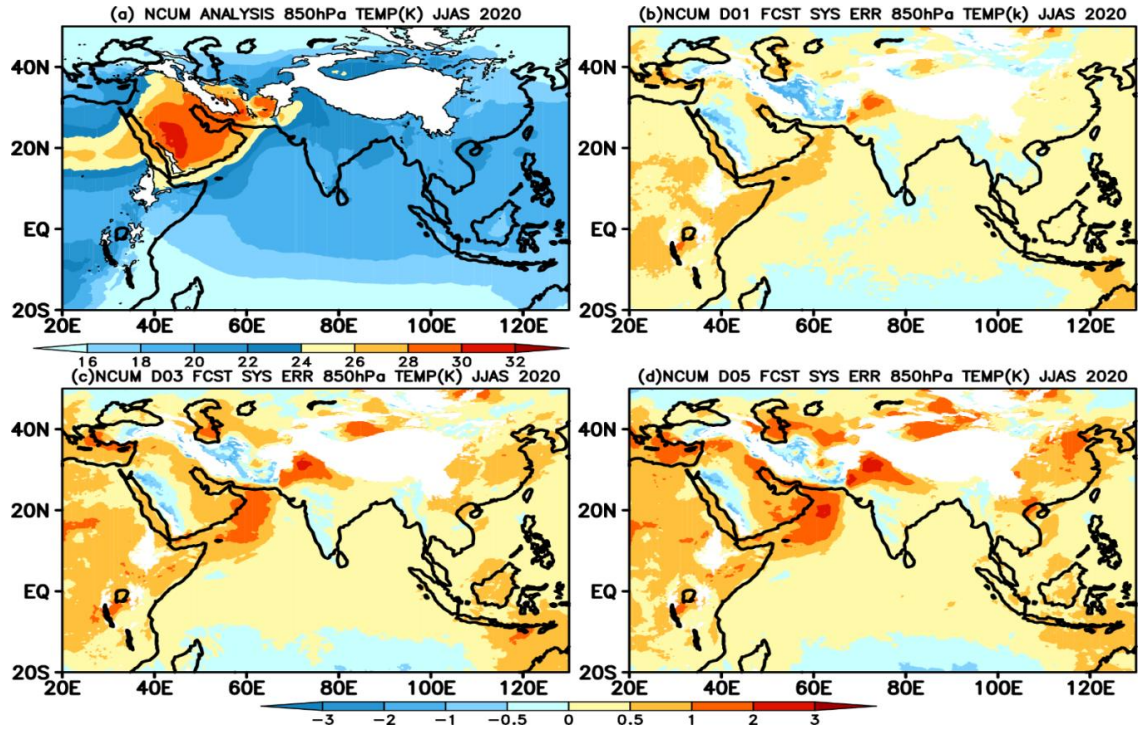

Figure 12. Mean 850 hPa analysis temperature and systematic errors in the (b) Day-1 (b) Day-3 and (d) Day-5 forecasts.

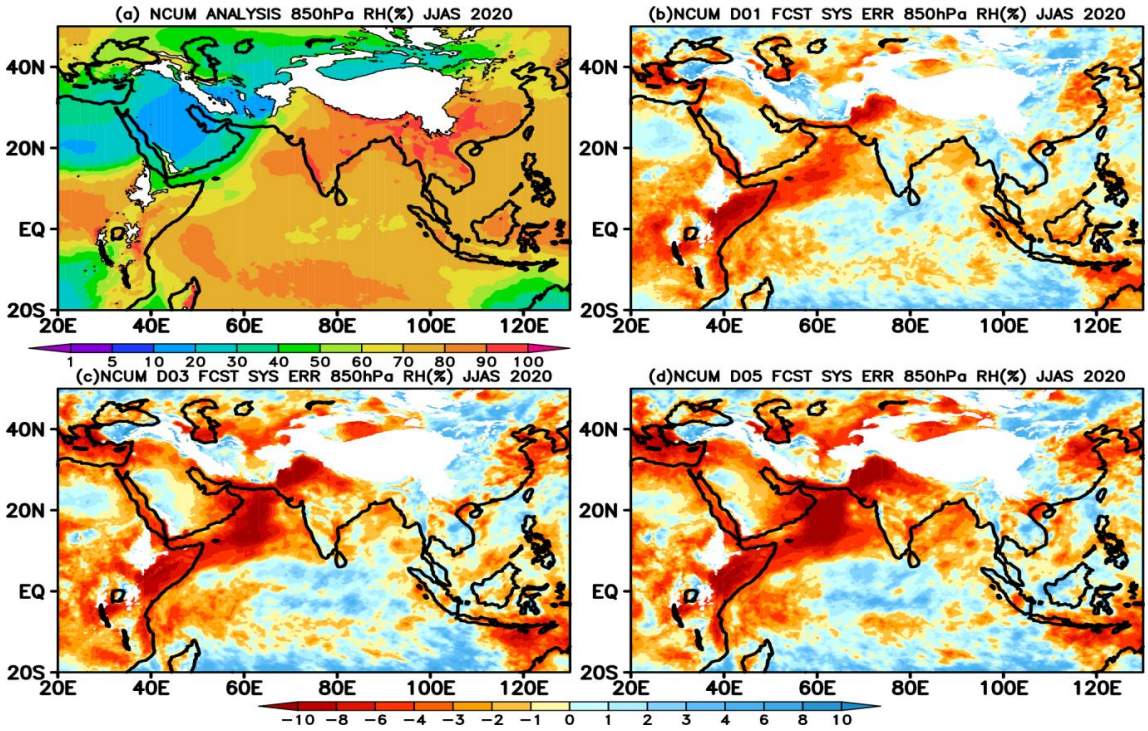

Figure 13. Mean 850 hPa analysis RH and systematic errors in the (b) Day-1 (b) Day-3 and (d) Day-5 forecasts.

4.3Vertically Integrated Moisture Transport (VIMT)

To further assess the role of biases in winds and moisture (RH) at different levels, systematic errors in VIMT are computed based on wind components (*U, V*) and specific humidity (*q*) at all levels (surface – 300hPa). The analyses mean VIMT and systematic errors are presented in Figure 14a-d. The biases in the forecasts show positive (blue) over the entire Indian land region, Arabian Sea

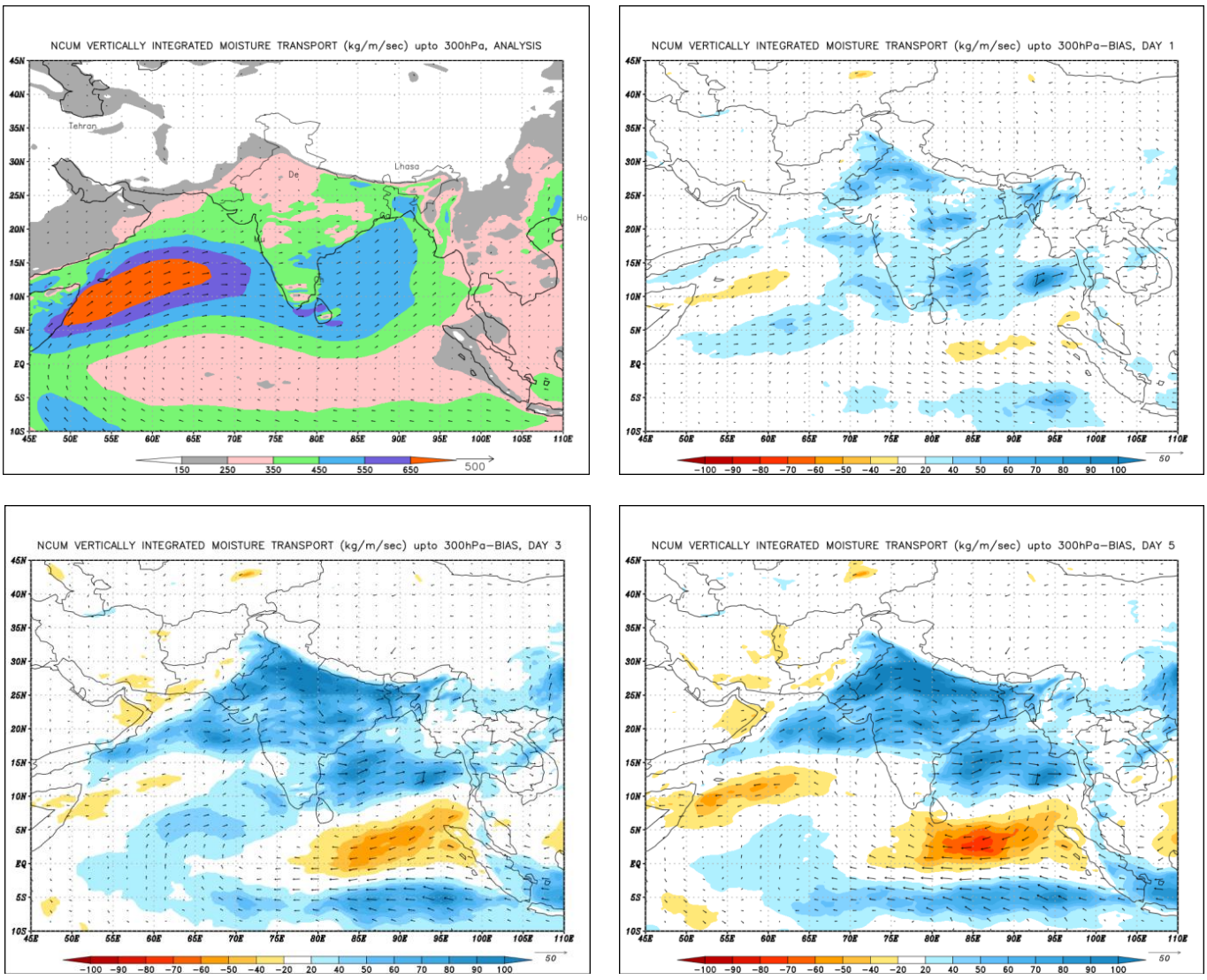

Figure 14 Mean Analysis VIMT and systematic errors in the Day-1, Day-3, and Day-5 NCUM forecasts

and the Bay of Bengal. The Day-3 and Day-5 forecasts show negative bias (red) over the southern Bay of Bengal consistent with the wind biases at 850, 700, and 500 hPa shown in Figures 3 and 4. Despite the strong dry bias indicated by systematic errors in RH (Figure 13), it can be inferred that the strong positive biases in winds have offset the impact of dry bias in the forecasts over India.

4.4RMSE in Wind and Temperature over India 2018-2020

The forecast errors in 850 and 200 hPa winds and temperature are quantified in terms of RMSE computed against the analysis over India in Figure 15. The results of JJAS 2020 are compared with scores from recent monsoons (2018-2020). Although there are some year-to-year variations in the RMSE, the lowest RMSE is evident during JJAS 2020 at all lead times (except for 200 hPa Temp, Figure 15(d)).

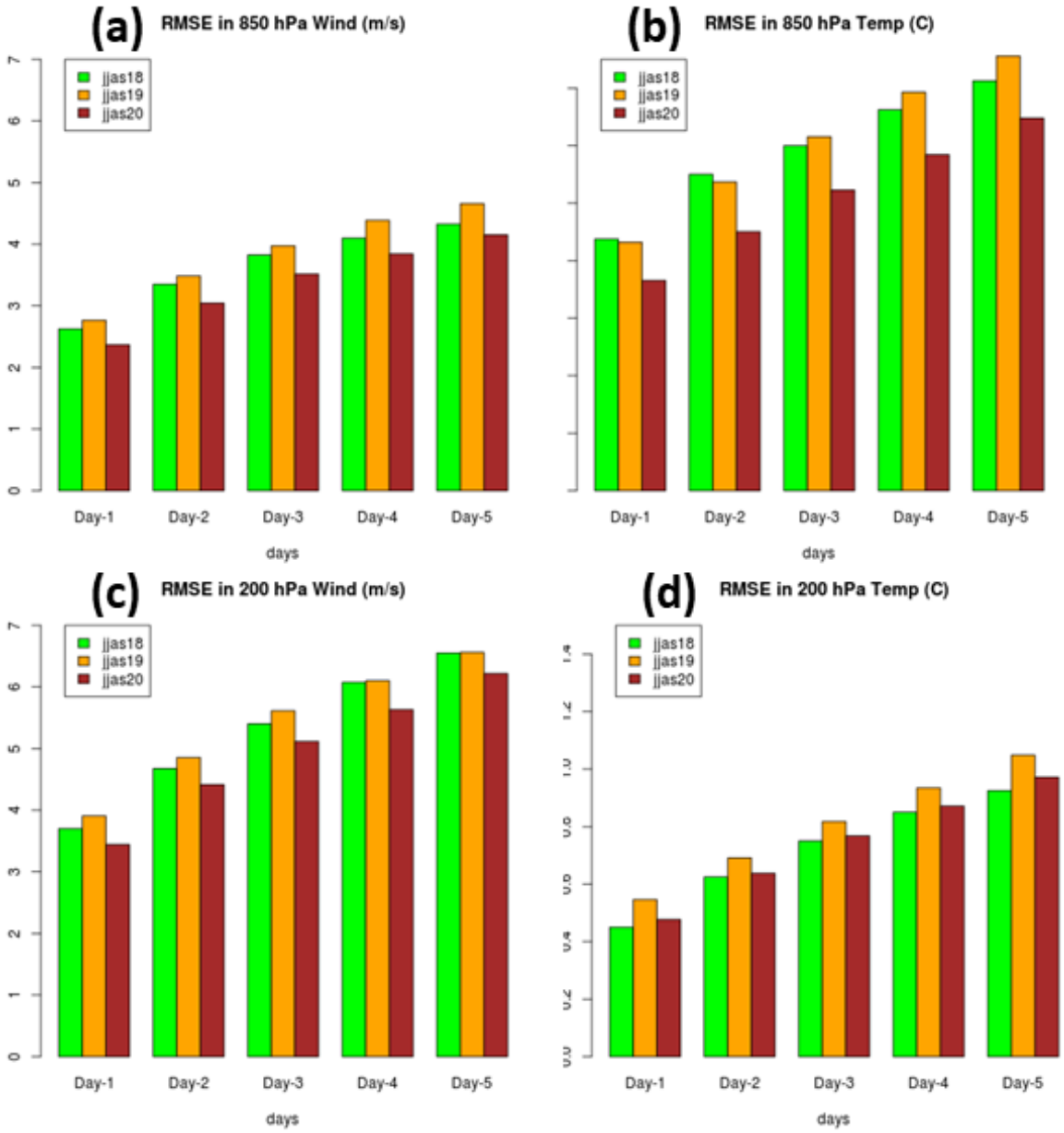

Figure 15. RMSE in the NCUM forecast 850 & 200 hPa wind and temperature during recent years.

5. Rainfall Forecast Verification

Verification of NCUM-G model rainfall forecasts is presented in this section for JJAS 2020. The 24 hour accumulated rainfall forecasts are verified against the NCMRWF-IMD merged Satellite and gauge rainfall analysis. The discussion presented in this section is aimed at highlighting the model biases and accuracy over India using mean error (ME), RMSE, correlation, forecast rainfall frequency at different thresholds, etc. Further, this section also quantifies forecast skill using standard verification metrics, namely, POD, FAR, CSI, BIAS, PSS which are described in standard text books (Wilks, 2011, Jolliffe and Stephenson, 2012), and new SEDI (Stephenson et al 2008, Ashrit et al 2015, Sharma et al 2021).

