NMRF/RR/06/2020

VERIFICATION REPORT

Quantitative precipitation forecast (QPF) verification for *iFlows* over Mumbai during southwest monsoon 2020

Mohana S. Thota, Niranjan Kumar Kondapalli, Sagili Karunasagar, Raghavendra Ashrit

January 2021

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

Verification of integrated Flood warning system (iFlows) over Mumbai during southwest monsoon 2020

Mohana. S. Thota, Niranjan Kumar Kondapalli, Sagili Karunasagar, Raghavendra Ashrit

January 2021

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

Ministry of Earth Sciences
National Centre for Medium Range Weather Forecasting
Document Control Data Sheet

1	Name of the Institute	National Centre for Medium Range Weather Forecasting
2	Document Number	NMRF/RR/06/2020
3	Date of Publication	January 2021
4	Title of the document	Verification of integrated Flood warning system (iFlows) over Mumbai during southwest monsoon 2020
5	Type of the document	Verification Report
6	Number of pages, figures, and Tables	29 Pages, 14 Figures and 1 Tables
7	Authors	Mohana. T. S, Niranjan Kumar Kondapalli, Sagili Karuna Sagar, Raghavendra Ashrit
8	Originating Unit	National Centre for Medium Range Weather Forecasting (NCMRWF)
9	Abstract (brief)	The aim of this report is to verify the rainfall forecasts from the global (NCUMG) and regional models operational at NCMRWF with an integrated Flood warning system over Mumbai region. Analyses suggest that models' performance is hindered with the gridded rainfall data during monsoon season (June-September) of 2020. In contrast, the verification with the station observations, the regional model (NCUMR) performance is significantly improved in terms of high values of equitable threat score, and low false alarm ratio. The simple correlation analysis shows that NCUMR has better correlations than NCUMG with station data indicating the strength of high quality observations for the rainfall verifications.
10	References	5
11	Security classification	Unrestricted
12	Distribution	General

Table of contents

Sl. No.		Page No.
	Abstract	5
1	Introduction	6
2	Data and Methodology	7
3	Results	8
4	Observed rainfall scenario	8
5	Verification – Gridded data	8
6	Verification – Station data	10
6	Summary and Conclusions	13
7	References	15
8	Appendix -I (Tables)	16
9	Appendix- II (Figures)	17

Abstract

This report presents a brief summary of the verification of the NCMRWF global and regional models' rainfall forecasts during the monsoon season (June-September, JJAS) 2020 with an integrated Flood warning system (iFlows) over Mumbai region. The verification of forecast is presented for NCMRWF Unified Model (NCUMG) (12 km grid resolution) and Regional models, NCUMR (4 km grid). Analyses suggest that models' performance is hindered when used the gridded rainfall data during monsoon season in 2020. In contrast, when the verification is carried with available high dense station observations, the regional model (NCUMR) performance is significantly improved, with high ETS and low FAR values. Simple point wise correlation statistics show that NCUMR has better correlations than NCUMG with station data indicating the strength of high quality observations for the rainfall verifications.

1. Introduction:

Indian Summer Monsoon Rainfall (ISMR), during June through September (JJAS), affects lives of large number of people and agriculture sector due to its large variability over subcontinent (ex. Webster 1998; Gadgil 2003). Among the rainfall coherent regions, the west coast of India receives very heavy rainfall (ex. Rao 1976; Francis and Gadgil 2006). Mumbai city, which is situated on the west coast, has the history of flooding during extreme rainfall events. The city had experienced massive floods in July 2005, and most recently in 2017. In general, the heavy rainfall over Western Ghats and surrounding regions is attributed to the ascent of moist air from Arabian Sea over the orographic regions. In addition, numerous studies have also indicated the role of mid tropospheric cyclones (MTC) and offshore vortices in generating heavy rainfall over Mumbai region (Pattanaik, and Rajeevan, 2010).

Recently, Ministry of Earth Sciences (MoES), in coordination with the Municipal Corporation of Greater Mumbai (MCGM) launched an integrated flood warning system known as *iFLOWS*. Mumbai is second city in India after Chennai to have such a system. The *iFLOWS* system comprises of seven modules, namely data assimilation, flood, inundation, vulnerability, risk, dissemination, and decision support system. The complete flood warning system is giThe system incorporates numerical weather prediction (NWP) models from National Centre for Medium Range Weather Forecasting (NCMRWF), India Meteorological Department (IMD) and observations from the rain gauge network stations.

The aim of this report is to evaluate the performance of rainfall forecasts from global and regional operational models of NCMRWF, and the verification is performed for 2020 south west monsoon season over Mumbai region.