5.1 Mean, Mean Error, RMSE, and Correlation

The observed and forecast mean monsoon rainfall during JJAS 2020 is shown in Figure 16. Observations indicate the highest mean rainfall exceeding 20mm/day over the west coast and parts of eastern India over Meghalaya and adjoining Bangladesh. Another region of high rainfall >10mm/day is prominent over the core monsoon region where the monsoon depressions bring in heavy rains. The panels in the middle row, Figure 16b-d show Day-1, Day-3, and Day-5 forecast rainfall averaged during the same period. The observed peak rainfall amounts (>15mm/day) along the Western Ghats and the Arrakkan coast are predicted. However, it is found that the NCUM forecast shows higher rainfall amount all over the west coast, NE India, and over the Himalayas. The forecasts overestimate the isolated high rainfall amounts (>10mm/day) over the core monsoon region indicating the overestimation of observed rainfall over land and the neighboring seas. The reduced rainfall amounts (<6mm/day) over the eastern parts of the peninsula and northwest India are predicted fairly well in the model. The panels in the bottom row show mean error (ME) in predicted rainfall indicating wet bias (blue) all along the west coast and over central India. Dry bias is prominent over the Arabian Sea and Bay of Bengal. Over the Indian land region, dry bias can be seen over western Ghats, Gujarat, western India, eastern peninsula parts of West Bengal, and adjoining Bangladesh.

The excessive forecast rainfall in the model over hilly regions of the west coast, NE Indian, and Himalayas is attributed to an enhanced model grid resolution of 12km. The wet bias (blue) over central India and the core monsoon zone can be understood as resulting from the positive bias in VIMT discussed in Figure 14. While the dry bias over the southern Bay of Bengal could be associated with negative bias in VIMT over that region, the dry bias over the Arabian Sea can’t be explained in terms of systematic errors in VIMT.

To further quantify and summarize the rainfall forecast performance over India, a spatial map of RMSE and correlation is presented in Figure 17. RMSE gives the average forecast error weighted according to squared error. While it does not indicate the direction of the forecast errors, it gives greater emphasis on relatively larger errors. It can be noted that the forecasts feature large errors over regions of high rainfall amounts along the west

coast, Himalayan belt, and NE India. Additionally, RMSE is high over parts of the core monsoon region, which can be seen to increase with lead time. However, the map of correlation (Figure 17d-f) shown in the bottom panels suggests a sharp decline in the forecast skill.

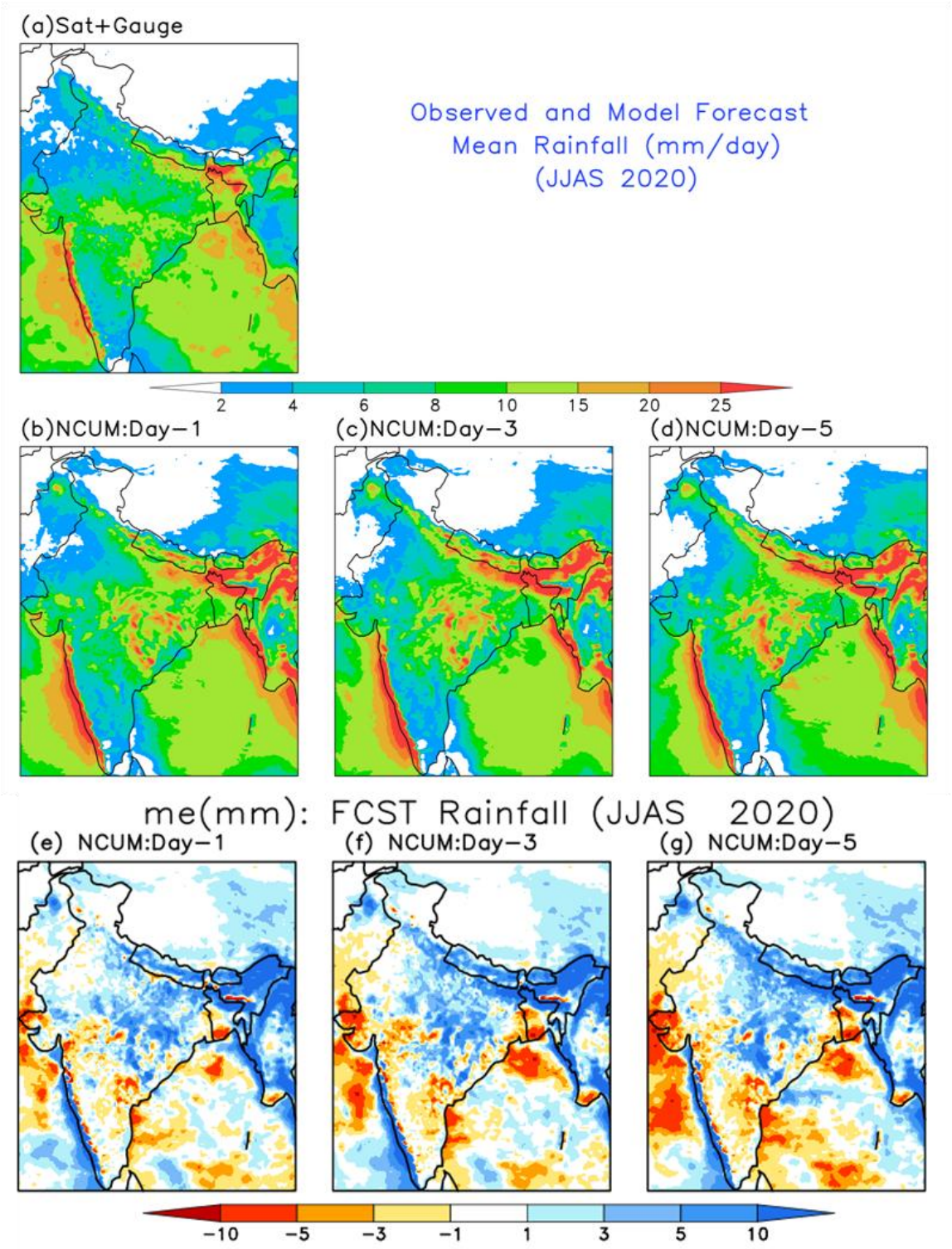

Figure 16. Observed model forecast rainfall (mm) averaged during Jun-Sept 2020. Panel (a) IMD-NCMRWF shows observed rainfall analysis. Panels (b), (c), and (d) correspond to Day-1, Day-3 and Day-5 forecast mean rainfall based on NCUM.

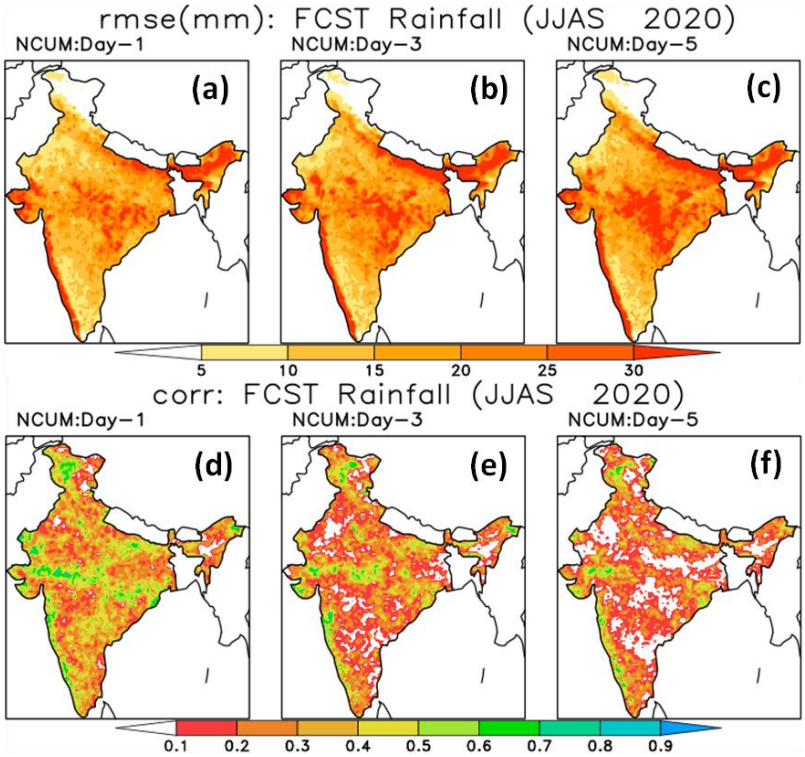

Figure 17. RMSE and correlation in the Day-1, Day-3, and Day-5 forecasts.

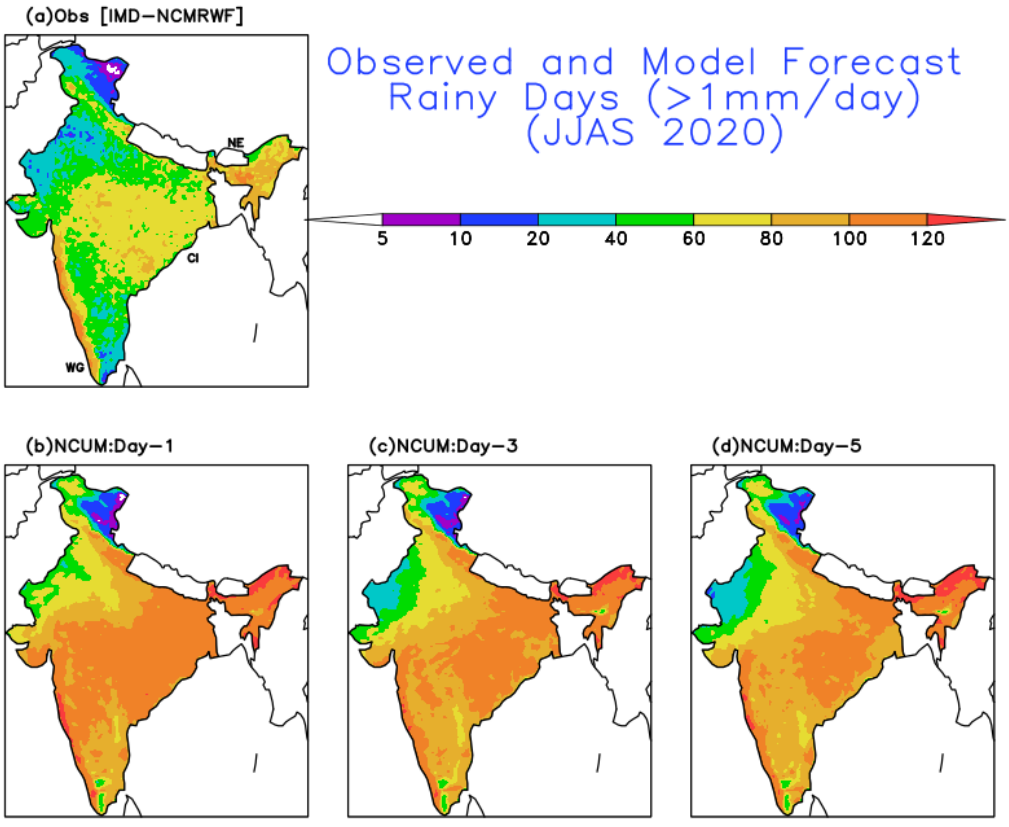

Figure 18. Observed model forecast number of rainy days during Jun-Sept 2020. Panel (a) Observed (IMD-NCMRWF) rainfall analysis. Panels (b), (c), and (d) correspond to Day-1, Day-3, and Day-5 forecasts based on NCUM

5.2 Frequency of Rainy Day, Moderate and Heavy Rain Days

To further assess the source of wet rainfall bias over the land regions in the model forecasts, observed and forecast counts of rainfall >1mm/day (rainy day) are presented in Figure 18. The observations show (a) a high number of rainy days (>80 days; orange shade) along the west coast and NE India and (b) 60-80 days (yellow shade) over the core monsoon region. Over the dry regions of NW India and the eastern peninsula, the number of rainy days is lower than 40-60 days (green). The model forecasts over most parts of India show an exceedingly high number of rainy days.

Similar counts for observed and forecast rainfall exceeding the "moderate" (>15.6 mm/day) and "heavy" (>64.5mm/day) rainfall thresholds are presented in Figures 19 & 20. Figure 19b-d shows that the model forecasts have a higher (60-80) number of moderate rainfall over the west coast and NE India. Over the dry regions of NW India, central and eastern peninsula, the model predicts a far lower number of moderate rain days. For the heavy rain category (Figure 20a-d), forecasts have a realistic number of counts along the west coast and over the core monsoon region (to some extent). The counts are overestimated over NE India and the foothills of the Himalayas. Thus it can be concluded that a significant part of rainfall biases over the Indian land region is purely statistical, due to the model's tendency to frequently predict higher rainfall amounts.

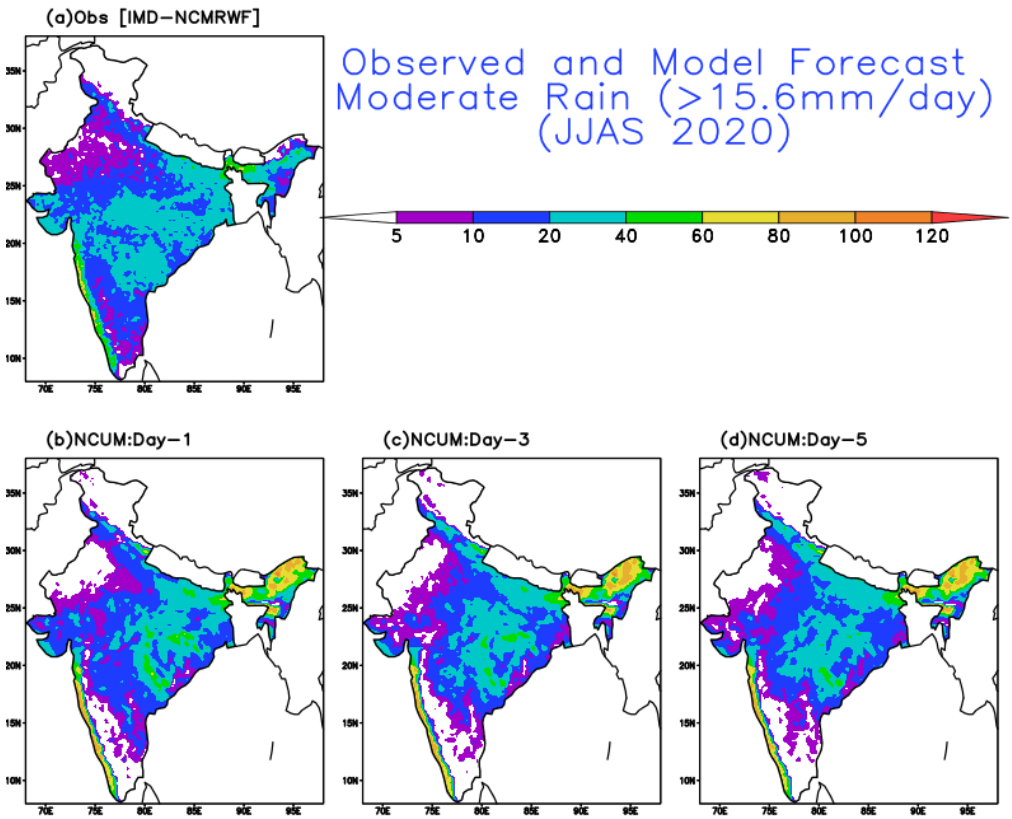

Figure 19. As in Figure 18 for rainfall threshold of 15.6 mm/day

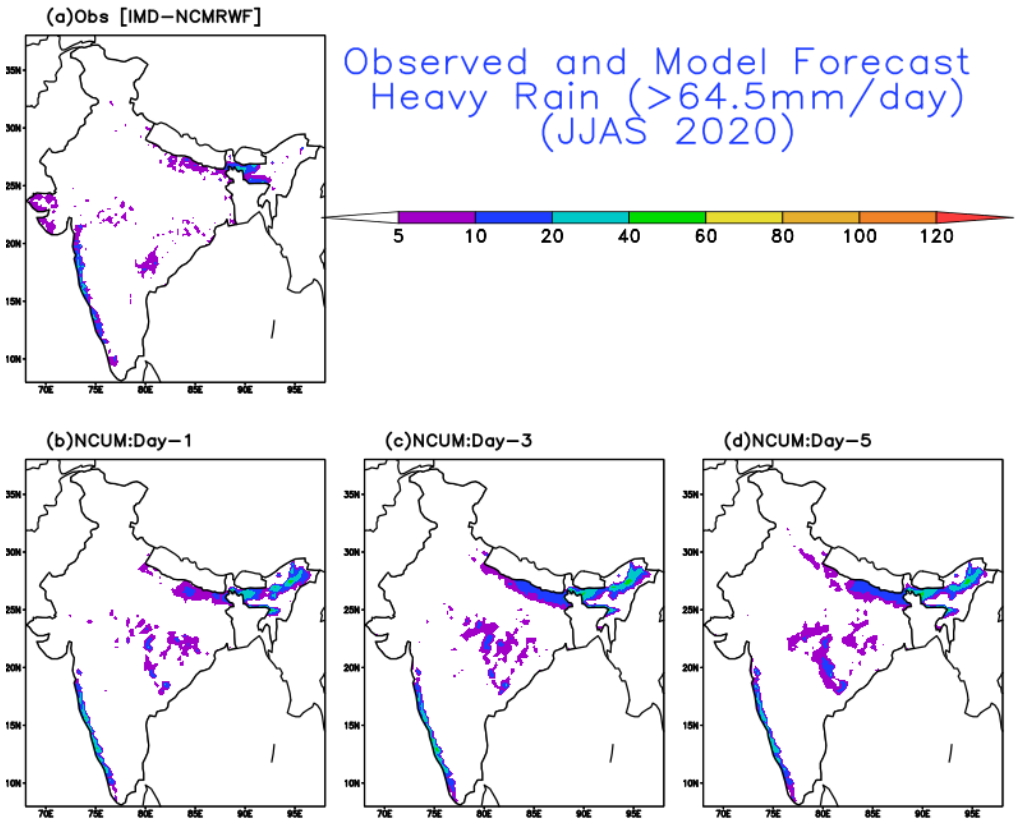

Figure 20. As in Figure 18 for rainfall threshold of 64.5 mm/day

5.3 Categorical Verification of Rainfall forecasts

The categorical approach of verifying QPF is generally based on the 2 x 2 contingency table which is evaluated for each threshold. A detailed description of the verification metrics is presented in Appendix and the results for JJAS 2020 are presented in Figure 21. Verification scores are presented for rainfall of up to 30mm/day. For rainfall threshold of up to 9mm/day, the forecasts have POD >0.5 (& FAR <0.6) in the Day-1 forecast. The BIAS score (frequency bias) indicates that forecasts overestimate the frequency of 9mm/day or lower. The values of CSI, PSS, and SEDI all are high for rainfall up to 9mm/day suggesting reasonable skill. However, the skill is not bias-free. For higher amounts (9-30mm/day), there is no frequency bias, but the skill is low as indicated by CSI, PSS, and SEDI.

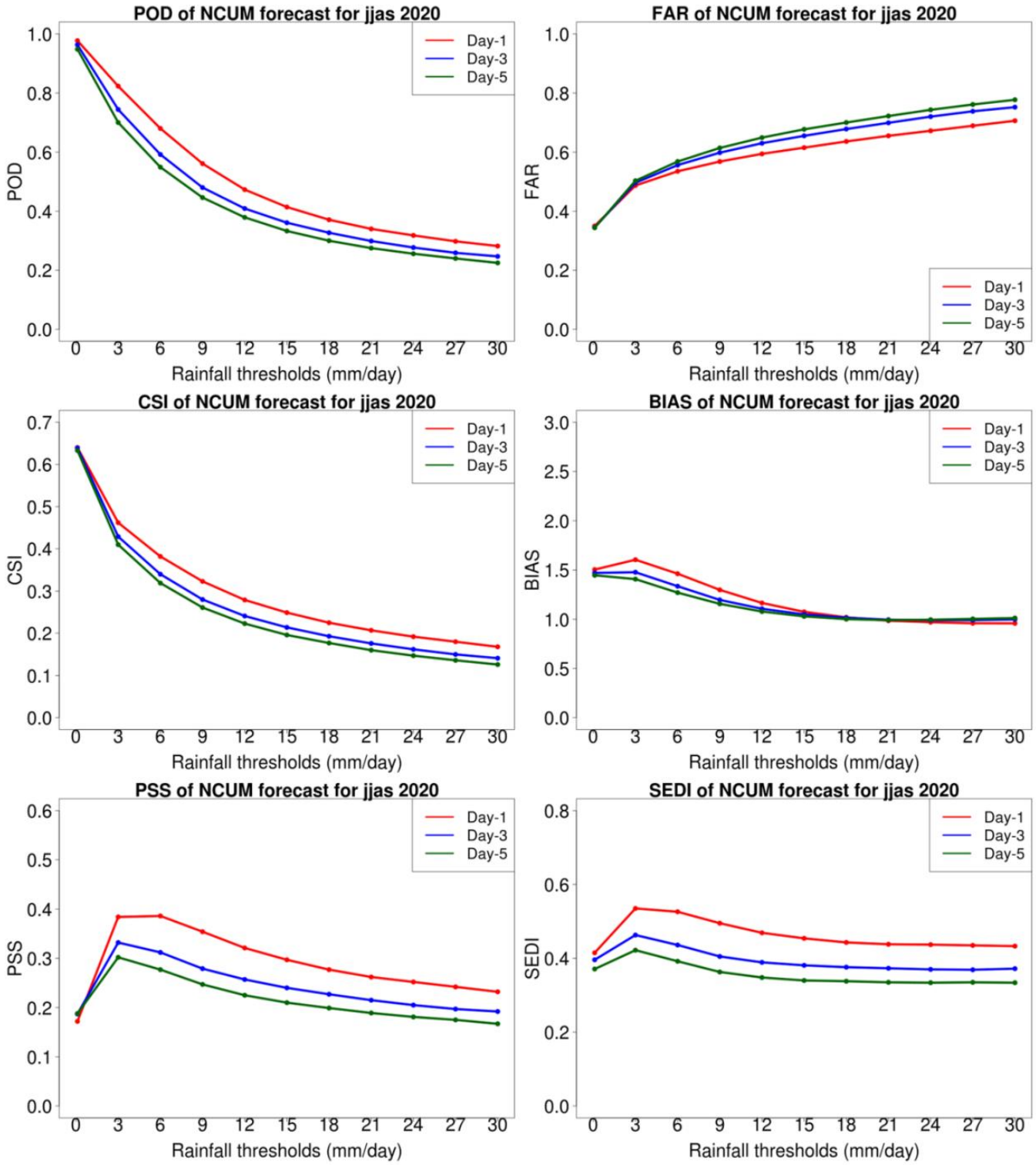

Figure 21. Categorical Rainfall scores POD (*top left*), FAR (*top right*), CSI(*middle left*), BIAS (*middle right*), PSS (*bottom left*) and SEDI (*bottom right*).

5.4Verification of Probabilistic Rainfall Forecasts

The verification of probabilistic forecast of rainfall obtained from NEPS over the Indian land region was carried out for rainfall exceeding 1, 2.5, 15.6, and 65.5 mm/d thresholds for June to September 2020. Standard metrics for the verification of probabilistic forecasts like Brier Score (BS), Reliability, and Resolution as components of BS, and Area under ROC (AROC) were calculated for rainfall exceeding all the above thresholds. Detailed description of the metrics can be found in standard text books (Wilks, 2011, Jolliffe and Stephenson, 2012) and example applications to NCMRWF model forecasts in Dube et al., (2016a,b) . BS is the difference between the forecast probabilities and the observed frequencies and therefore the perfect value for this score is '0'. Panels in Figure 22 (a-d) show the BS, reliability, resolution, and BSS for the rainfall forecasts from NEPS. It is seen from Figure 22(a) that the BS for higher rainfall thresholds is lower implying a better match between the forecast probability and observed frequency for heavier rainfall. For lower rainfall thresholds it is seen that the BS increases slightly with an increase in forecast lead time which is indicative of poor model skill. This could be attributed to excessive number of rainy days seen in the NCUM-G (Figure 18). Figure 22(b) shows the reliability as a component of BS and for a perfect forecast, a lower value of reliability is desirable. As in the case of BS, it is seen that forecasts for higher rainfall thresholds show lower reliability values as compared to smaller amounts of rainfall. Resolution on the other hand should have a higher value for a better forecasting system and it is seen from Figure 22(c) that the resolution values for lower rainfall thresholds are higher and the resolution values

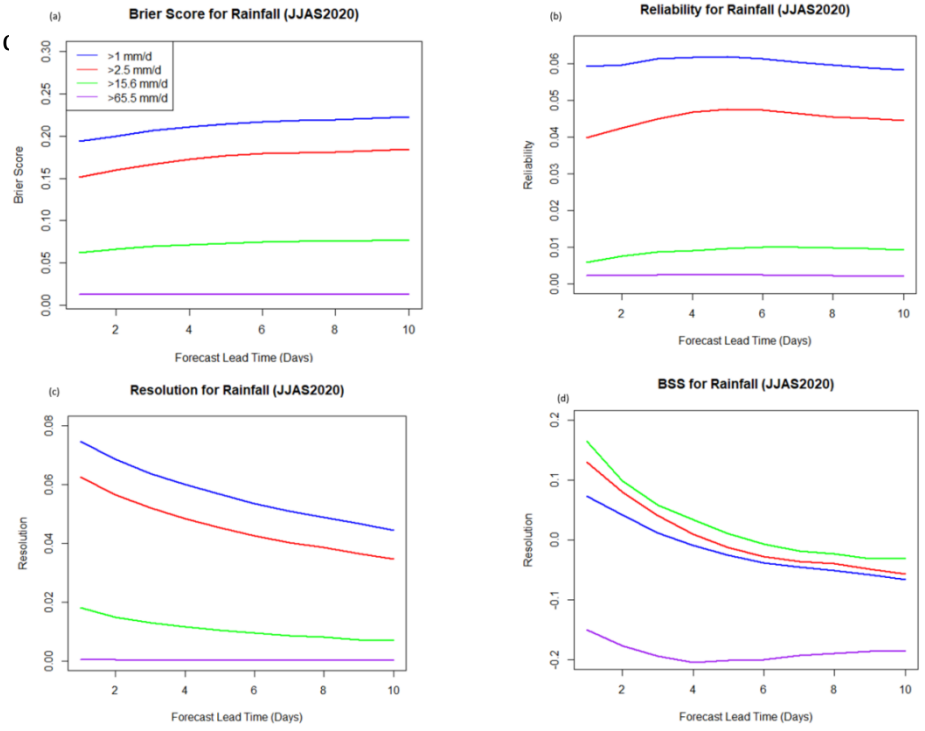

Figure 22 : (a) BS, (b) reliability, (c) resolution and (d) BSS for JJAS 2020 rainfall forecast from NEPS

The area under the ROC (AROC) curve is also a verification metric that can be used for analyzing the skill of the model. For a perfect forecasting system, the value of the AROC should be closer to 1 and an area of 0.5 indicates no skill in the model (or the ROC curve is aligned along the diagonal line). Figure 4 shows the ROC curve for the above rainfall thresholds. It is seen from the figure that the higher rainfall thresholds (65.5 mm/d) show the lowest AROC values followed by 15.6 mm/d & 2.5 mm/d threshold shows the higher AROC among all the rainfall thresholds. Also, the AROC values for all thresholds decrease with increasing lead times, which is also seen from Figure 23. (ROC curves shift closer to the diagonal lines with increasing lead time).

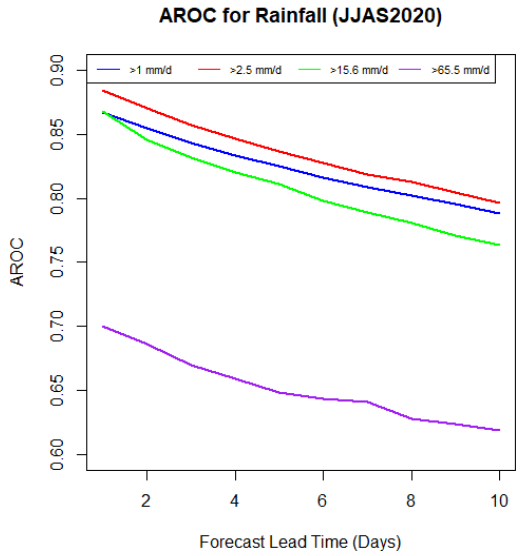

Figure 23: AROC for JJAS 2020 rainfall forecast from NEPS

5.5 Rainfall forecast skill during 2018-2020

Forecast skill for higher rainfall amounts is rather poor as noted in an earlier section. Additionally, for higher rainfall amounts, it is a challenge to derive any meaningful conclusion when verification is carried out based on smaller sample size. It would be still useful to monitor relative changes/improvements in the forecast skill of the model for higher rainfall thresholds. Categorical scores for rainfall > 2 and 5 cm/day in the recent five years are shown in Figure 24. During 2018-2019, the forecasts consistently show improved skill as evidenced in increasing POD, PSS SEDI, and decreasing FAR. During JJAS 2020 the metrics indicate a marginal dip in the skill which could be due to year-to-year variations in the number of events.