2. Data and Methodology:

In the present analysis, the verifications are performed with the NCMRWF Global Unified Model (NCUMG) with a horizontal resolution of 12 x18 km (Sumit et al. 2018), and the NCMRWF Regional Unified Model (NCUMR) with 4x4 km (Abhishek et al. 2020). The gridded rainfall observations are taken from daily merged gauge and satellite product available at NCMRWF (Mitra et al. 2009) for the monsoon season (JJAS) 2020. Note that, the verification of the forecast is carried out upto Day-3 forecasts only. Standard scores (equitable threat score (ETS), hits) and statistics (mean error) used in the forecast verification are given below (https://www.cawcr.gov.au/projects/verification/#Types_of_forecasts_and_verifications).

$$Mean\ error\ (ME) = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i) \ldots\ldots (1)$$

$$ETS = \frac{hits - hits_{random}}{hits + misses + false\ alarams - hits_{random}} \ldots\ldots (2)$$

$$hits_{random} = \frac{(hits+misses)(hits+fasle\ alarms)}{total} \ldots\ldots\ldots (3)$$

ETS range from -1/3 to 1, with 1 indicating prefect skill and 0 as no skill.

Hits, false alarms and *misses* are calculated based on 2x2 contingency table between forecast and observations, given below.

<table>
<tr><td rowspan="5">Forecast (F)</td><td colspan="4">Observed (O)</td></tr>
<tr><td></td><td>Yes</td><td>No</td><td>Total</td></tr>
<tr><td>Yes</td><td>Hits</td><td>False alarms</td><td>F - Yes</td></tr>
<tr><td>No</td><td>Misses</td><td>Correct negatives</td><td>F - No</td></tr>
<tr><td>Total</td><td>O –Yes</td><td>O - No</td><td>Total</td></tr>
</table>

3. Results

a) Observed rainfall scenario:

Figure 1 shows the time series of JJAS rainfall occurred over the two stations (Coloba and Santacruz) of Mumbai. For completeness, in Figure 1, the time series of gauge and satellite merged (Mitra et al 2009) and global precipitation mission (GPM) rainfall (https://gpm.nasa.gov/data/directory) values are also shown. Here, in the gridded observations the nearest station data is obtained by using the minimum distance method (Pythagorean theorem). Daily accumulations of the rainfall products are done from 03Z-03Z to match with stations obseravtions. It is clear from Figure 1 that majority of the days are rainy with rainfall amounts > 20mm/day. However, in June, rainy events are relatively low with threshold not exceeding 30 mm/day. As the monsoon flow getting strengthened, rainy events with threshold 100mm/day are seen during July and August. There are occasions during the months of July and August where rainfall amounts reached as high as 200mm/day. The seasonal total rainfall is ~3670mm, third highest seasonal rain in past 62 years (IMD report). These extremely heavy rainfall events caused lots of flooding activity and affected several regions over Mumbai. Also from the Figure 1 it is apparent that though the gridded satellite products are able to pick the heavy rainfall episodes reasonably well (GPM overestimates slightly at lower observed values), there are instances when the gridded products failed to pick up the phase and magnitude of the rainfall event, highlighting the limitation of using gridded products in rainfall verification.

b) Verification - Gridded data

As a preliminary step, to assess the performance of the modeling systems (both global and regional) operational at NCMRWF, we have used the rainfall from the gauge and satellite merged gridded observations (Mitra et al. 2009) and a few statistics are computed. Note that, here the model forecasts are re-gridded to 0.25^{o} x 0.25^{o} grid to match with the gridded observations (IMD-NCMRWF merged rainfall hereafter referred as gridded observations). Figure 2 shows the JJAS mean rainfall from observations and global and regional model forecasts. It is clear from the figure that the observed rainfall maximum (>30 mm/day), which is on the Eastern parts of the Mumbai region, is not able to predict well by both NCUMG and NCUMR models. For instance, the rainfall maxima, seen in gridded observations, are shifted to westward in the NCUMG and it is more intense in Day-2 forecast. In contrast the rainfall maximum is large in Day-1 forecast in the regional model and with increased lead time reduction in rainfall magnitudes are noticed. On a similar note, mean error (ME) of model forecasts is also showing considerable difference among the model forecasts (Figure 3). ME for the NCUMG show large underestimation (~-10mm) in all the forecast days' w.r.t gridded observations (Figure 3). This underestimation is also seen in the regional model for Day-2, Day-3 forecasts except for Day-1 forecast. Notably, regional model Day-1 forecast exhibits overestimation with magnitudes reaching upto 5-10mm. It is interesting to note that, the rainfall underestimation is more in the regional model for Day-2 and Day-3 forecasts. The correlation coefficients between gridded observations and models are computed shown as spatial maps (Figure 4). The figure shows slightly better correlations (>0.5) in NCUMG than NCUMR. As expected, the correlations decrease with increase forecast lead time. Overall, the comparison of the mean rainfall and the statistics computed with the gridded observations showing considerable discrepancies, indicating the limitations in using the gridded observations to assess the model performance (Figure 2, 3, 4).