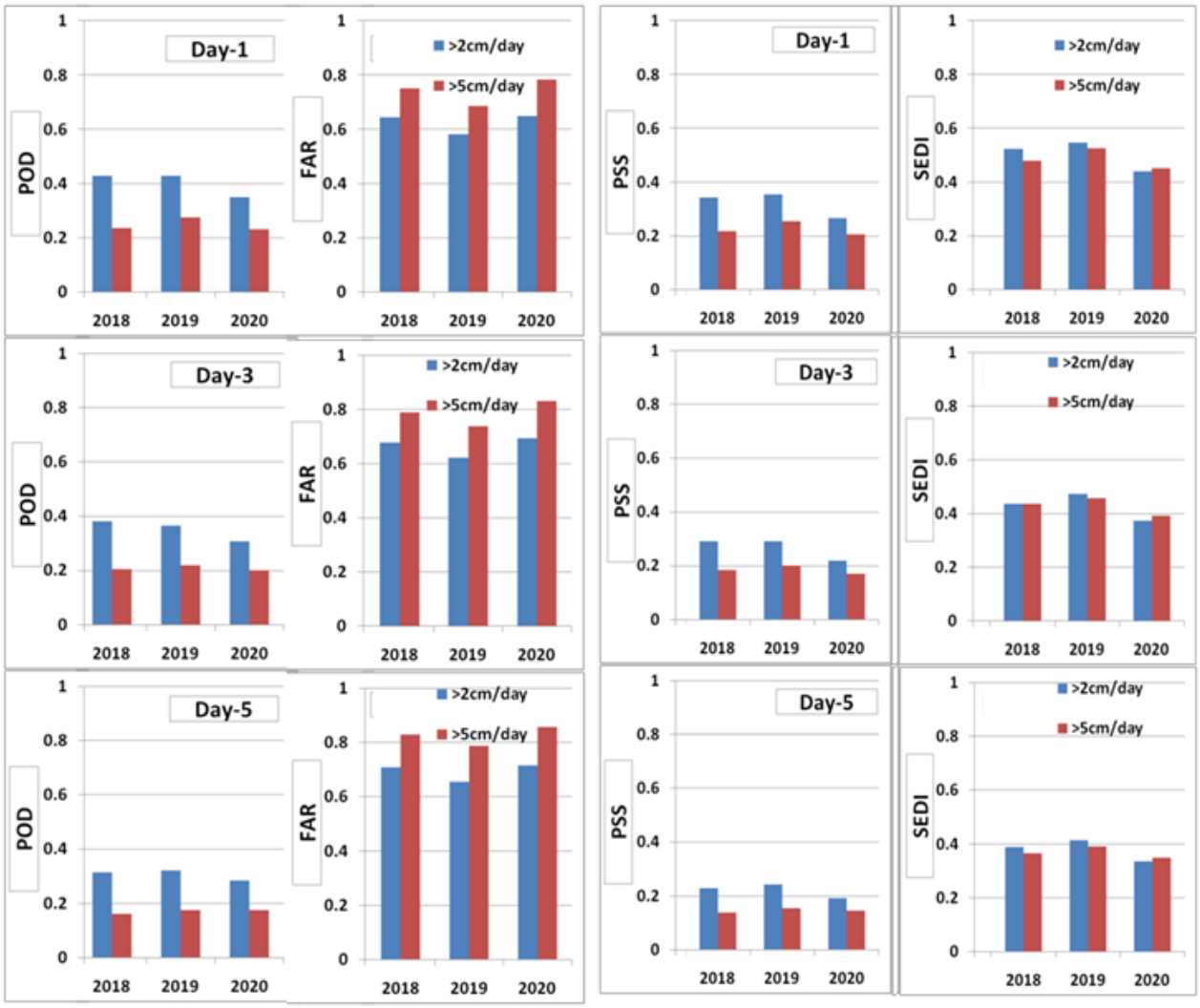

Figure 24. Categorical verification of rainfall forecasts for rainfall > 2cm/day and 5cm/day during five recent monsoons .

6. Verification of Onset, Active/Break Spells, and Synoptic Features

6.1 Onset of Monsoon: Onset Circulation Index (OCI)

In this section, we identify the onset date based on the NCUM Global model forecasts. The objective identification of monsoon onset is not an easy task as it involves many meteorological variables. For instance, the IMD defines the onset based on several variables that include stations rainfall in Kerala state, wind field, and outgoing longwave radiation, etc. (https://mausam.imd.gov.in/imd_latest/contents/onset.php). However, there are several studies in the past tried to identify simple objective definitions by correlating with IMD monsoon onset criteria. For instance, Wang et al., (2001) identified a circulation index by taking the difference between 850hPa wind fields averaged over 5^{o}–15^{o}N, 40^{o}–80^{o}E, and 20^{o} –30^{o}N, 70^{o} –90^{o}E. This index describes the low-level shear vorticity over the monsoon trough which is well correlated (0.79) with the IMD onset date along with the rainfall variability of the Indian monsoon (0.74). Later, a further simple and robust circulation index is estimated using the 850 hPa zonal wind averaged over the Southern Arabian Sea (SAS) region i.e., 5^{o}–15^{o}N, 40^{o}–80^{o}E (Wang et al., 2009). The index here is defined as the *onset circulation index (OCI)* which has a correlation of 0.81 with the IMD monsoon onset date over the 60-yr period during 1947-2007. Thus, this OCI index is used here for identifying the monsoon onset over Kerala for the year 2020. Further, the onset date is taken as the first day when OCI exceeds 6.2 ms^{-1} and persists for about 7 consecutive days. This threshold is chosen as the climatological mean onset date at Kerala is 1 June where the 850-hPa zonal wind averaged over SAS is 6.2ms^{-1} on 1 June (Wang et al., 2009). Moreover, the 7 consecutive days are taken ensuring that the strong westerlies are not induced by the synoptic event.

Here, we have taken the NCUM-G model forecasts for Day-1, Day-3, and Day-5 along with NCUM-G analysis and ERA5 reanalysis for a comparison. Figure 25 below indicates the temporal evolution of OCI time series from NCUM-G analysis (NCUM-ANA) with forecasts and ERA5 reanalysis. The IMD onset date for the year 2020 is on 1 June, which is indicated by an upward arrow in Figure 25. The temporal evolution of NCUM-G derived OCI is closely matched with the ERA5 reanalysis and NCUM-ANA analysis. While the onset date based on the above methodology from ERA5 and NCUM-ANA is on 31st May 2020 and 28th May 2020, respectively, the NCUM model forecasts indicate the onset day of 28th May 2020 for all the lead days. Hence, the reanalysis-based onset date indicates one day ahead while 3 days ahead in NCUM-G model analysis and forecasts than the actual onset date defined by the IMD.

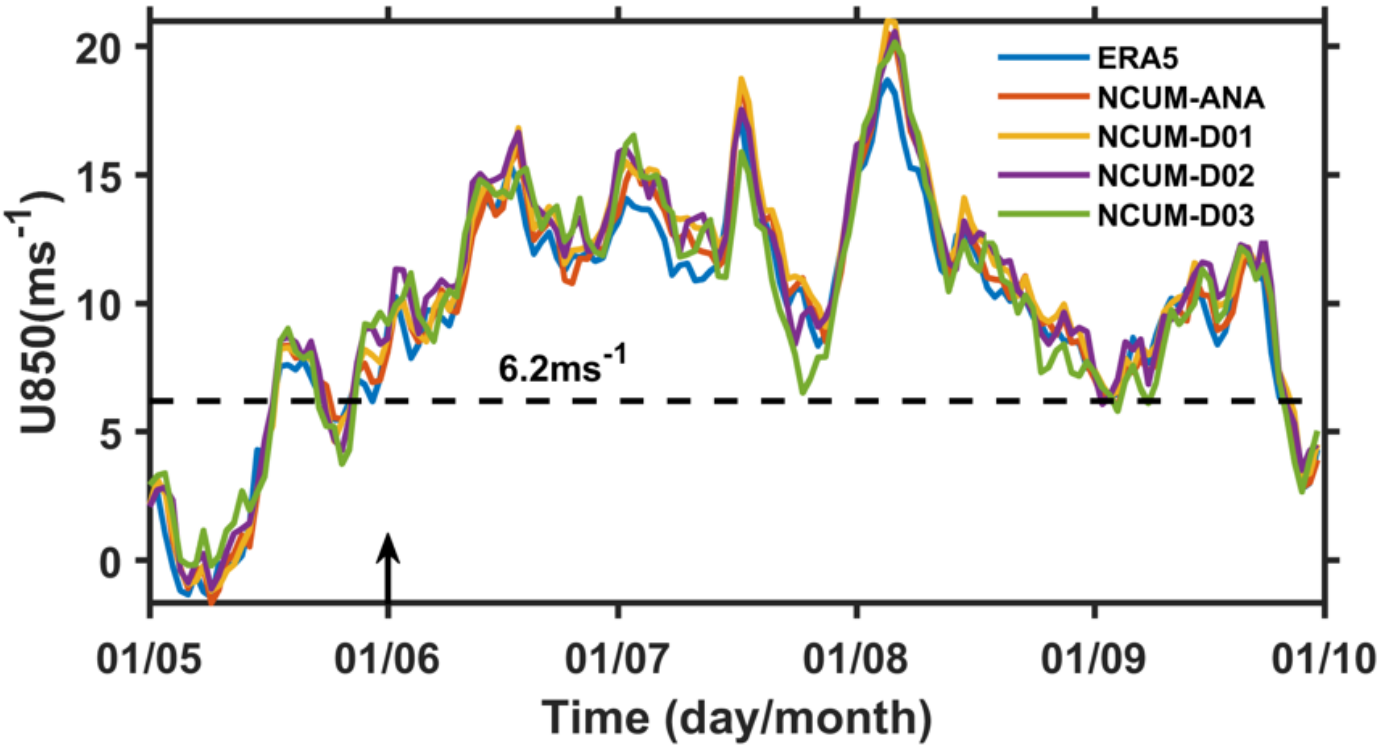

Figure 25: Time series of OCI from 01 May 2020 to 30 September 2020 estimated from the ERA5 Reanalysis along with NCUM-G model analysis and forecasts for Day-1, Day-3, Day-5 lead days.

6.2 Monsoon active and break spells – 2020:

During monsoon season a major component of variability exists from the intraseasonal oscillation of active (enhanced convection) and break (subdued or no convection) spells over the Indian subcontinent. Active and break spells over the Indian subcontinent during monsoon season (June through September, JJAS) 2020 were identified from NCMRWF and IMD merged gridded daily rainfall data over central India (15-25N, 72-85E; CI) following the methodology given by Rajeevan et al (2010). Based on the criterion, 5 active spells and 1 break event were identified. The occurrence time and the duration of the active and break spells are listed in Table 2. The spells are short and occurred mostly in the second half (i.e., during August and September) of the monsoon 2020. Among the active spells, the active spell during the first week of June is due to the *Nisarga* cyclone that passed over the Indian subcontinent.

Table 2: List of Active and break spells over Central India (CI) during JJAS 2020

SNo	Active	Break
1.	1-5 June 2020	26-28 July 2020
2	13-18 August 2020	
3.	20-22 August 2020	
4.	27-30 August 2020	
5.	21-23 September 2020	

6.3 Active and Break spells in NCUM-G model forecasts:

In this section, we have assessed the NCUM-G model rainfall forecast in the representation of active and break spells during the monsoon 2020. For this analysis, we have considered daily rainfall data from the observations averaged over CI model forecasts. The temporal evolution of rainfall time series over the core monsoon region (Figure 26) shows good agreement despite having small discrepancies in rainfall amplitudes especially in Day-3 and Day-5 forecasts.

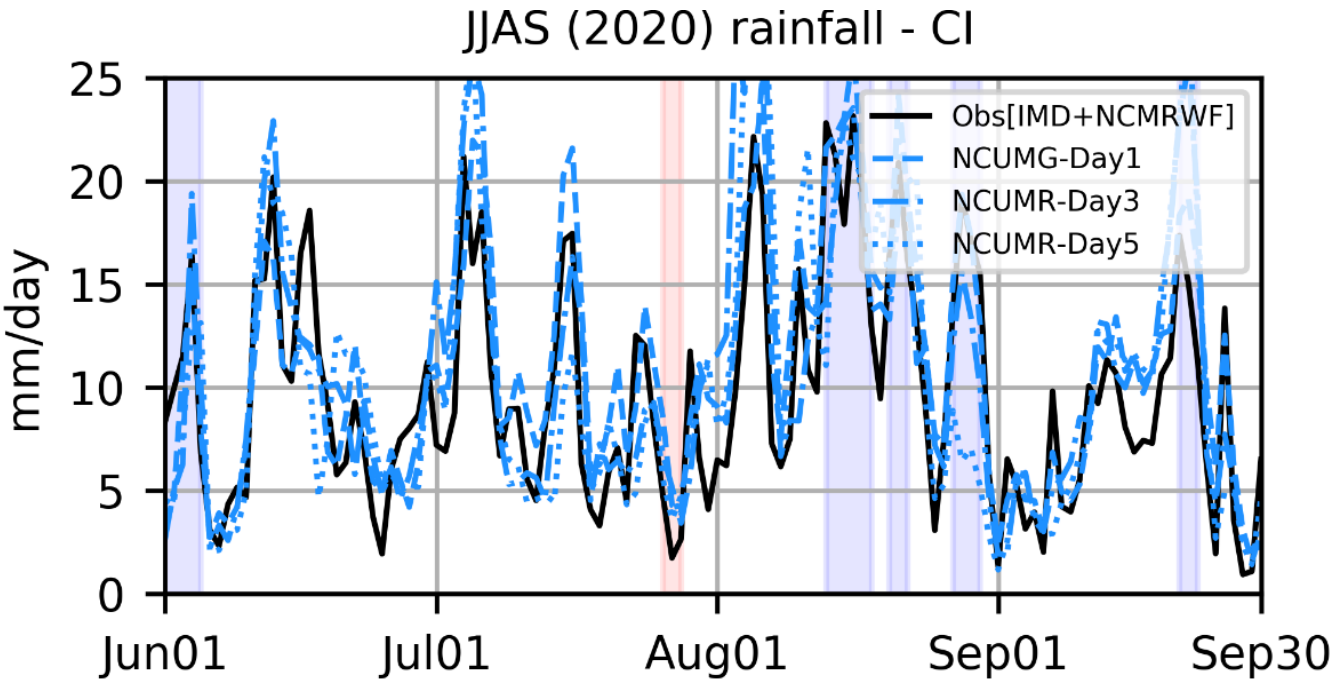

Figure 26: Time evolution daily rainfall from observations (solid black) and NCUMG model forecasts (blue) up to Day 5. Duration of the active (break) spells that occurred over CI are indicated in blue (red) patches.

6.4 Active and break composites:

The spatial maps of active and break days composite from observations and model forecast are shown in Figures27 and 28 respectively. Note that, while compositing the active spell days we have excluded the rainfall spell that occurred in June. The observed rainfall composite during active spell days shows rainfall maxima over the west coast, CI, and part of Bay of Bengal (BoB), Arakan coast (Figure 27a). These enhanced rainfall regions are well represented in the rainfall forecasts. However, with lead time overestimation of rainfall is seen over foothills of the Himalayas, Gujarat coast, and in BoB (Figures 27c and 27d). Over CI the overestimation in magnitude and rainfall area is very prominent. In all the forecast days rainfall is underestimated over the Telangana region (Figure 27d).

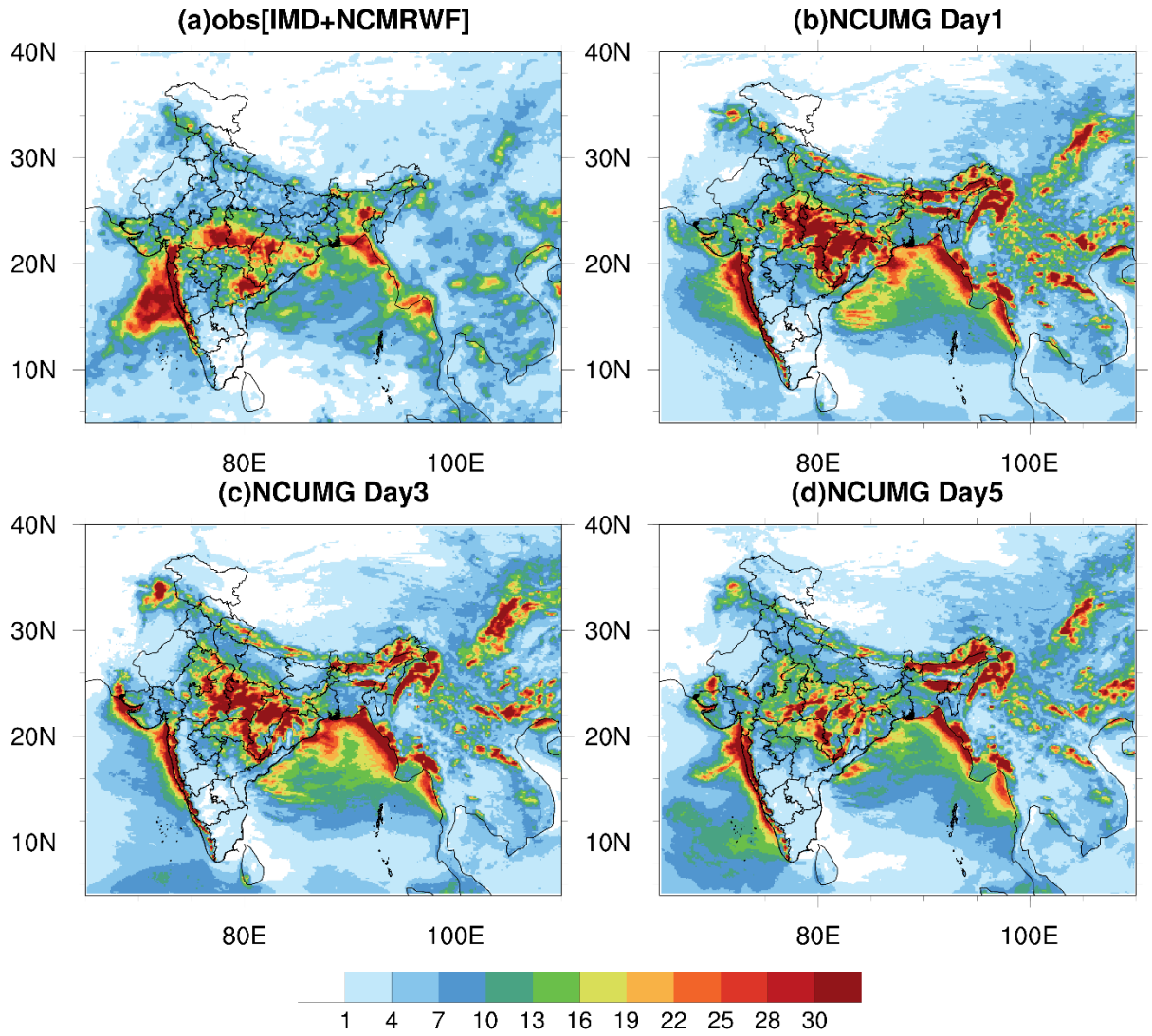

Figure 27: Composite rainfall (mm/day) map during the active spell days from (a) observations (IMD + NCMRWF merged product) and NCUM-G forecasts (b) Day-1, (c) Day-3 and (d) Day-5

The break composite shows no rainfall over most of the Indian land region. There are isolated and enhanced convection over the Arabian Sea and eastern equatorial Indian Ocean regions (Figure 28a). Model forecasts also exhibit suppressed convection over most of the Indian region, except over the foothills of the Himalayas (Fig 28b-28d). The isolated convection over AS appears more organized in the forecasts (Figures 28c and 28d) and occupied over a larger area. In addition, enhanced convection over the eastern equatorial Indian Ocean (EEIO) moving northward with lead time, which is noteworthy. In crux, though the active/break mean composites are reasonably represented in NCUM-G model forecasts, the spatial biases in rainfall need attention.

6.5Synoptic features:

During the monsoon 2020, there are 5 low-pressure systems (LPS) formed over the Indian region. Among the LPS none of them intensified into depression or deep depression. The LPS that formed over the Arabian sea

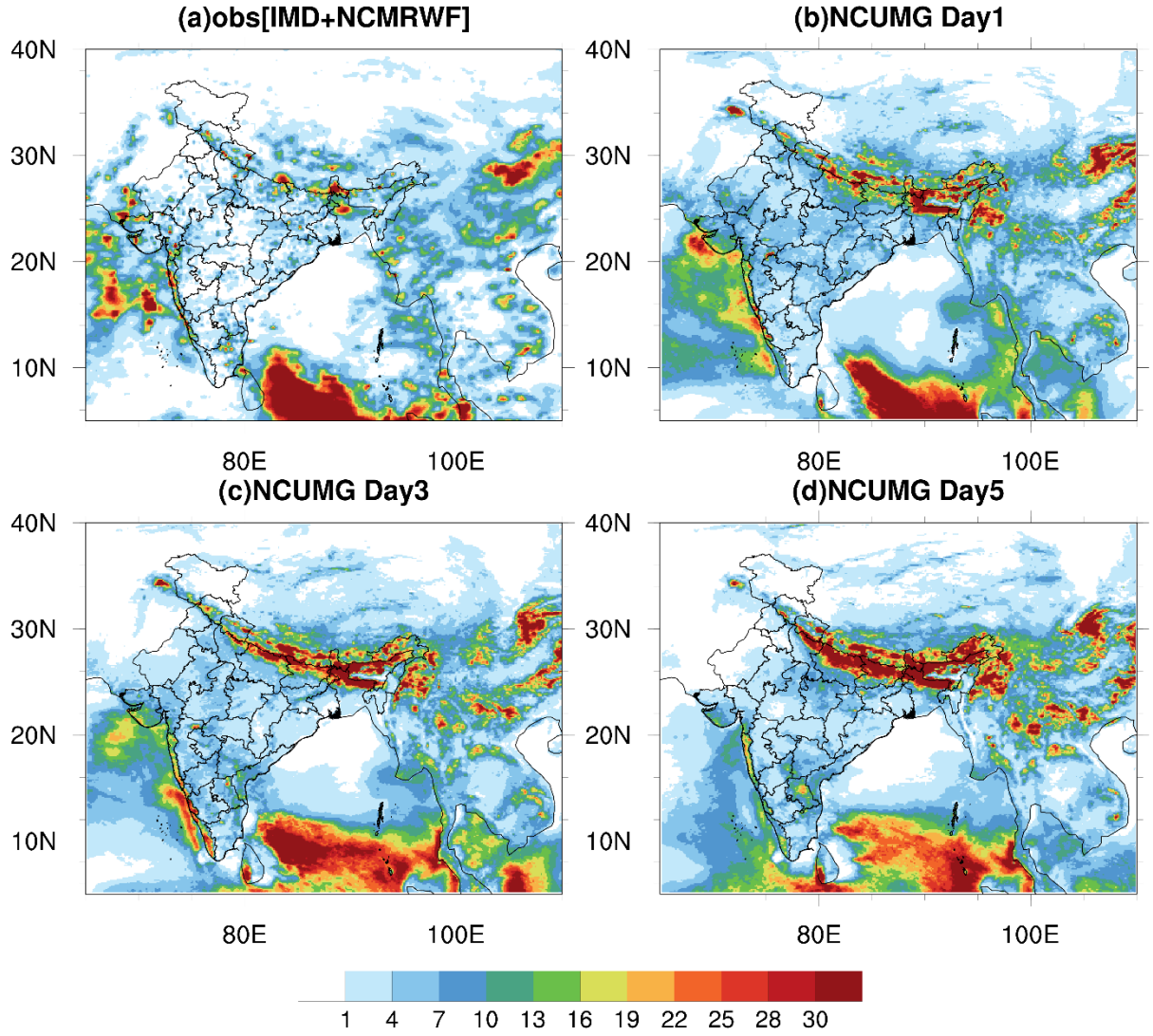

Figure 28: Same as Figure 3, but for rainfall composite during the break phase.

(AS) on 31st May 2020 has intensified into Severe Cyclonic Storm (SCS), NISARGA, on 2nd June 2020. The storm continued to move Northeastward and on 3rd June it crossed the Maharashtra coast.

6.5.1 Model forecasts – NISARGA:

Spatial maps of wind direction and magnitude (shading) for NCUM-G analysis and forecasts during 3rd June 2020 (landfall) time are shown in Figure 29. Here for better legibility, Day-1, Day-3, and Day-5 forecasts are presented along with the analysis. The best track is also superimposed on the spatial maps. It is clear from Figure 29that Day-1 forecasts from NCUM-Gare well matching with the analysis (Figure 29(a), however system forecast in Day-3 relatively intense with wind magnitudes reaching >30m/s. The position of the system in NCUMG Day-5 forecasts looks very feeble and decaying when compared with the analysis.

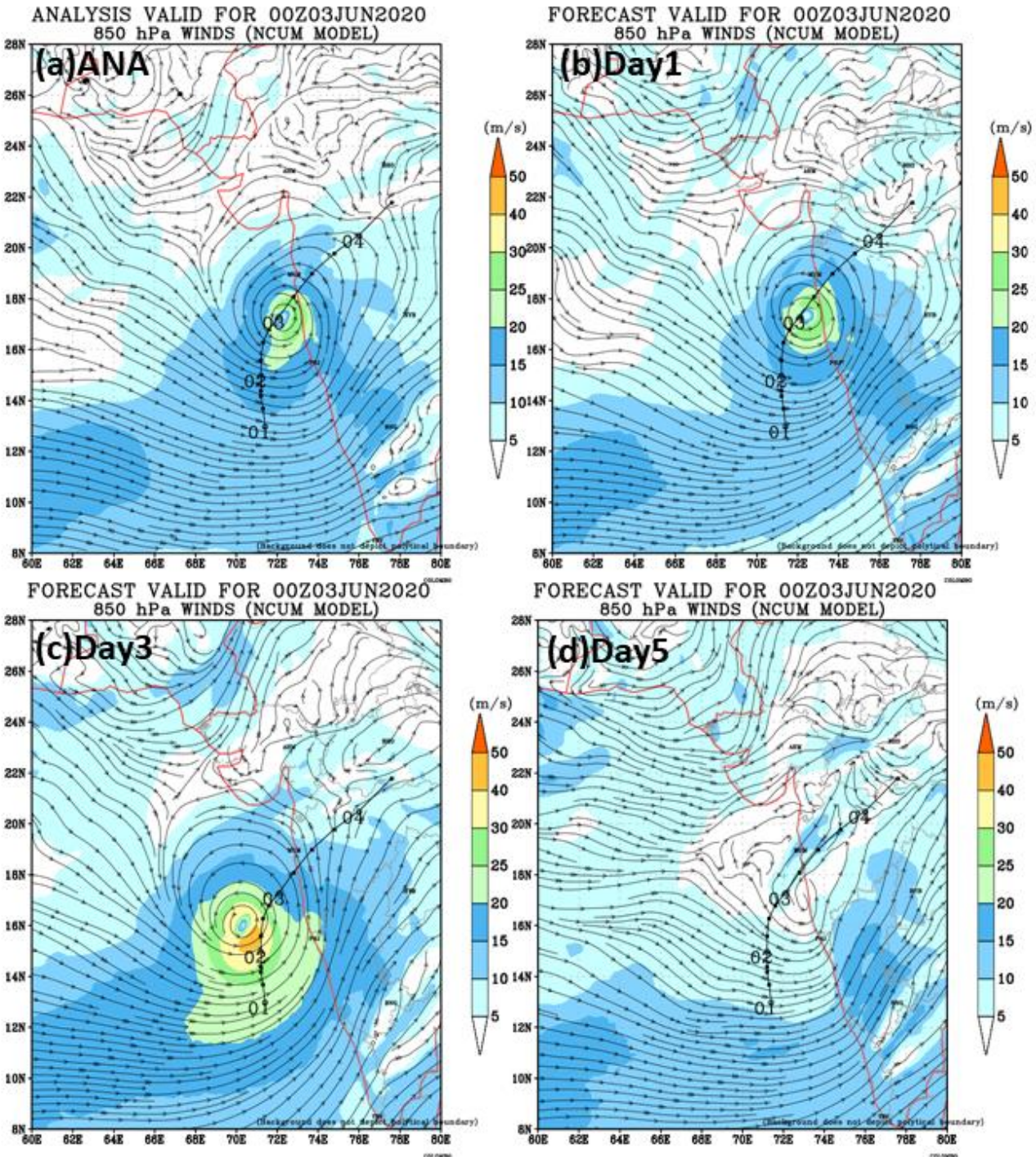

Figure 29: Spatial maps of Analysis; Day-1; Day-3 and Day-5 forecasts of wind speed (shading; units m/s) and direction from NCUM-G during the NISARAG cyclone on 3rd June 2020.

6.5.2 Low Pressure System in July 2020:

In July 2020, monsoon progression is weak due to many unfavorable features resulting in deficient rainfall for the country. The weak monsoon in July was mainly due to the absence of any major monsoon disturbance over the Bay of Bengal and due to the prevalence of a weak cross-equatorial (IMD report). The monsoon trough lay to the north of the normal position or close to the foothills of the Himalayas on many days

during July month. As a result, frequent and prolonged floods over northeastern India, Bihar, and adjoining areas of East Uttar Pradesh occurred and major parts of central and northwest India received deficient rainfall. However, in the first week of July, two low-pressure areas; one over coastal Saurashtra and neighborhood and the other over the Northwest Bay of Bengal off Odisha-Gangetic West Bengal coast (5-6 July) and their associated cyclonic circulation has caused wide-spread rainfall activity over eastern parts of Indian region.

Typical NCUM-G forecasts of the low-pressure system (LPS) that occurred over the Bay of Bengal region during the first week of July 2020 are shown in is given in Figure 30. Compared to NCUM-G analysis (Figure 30(a)), the Day-1 forecast shows the LPS over the head Bay of Bengal and a well-marked trough is extending along the east coast. In contrast, the Day-3 forecast in NCUM-G exhibits intense LPS more intense and shifted more towards the inland. With lead time the shift is further towards northwest towards land regions (Figure 30(d)). The strength of low-level jet not varying much in the forecast times and maximum wind speeds are situated over the Arabian sea region. Another noteworthy feature is that the strength of the LPS, as evidenced by the density and orientation of isobars, is enhancing in all the forecast timings.