Thus, as an alternative, we have used the available rain gauge (ground based) station observations to verify both the model rainfall forecasts. Since the station data during the JJAS season is varying, instead of the seasonal mean rainfall, a few extreme rainfall (ER) events (~200 mm/day) that occurred over Mumbai region are verified. As seen in Figure 1, there are 4 ER events occurred during JJAS 2020 (4th July, 16th July, 5th August, and 23rd September). Among these 4 events we have excluded 16th July, due to non-availability of observations, hence the verification is performed for the remaining three events. Large difference in magnitude and spatial distribution of rainfall amounts are noticed in NCUMG for the 3 ER events. This could be due to the coarse model resolution of NCUMG which is not able to resolve the localized rainfall events. In contrast, NCUMR represent the rainfall magnitude and distribution slightly better. Despite the mismatch in rainfall amounts, the "eye ball" verification of these events are well predicted, especially in high resolution NCUMR model (Figure 5 - 7). The crux is though the global and the regional models depict significant differences in predicting the seasonal rainfall over Mumbai region, models performance in representing the magnitude and spatial distribution of rainfall for extreme rainfall events (~200mm/day) are reasonably well. It reaffirms the point that strength of high-resolution regional model for understanding and predicting the extreme rainfall events.

c) Verification – Station data

Along with the gridded observations, we also used the station data to evaluate the models' performance. The time series of total and available stations over Mumbai region is given in Figure 8. Since the point-wise comparison of rainfall between observations and model grids are difficult, we have averaged the observations over a fixed circular area with the radius of ~9 km. The radius is chosen in such a way that the diameter of the circles (solid gray circles in Figure 9) coincided with the NCUMG grid resolution (~18 km). A total of six circles covered the Mumbai region, and in each circle there are ~15-20 raingauge stations available. However, the stations number decreased to ~10 from the last week of July due to non-availability of daily data from

MCGM stations over Mumbai region. Figure 9 represent the typical spatial distribution of raingauge observations (black solid squares) along with model grids. NCMRWF regional (NCUMR) model grids are represented by red "+" and blue color alphabets (A, B, C, D, E) shows the global model (NCUMG) grids, respectively.

Since the stations available at each circle varied with time, we have carried out the model verification and computed standard verification scores for each region with the available station data. Region wise correlation and bias (Figure 10) shows that over the regions "B" and "C" the model biases are relatively less due to the large density of the stations. As mentioned above, over the selected circles, covering most of the Mumbai region, the rainfall amounts from observations and model forecasts are averaged and used for further analysis. Here we have used standard and categorical verification methods (False Alarm Ratio, FAR and ETS etc.) to assess skill of the global and regional models. Note that, 3-point running mean is applied to remove the noise for both forecasts and observations.

Despite the mismatch in timing and magnitude, time series comparison between observed and NCUMR forecast rainfall shows a reasonable agreement in Day 2 compared to Day 1 and Day 3. It is seen that NCUMG overestimate the rainfall amounts specifically over A and B grids and the over estimation is more prominently seen from second week of July onwards (Figure 11). On the other hand, NCUMR is closer to the observations in all the grids and forecast days. However, overestimation of rainfall is also seen in regional model forecasts during the heavy rains that occurred in the 04-12 July 2020. Notably there are few occasions where both the models exhibit large underestimations. To better represent the discrepancy in the model forecasts weekly bias values computed over the 5 grids are shown in Figure 12. Quantitatively rainfall bias is relatively large in NCUMG compared to NCUMR over most of the grids. Lack of consistency in the

NCUMG rainfall is clearly noticed from Figure 4. Weekly biases are slightly lower in NCUMR Day 2 forecast compared to the other forecast times. The association between the observations and forecast rainfall is further examined by linear correlation analysis (top panel in Figure 13) and the coefficients are shown in Table 1. It is clear from Figure 13 and Table 1 that the regional model rainfall forecast is better correlated with observations compared to NCUMG. The correlations are reduced by ~50% in NCUMG compared to NCUMR forecast times, indicating the strength of high resolution of regional models in better simulating the rainfall.