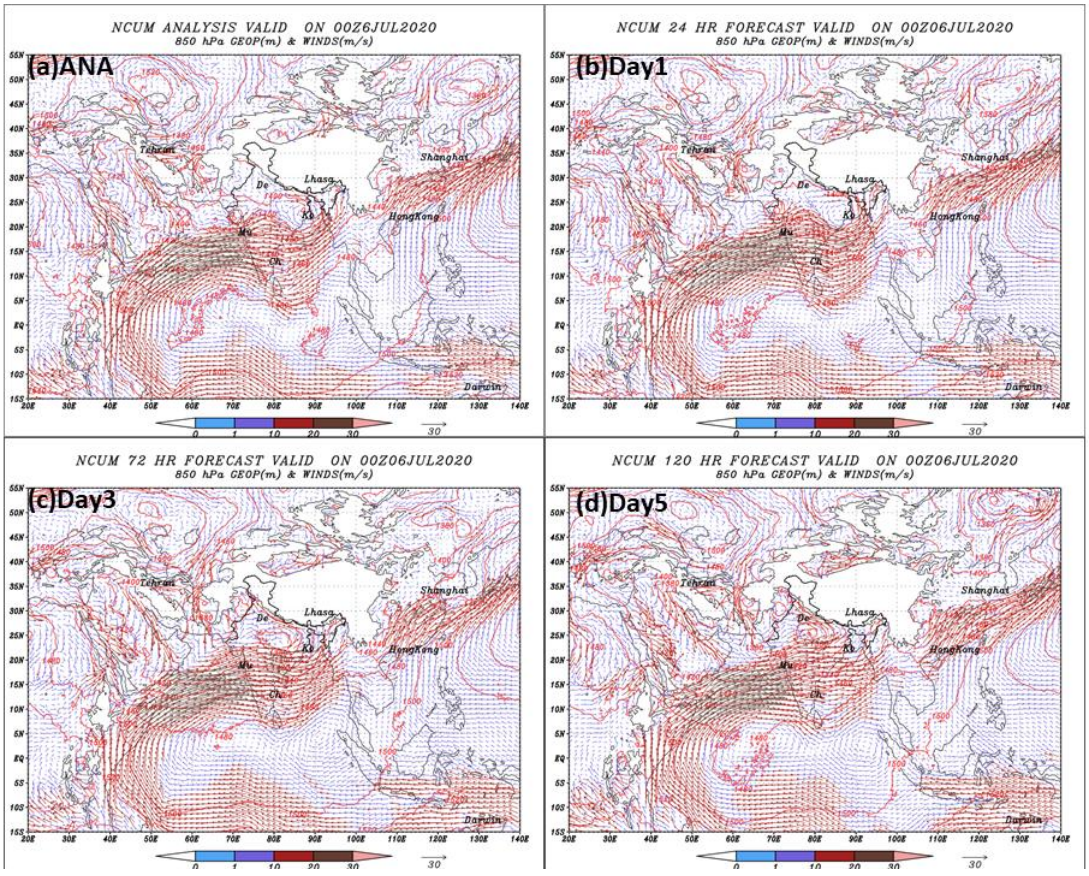

Figure 30: Spatial maps of Day-1 and Day-3 forecasts of wind speed (shading; units m/s) and direction (blue vectors) and geopotential height contours at 850 hPa level from NCUM-G during the Low pressure system (LPS) occurred on 6th July 2020.

6.6Heavy Rainfall frequencies in NCUM-G

In this section to assess NCUM-G forecasts in depicting the rainfall frequencies, the spatial distribution of maximum rainfall at each grid is shown in Figure 31. The figure shows the highest rainfall at each grid received during the season to compare the seasons' highest rain in observation and the two models. Over the land, region observed, and forecast rainfall events >10mm/day are considered. Observations show a near east-west belt over central India and the foothills of the Himalayas. Both regions have higher rainfall events with >20mm/day at isolated locations. NCUM-G forecasts show the highest number of rainfall >10 mm/day and 20 mm/day at all lead times over central India and hilly regions. NCUM-R forecasts, in contrast, have a much higher number >10mm/day and very few events of rainfall >20mm/day. To compare the rainfall frequencies over five different regions (Figure 32), we showed the normalized frequencies of rainfall shown for rainfall thresholds of up to 5 cm/day (Figure 33(a-e)). The plots show that NCUM-G overestimates the rainy days' frequencies in JJAS 2020 for up to 2cm/day rainfall threshold over most of the regions.

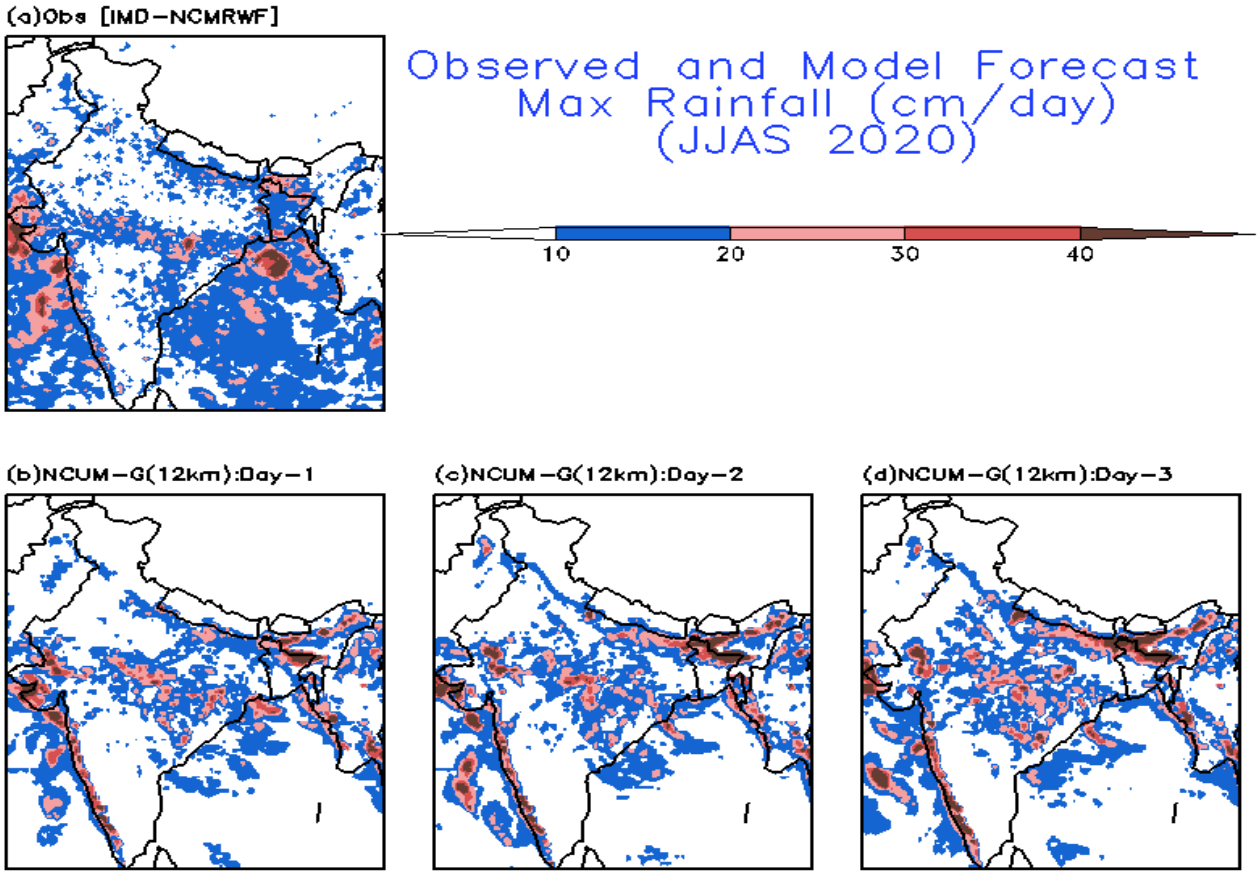

Figure 31 Observed and forecast highest seasonal rainfall in NCUM-G (b,c,d)

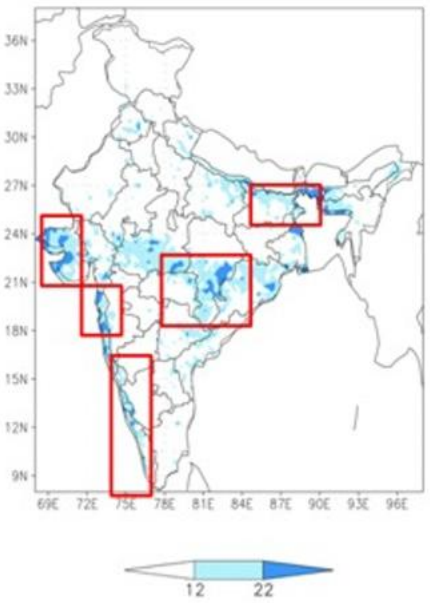

Figure 32: The five difference regions (outlined "red" color) selected for the estimation of rainfall skill from NCUM-G

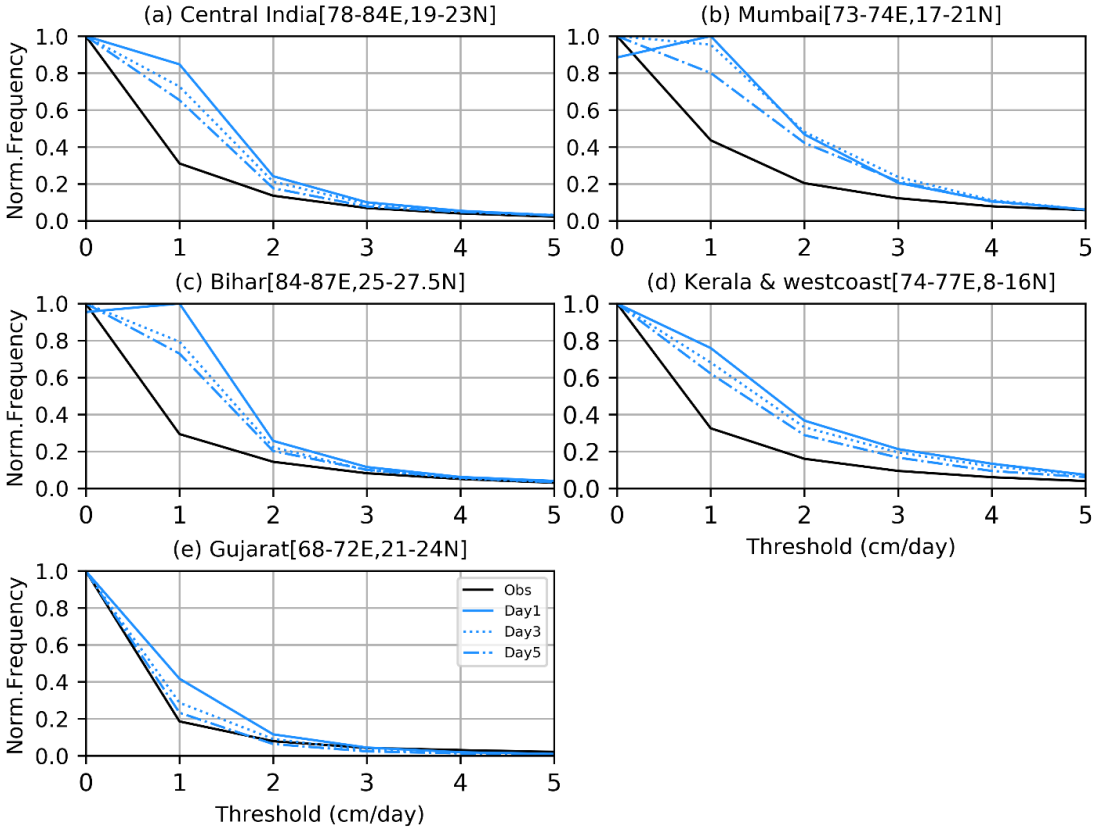

Figure 33: Observed and forecast rainfall frequencies over five regions of India in NCVUM-G forecasts

6.7 Investigation of sources of Rainfall Biases in NCUM-G

To further assess if the NCUM-G forecasts Figure 34 shows Day-1; Day-3 and Day-5 rainfall & 850 hPa wind biases in NCUM-G for monsoon JJAS2020. Note that here biases are computed against NCUM analysis. Like any other models both NCUMG model forecasts also exhibit biases in key variables. The wind biases in the NCUM-G show a prominent difference in both rainfall and 850hPa winds. For instance, wet bias is seen in NCUM-G over the west coast, foothills of the Himalayas, and south Bay of Bengal (BoB), and some parts over Central India (CI). The magnitude of these rainfall biases is increasing with lead time from Day-1 to Day-5 (Figure 34). Further, in NCUM-G forecasts the open ocean dry bias over the Arabian sea is persistent and enhancing lead time, perhaps due to the misrepresentation in surface fluxes. To further assess the source of biases, vertical profiles of T, q, and w averaged over central India (CI) are shown for NCUM-G Figure 35 with profiles from analysis as reference (red curve in Figure 35). While there is good agreement between models for T and q, vertical velocity (w) in NCUM-G is underestimated in forecasts compared to analysis at all levels above 600 hPa.Also, since the vertical profile of vertical velocity dictates radiation-convection feedbacks, the differences seen here i.e, weak ascent in the mid and upper troposphere in NCUM-G forecasts perhaps a likely cause of rainfall biases seen in earlier sections. In addition, low-level wind bias in the model forecasts is westerly over most of the Indian subcontinent and surrounding regions.

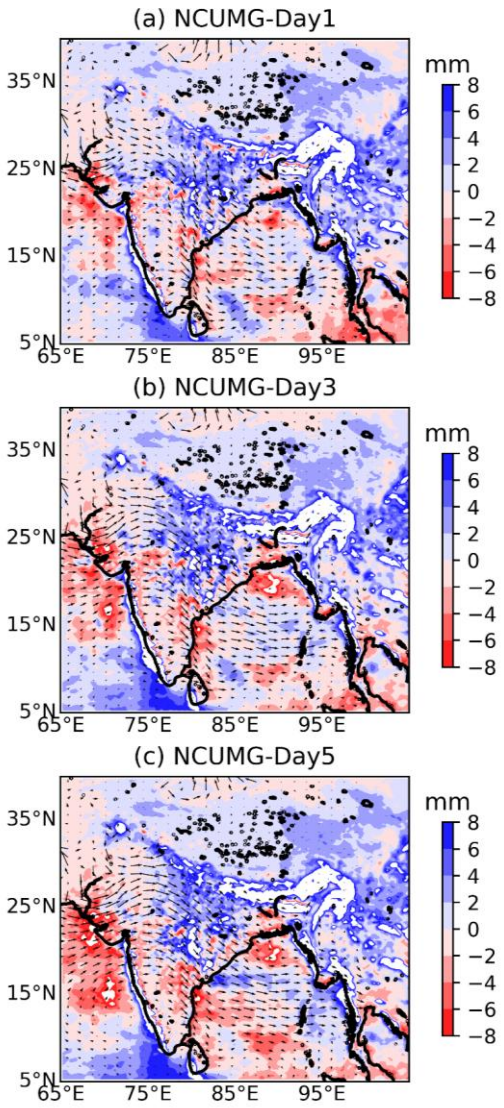

Figure 34:Rainfall and wind biases in NCUM-G during (a) Day-1; (b) Day-3 and (c) Day-5during JJAS 2020

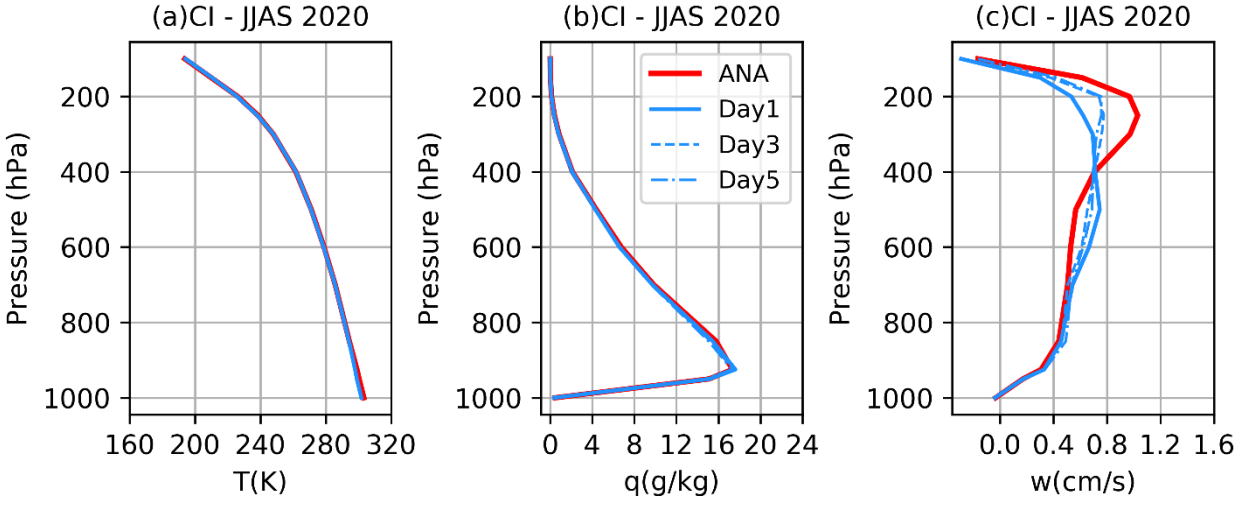

Figure 35:Mean vertical profiles of T, q, and w averaged over the core monsoon region

7. Subseasonal Variability of monsoon rainfall 2020:

Monsoon rainfall over the Indian landmass is crucial for agriculture and many other socio-economic benefits. Hence, previous sections (section 5 and section 6) discussed spatial and temporal variability of the rainfall and related biases in model forecasts. It is observed that the temporal variability of the rainfall from the model is quite well matching with the observed variability with some day-to-day biases. The evolution of daily biases, specifically over the core monsoon region, between model forecasts and observations shown in Figure 26. Overall, the bias varies between ±10mm/day with day-to-day variability with NCUM-G overestimating the rainfall in space and time (Figure 34 & Figure 26) with respect to observations. Moreover, the temporal evolution of rainfall also indicates some high- and low-frequency variability along with synoptic variability. Indeed, it is well known that the south Asian monsoon rainfall primarily varies at three scales: i) Synoptic scale (<7 days); ii) Quasi-biweekly scale (QBWO;10-20days); iii) Low-frequency instar-seasonal scale (ISO;30-60days) (ref: *Krishnamurti and Bhalme, 1976*).

The synoptic variability includes various systems such as the monsoon trough, low-pressure systems, monsoon depressions, cyclones, offshore vortices, etc. Here we first show that the synoptic variance from the model and observations (IMD+NCMRWF merged rainfall product) data. We have filtered 2-7 days (*bandpass filtered*) daily rainfall at each grid point over India. Figure 36 shows the synoptic variance from the observed and model forecasts. We can notice that the observed variance is extremely high over central India and the complex Himalayan region. This variance is particularly associated with the movement of monsoon trough and monsoon depressions. While the variance is also high in the head Bay of Bengal and the west coast of India. The variance in the Bay of Bengal is probably associated with the strong convection and evolution of low-pressure systems such as cyclones while the high variance on the west coast is primarily associated with the offshore vortices. The model-derived synoptic variance over the CI and complex Himalayan region overestimates the observed variance. It is also important to note the variance in the western ghats on the lee-ward side is very high in the NCUM-G forecast which is absent in observations. Further, the model forecast is also unable to represent the high variance over the west coast adjacent to the state of Maharashtra, especially in Day3 and Day5 forecasts. Further, the high variance in the head Bay is also absent in the model forecasts. Hence, the synoptic variance over the oceanic regions is severely underestimated by the NCUM-G model.

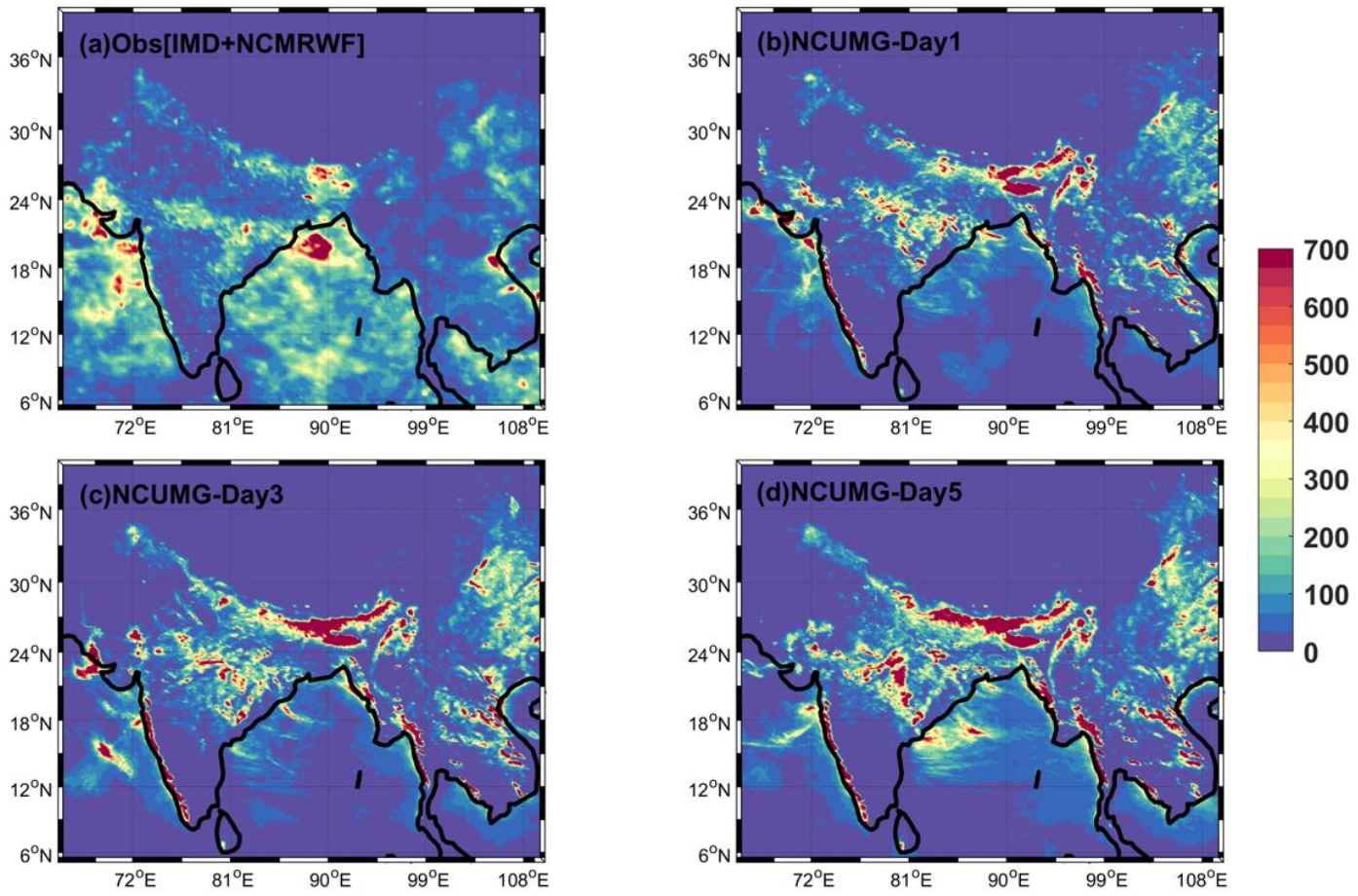

Figure 36:The spatial distribution of the synoptic variance (mm2) of rainfall from (a) Observed and NCUM-G (b) Day-1 (c) Day-3, (d) Day-5 forecasts

In Figure 37, we show the spectral analysis of the rainfall to understand the energetics of the intraseasonal variability of south Asian monsoon rainfall. Several studies have noted two modes (QBWO and ISO along with 20-30 day oscillation) at the intraseasonal scale in the Indian monsoon rainfall. The spectrum in Figure 37 depicts the above modes along with synoptic variability in the monsoon rainfall. It is very interesting to note that the QBWO is very well captured by the model with respect to observed rainfall, except in the Day5 forecast. However, there are significant differences in the power of the ISO between the observations and NCUM-G. The model power at the ISO band is extremely high relative to observed rainfall. Moreover, the ISO power is increasing with the lead day. In Figure 37, we do not have the time information when the two prominent modes are strong, therefore, we further did the wavelet analysis of daily average rainfall over the south Asian monsoon region. The wavelet spectrum preserves the time information giving the evolution of these two modes.

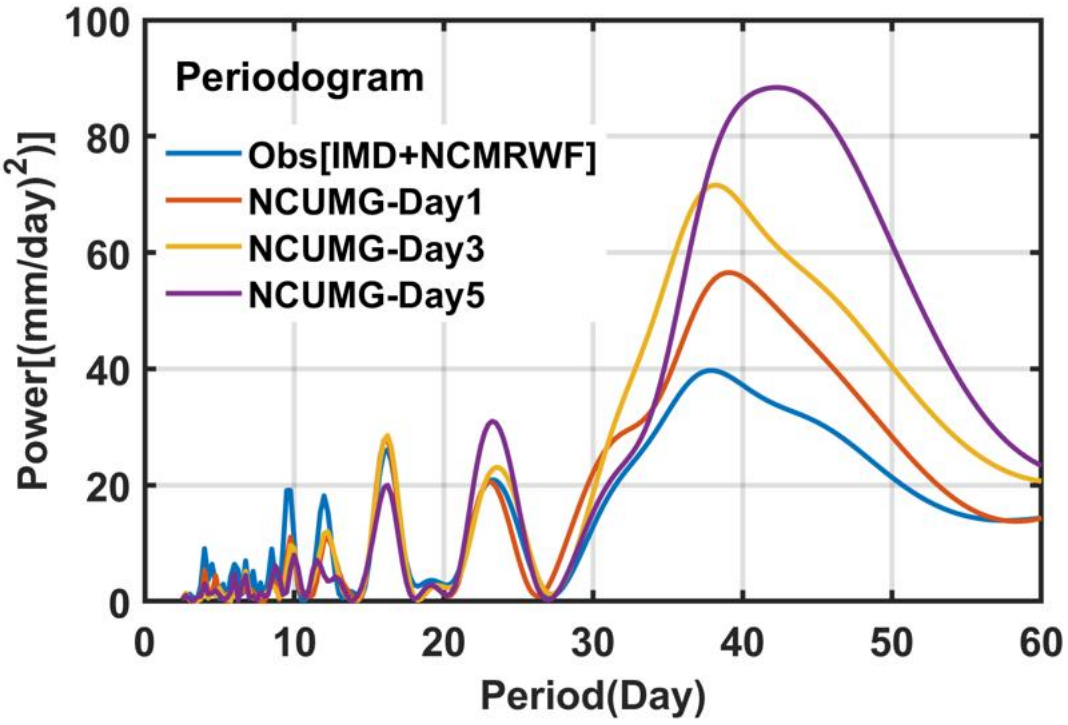

Figure 37:The spectral analysis of JJAS 2020 daily monsoon rainfall averaged over the south Asian monsoon region (5^{o}N-38^{o}N;65^{o}E -110^{o}E) from Observedand NCUM-G forecasts.

For instance, Figure 38 shows the wavelet spectra of observed and model forecast at different lead days. The observed spectra in Figure 38a indicate a prominent peak of ISO but not significant at 95% confidence and QBWO which is strong during August and September 2020. The model spectra also show these two prominent ISO and QBWO. It is also interesting to note the phase of ISO from the model forecast is very well matched with the observed spectra, however, the magnitudes are higher in the model which is also evident in Figure 37. The QBWO, on the other hand, shows some differences between the observed and model spectra. The model forecast at all lead days underestimated the magnitude of QBWO in September 2020. Further, the observed spectra in Figure 38a also show the significant synoptic variability of less than 7 days. The Day1 (Figure 38b) model forecast reasonably predicted this variability. Overall, the overestimation of the mean monsoon rainfall (Figure 16) in the model forecasts is primarily associated with the strong ISO modes in the model as compared to the observations. Since the ISO is one of the prominent northward propagating rainfall bands providing a significant amount of rainfall over the south Asian monsoon region. The model is unable to represent the spatial patterns of the synoptic variability, especially over the oceanic regions.

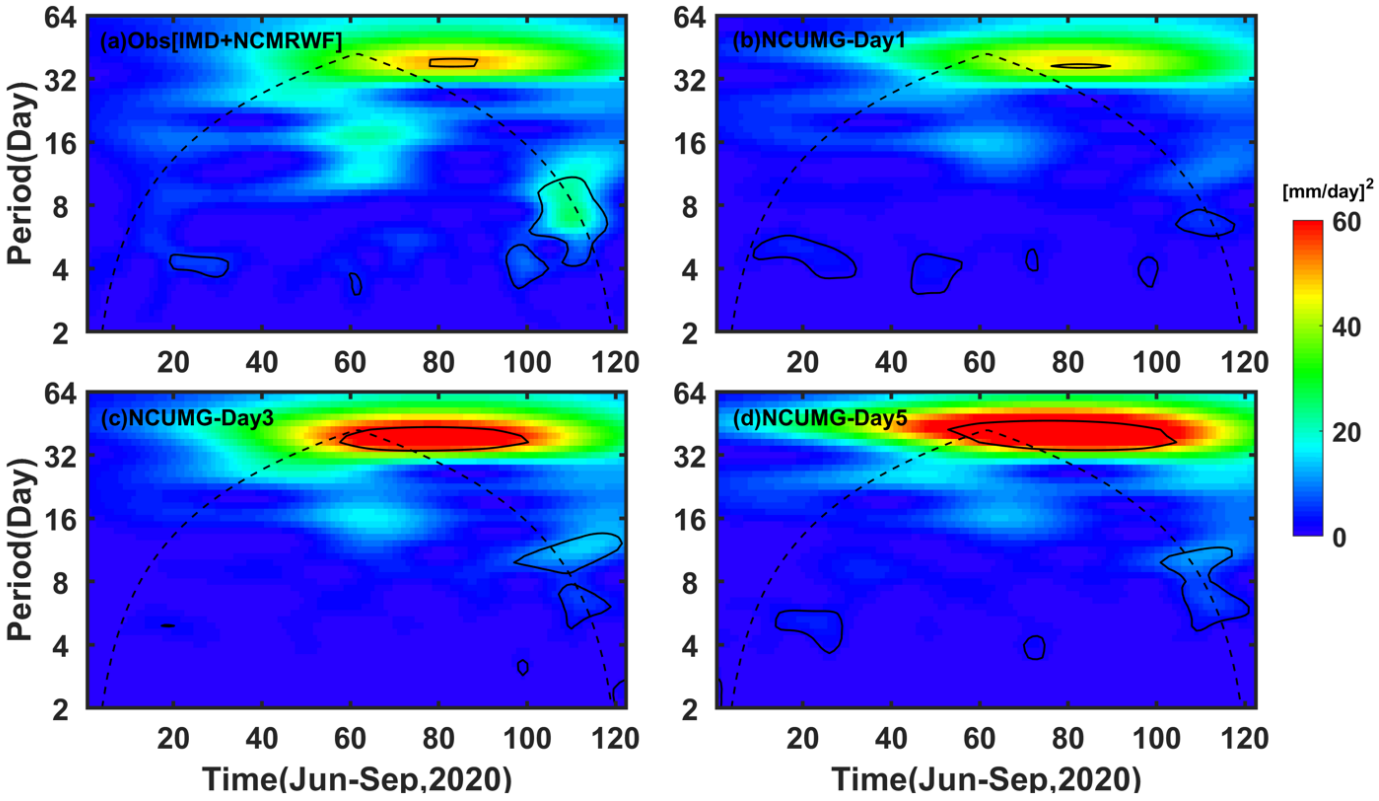

Figure 38:Wavelet Spectrum of Daily area-averaged rainfall over the south Asian monsoon region from (a) Observed and NCUM-G (b) Day-1 (c)Day-3, and (d)Day-5 forecasts. The dashed line indicates the cone of influence anything above is dubious. The solid lines indicate the 95% confidence interval based on chi-squared statistics.

8. Summary and Conclusions

This report documents performance of the NCMRWF model forecasts during monsoon season (JJAS) 2020. The verification results are presented to address (a) forecasters and (b) model developers. The information on biases in the forecast winds, temperature humidity, rainfall, etc., are crucial for the forecasters to interpret the model guidance for forecasting. Additionally, information on recent improvements in the model skill adds to confidence in the model forecasts. The results of the study can be summarized below.

NCUM-G Mean analysis and anomalies during JJAS 2020

- *The low-level wind anomalies at 850 and 700 hPa show prominently cyclonic circulation over the Arabian Sea indicating stronger monsoon current and relatively weaker winds over peninsular India and western India. Anomalous anticyclonic flow at 200hPa level depicts stronger Tibetan anticyclone with an eastward shift in the model analysis w.r.t ERA5 climatology.*
- *Low level (850 and 700 hPa) temperature anomalies show a dipole pattern with positive (negative) anomalies over the south and central (north) Arabian sea. The negative anomalies can be attributed to the strong northeasterly flow from land, particularly at 700 hPa. In contrast, the negative anomalies seen at mid-tropospheric levels (700-500 hPa) levels perhaps due to the enhanced mid-latitude westerlies in the analysis.*
- *The positive RH (wet) anomalies at 850 and 700 hPa over the entire Arabian Sea and parts of central and northern India and neighboring seas typically indicating the humid monsoon climate. However, during JJAS 2020, the Bay of Bengal and large parts of the equatorial Indian Ocean (central and eastern) exhibits negative (dry) RH anomalies.*

NCUM-G Systematic Errors

- *At 850hPa level, winds show enhanced westerlies (stronger monsoon flow), and the magnitude of bias increases with increasing lead time from Day-1 to Day-5. In addition, an easterly bias over the southern Bay of Bengal indicating a weakened westerly/southwesterly flow in the forecasts. At 200 hPa positive easterly bias (blue) over Indian peninsula along with neighboring Arabian Sea and Bay of Bengal in conjunction with negative westerly bias (red) over entire equatorial region suggests weakened tropical easterlies.*
- *A systematic warm bias (>1℃) over most of the Indian land region and the neighboring Arabian Sea, Bay of Bengal, and the equatorial Indian Ocean, which increases with lead time is noticed at lower levels (850 hPa). At 200 hPa level, the land (sea) regions partly feature cold (warm) bias in all the forecasts with magnitudes of about 0.5℃. On a similar note, RH systematic errors also show strong dry bias over land regions and mild wet bias over Sea.*
- *Systematic errors in the Vertical Integrated Moisture Transport(VIMT) show positive (wet) bias over India suggesting that strong positive wind biases offset the impact of negative (dry) bias in RH forecast.*

Rainfall Forecast Verification

- *Model forecasts overestimate the isolated high rainfall amounts (>10mm/day) over the core monsoon region indicating overestimation of observed rainfall over land and neighboring seas. Interestingly, model forecasts over most parts of India show an exceedingly high number of total rainy days. The excessive*

forecast rainfall in the model over hilly regions of the west coast, NE Indian, and Himalayas is attributed to an enhanced model grid resolution of 12km.

- *Over the dry regions of NW India, central and eastern peninsula, the model predicts a far lower number of moderate rain days. The counts are overestimated over NE India and foothills of the Himalayas, indicating the model's tendency to frequently predict higher rainfall amounts. In addition for higher amounts (9-30mm/day), there is no frequency bias, but the skill is low as indicated by CSI, PSS, and SEDI in the model forecasts.*
- *Verification of PQPF presented in terms of Brier Score (BS) for higher rainfall thresholds (15.5 & 65.4 mm/day) is lower implying a better match between the forecast probability and observed frequency for heavier rainfall. For lower rainfall thresholds(>1 and 2.5mm/day) it is seen that the BS increases slightly with an increase in forecast lead time which is indicative of decreasing model skill.*

2018-2020: Verification scores

- *A comparison of RMSE in the 850 & 200 hPa winds and temperature during 2018-2020 show a general decreasing trend and lowest RMSE values during JJAS 2020.*
- *Verification of rainfall forecasts for higher rainfall amounts (2cm and 5cm/day) during 2018-2020 show an overall improvement in the skill with some interannual variations.*

Representation of Active/Break cycles, ISO Modes and Synoptic Variance

- *The summer monsoon system varies at intraseasonal scales comprising of active/break cycles. During the 2020 monsoon period (JJAS) there are 5 active cycles and 1 break condition. The NCUM model can predict these active/break conditions very well for different forecast lead times (Day 1- Day 3). However, the model forecast has a bias in terms of rainfall intensities and its spatial distribution, specifically in Day-2 and Day-3 forecasts, during active/break cycles.*
- *The active/break cycles are the manifestation of westward propagating quasi-biweekly mode (QBWO;10-20days) and northward propagating instar-seasonal (ISO) events (30-60 days). The QBWO is very well captured by the model forecasts with respect to iMERG rainfall. However, the NCUM model energetic at the ISO band is extremely high relative to observed rainfall. The overestimation of the accumulated monsoon rainfall in NCUM-G model forecasts might be associated with the strong ISO modes in the model as compared to the observations.*
- *The model-derived synoptic variance over the core monsoon region and complex Himalayan region overestimates the iMERG derived variance. While the synoptic variance over the oceanic regions is severely underestimated by the NCUM-G model.*

References.

Ashrit R, Sharma K, Dube A, Iyengar G R, Mitra A K and Rajagopal E N 2015: Verification of short-range forecasts of extreme rainfall during monsoon; Mausam 66 375–386, 607 https://metnet.imd.gov.in/mausamdocs/16633˙F.pdf.

Auligne, T., McNally, A. P., Dee, D. P., 2007. Adaptive bias correction for satellite data in a numerical weather prediction system.Q J ROY METEOR SOC, 133, 631-642. https://doi.org/10.1002/qj.56

Barker, D., 2011. Data assimilation-progress and plans, MOSAC-16, 9-11 November 2011, Paper16.6.

Buizza, R., P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of theECMWF, MSC, and NCEP Global Ensemble Prediction System. Mon. Wea. Rev., 133, 1076-1097.

Bush, J. S., Turner, A.G., Woolnough, S. J., Martin, G. M., and Klingaman, N. P (2015); The effect of increased convective environemnt on Asian monsoon biases inthe Met UM global circulation model Q. J. R. Meteorol. Soc. 141: 311–326, DOI:10.1002/qj.2371

Cameron, J., Bell, W., 2018. The testing and planned implementation of variational bias correction (VarBC) at the Met Office 21.

Clayton, A.M., Lorenc, A.C., Barker, D.M., 2013. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office: Hybrid Ensemble/4D-Var Data Assimilation. Q.J.R. Meteorol. Soc. 139, 1445–1461.https://doi.org/10.1002/qj.2054.

Dube, A., Raghavendra Ashrit, Harvir Singh, G. R. Iyengar, and E. N. Rajagopal, 2016a: Verification of Medium Range Probabilistic Rainfall Forecasts over India. Pure and Applied Geophysics, 173(7), 2489-2510

Dube, A., Raghavendra Ashrit, Harvir Singh, Kopal Arora, G. R. Iyengar, and E. N. Rajagopal, 2016b: Evaluating the performance of two Global ensemble forecasting systems in predicting rainfall over India during the southwest monsoons. Meteorological Applications 24(2), 230-238

Harris, B.A., Kelly, G., 2001. A satellite radiance-bias correction scheme for data assimilation. Q.J Royal Met. Soc. 127, 1453–1468.https://doi.org/10.1002/qj.49712757418.

Hersbach H, Bell B, Berrisford P, et al. 2020: The ERA5 global reanalysis. Q J R Meteorol Soc. 2020;146:1999–2049. https://doi.org/10.1002/qj.3803

John P George, S. Indira Rani, A. Jayakumar, Saji Mohandas, SwapanMallick, R. Rakhi, M. N. R. Sreevathsa and E. N. Rajagopal 2016: NCUM Data Assimilation System. Technical Report, NMRF/TR/01/2016., 20p.

George, J.P., Rani, S.I., Jayakumar,A., Mohandas, S., Mallick,S., Lodh, A., Rakhi,R., Sreevathsa, M., N., R., Rajagopal, E., N., 2016. NCUM-G Data Assimilation System, NMRF/TR/01/2016, 20p.

Gregory, D., Rowntree, P. R.,1990. A massflux convection scheme with representation of cloud

Jolliffe, I. T., and D. Stephenson, 2012: Forecast Verification: A Practitioner's Guide in Atmospheric Science, John Wiley & Sons, Ltd

Kumar Sumit, A. Jayakumar, M. T. Bushair, Buddhi Prakash J., Gibies George, Abhishek Lodh, S. Indira Rani, Saji Mohandas, John P. George and E. N. Rajagopal 2018: Implementation of New High Resolution NCUM Analysis-Forecast System in Mihir HPCS. NMRF/TR/01/2019, 17p.

Kumar Sumit, M. T. Bushair, Buddhi Prakash J., Abhishek Lodh, Priti Sharma, Gibies George, S. Indira Rani, John P. George, A. Jayakumar, Saji Mohandas, Sushant Kumar, Kuldeep Sharma, S. Karunasagar, and E. N. Rajagopal 2020: NCUM Global NWP System: Version 6 (NCUM-G:V6), NMRF/TR/06/2020

Krishnamurti, T. N., &Bhalme, H. N. (1976). Oscillations of a MonsoonSystem. Part I. Observational Aspects, Journal of Atmospheric Sciences,33(10), 1937-1954. DOI:https://doi.org/10.1175/1520-0469(1976)033<1937:OOAMSP>2.0.CO;2

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

Lorenc, A.C., 2003. Modelling of error covariances by 4D-Var data assimilation. Q.J.R. Meteorol. Soc. 129, 3167–3182.https://doi.org/10.1256/qj.02.131

Levine, R.C. and Martin, G.M. (2018) On the climate model simulation of Indian monsoon low-pressure systems and the effect of remote disturbances and systematicbiases. Climate Dynamics, 50, 4721– 4743

Mamgain, A., Sarkar, A., Dube, A., Arulalan, T., Chakraborty, P., George, J. P., and Rajagopal, E. N., 2018: Implementation of Very High Resolution (12 km) Global Ensemble Prediction System at NCMRWF and its Initial Validation, NMRF/TR/02/2018, 21p.

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

Mitra, A. K., I. M. Momin, E. N. Rajagopal, S. Basu, M. N. Rajeevan and T. N. Krishnamurti, 2013, Gridded Daily Indian Monsoon Rainfall for 14 Seasons: Merged TRMM and IMD Gauge Analyzed Values, J. of Earth System Science, 122(5), 1173-1182.

Murphy A. H., 1988. Skill Score based on Mean Squared Error and their relationship to the correlation Coefficient. Mon. Wea. Rev., 124, 2417-2424

Murphy A. H., 1996. General decomposition of MSE based Skill Scores: Measures of some Basic Aspects of Forecast Quality. Mon. Wea. Rev., 116, 2353-2369.

Rabier, F., Thépaut, J.-N., Courtier, P., 1998. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. Q.J Royal Met. Soc. 124, 1861–1887.https://doi.org/10.1002/qj.49712455005

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., Simmons, A., 2007. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. Q.J.R. Meteorol. Soc. 126, 1143–1170.https://doi.org/10.1002/qj.49712656415

Rajagopal, E. N., Iyengar G. R., George J.P.,Gupta, M.D., Mohandas, S., Siddharth, R., Gupta, A., Chourasia, M., Prasad, V. S., Aditi, Sharma, K. Ashish, A., 2012. Implementation of the UM model based analysis–forecast system at NCMRWF, NMRF/TR/2012, 45p.

Rani, S.I., Taylor, R., Sharma, P., Bushair, M.T., Jangid, B.P., George, J.P., Rajagopal, E.N., 2019. Assimilation of INSAT-3D imager water vapour clear sky brightness temperature in the NCMRWF's assimilation and forecast system. J Earth Syst Sci 128, 197.

Rawlins, F., Ballard, S.P., Bovis, K.J., Clayton, A.M., Li, D., Inverarity, G.W., Lorenc, A.C., Payne, T.J., 2007. The Met Office global four-dimensional variational data assimilation scheme. Q.J.R. Meteorol. Soc. 133, 347–362.https://doi.org/10.1002/qj.32

Sarkar, A., Chakraborty, P., George, J. P., and Rajagopal, E. N., 2016: Implementation of Unified Model Based Ensemble Prediction System at NCMRWF (NEPS), NMRF/TR/02/2016, 26p.

Sharma, K., Ashrit, R., Kumar, S. et al. Unified model rainfall forecasts over India during 2007–2018: Evaluating extreme rains over hilly regions. J Earth Syst Sci 130, 82 (2021). https://doi.org/10.1007/s12040-021-01595-1

Stephenson D.B., B. Casati, C.A.T. Ferro and C.A. Wilson, 2008: The extreme dependency score: a non-vanishing measure for forecasts of rare events. Meteorol. Appl., 15, 41-50.

Wang, B., Wu, R., & Lau, K-M. (2001). Interannual Variability of theAsian Summer Monsoon: Contrasts between the Indian and the Western NorthPacific–East Asian Monsoons, Journal of Climate, 14(20), 4073-4090.DOI:https://doi.org/10.1175/1520-0442(2001)014<4073:IVOTAS>2.0.CO;2

Wang, B., Ding, Q., & Joseph, P. V. (2009). Objective Definition of the Indian Summer Monsoon Onset, Journal of Climate, 22(12), 3303-3316.DOI:https://doi.org/10.1175/2008JCLI2675.1

Walters, D., and co-authors: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geosci. Model Dev., 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, 2017

Wilks D S 2011 (eds) Statistical methods in the atmospheric 807 sciences; 3rd edn, Elsevier, 676p