Statistical verification of the rainfall forecasts from NCUMG and NCUMR are performed at different rainfall threshold from 0.05mm to 150 mm with increment of 10mm by computing categorical scores ETS and FAR (Figure 13). The performances of the forecast are evaluated over JJAS season. Despite of having the biases in the rainfall forecasts, the ETS values are greater than zero for all the thresholds indicating the effectiveness of the forecasts. Implies, while the model under/overestimates the rainfall area, it correctly predicts its position. The ETS computed here is ranging between 0.0 to 0.4 upto 50mm threshold then it decreases sharply to zero 100mm threshold in NCUMR. In contrast, in NCUMG the ETS values are peaking at ~25mm threshold and becoming zero at around 75mm threshold and the ETS values are decreasing with lead times. And this feature is seen in both Day-1 and Day-2 NCUMR forecasts. Another notable feature is that the regional model exhibits higher ETS values in all the lead times. FAR also indicates that overall, the performance of NCUMR is relatively better than NCUMG. However, close examination depicts that for low rainfall amounts (less than 10mm) threshold FAR in NCUMR Day-1 forecast is slightly larger than global. On the other hand, for higher rainfall amounts ~75-100mm, the FAR is much less in NCUMR in all the forecast days indicating the model performance is better in simulating the heavy rainfall events with greater skill.

Lastly the biases found in both global and regional model forecasts indicate a close examination of the other key meteorological variables like winds and fluxes. Also, most of these extreme rainfall events are largely governed by the synoptic scale variability, one needs to examine the representation of the synoptic systems in the model forecasts for better understanding and prediction of these events.

4. Summary and Conclusions

a) The aim of this report is to evaluate the performance of NCMRWF global (NCUMG) and regional (NCUMR) for the integrated flood warning system (iFlows) over Mumbai region during monsoon season (June to September, JJAS) 2020.

b) Rainfall scenario over Mumbai exhibits 4 extreme rainfall (ER) (>200mm/day) events.

c) Rainfall maxima observed with the gridded observations are shifted to westward in the NCUMG and it is more intense in Day-2 forecast. In contrast the rainfall maximum is large in Day-1 forecast in the regional model and with lead time reduction in rainfall magnitudes are seen.

d) Verification (mean, mean error and spatial correlation) with satellite gauge merged data are showing considerable discrepancies, indicating the limitations in using the gridded observations to assess the model performance.

e) Although, global and regional models depict significant differences in predicting the seasonal rainfall, models performance in representing the magnitude and spatial distribution of rainfall for extreme rainfall events (>200mm/day) are reasonably well.

f) The categorical verification methods (FAR and ETS) showed a better skill in NCUMR compared to NCUMG.

5. References

1. Sumit Kumar, 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/01/2018, 17p. (https://www.ncmrwf.gov.in/NCUM-Report-Aug2018_final.pdf).

2. Mitra A. K., A. K. Bohra, M. N. Rajeevan and T. N. Krishnamurti, 2009: Daily Indian precipitation analysis formed from a merge of rain-gauge data with the TRMM TMPA satellite-derived rainfall estimates. *J. Meteorol. Soc. Jpn*. 87A, 265-279.
3. Gadgil, S. 2003: The Indian monsoon and its variability. *Annu. Rev. Earth Planet. Sci.*, 31, 429-467.
4. Webster, P. J., V. O. Magna, T. N. Palmer, J. Shukla, and R. A. Tomas 1998: Monsoons: processes, predictability and the prospects for prediction, *J. Geophys. Res.*, 103, 4451-4510.
5. Rao, Y. P.,1976: Southwest Monsoon. Met. Monograph, India Meteorological Department, 367.
6. Francis, P., Gadgil, S. 2006: Intense rainfall events over the west coast of India. *Meteorol. Atmos. Phys.* **94,** 27–42 https://doi.org/10.1007/s00703-005-0167-2.
7. Pattanaik, D.R. and Rajeevan, M 2010: Variability of extreme rainfall events over India during southwest monsoon season. *Meteorol. Appl*. 17: 88–104.

8. Sumit Kumar, 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/01/2018, 17p. (https://www.ncmrwf.gov.in/NCUM-Report-Aug2018_final.pdf).
9. Abhishek, L., Routray, A., Dutta, D. George, J.P. and Rajagopal, E.N. 2020: A Regional Land Data Assimilation System for NCUM-R. NMRF/TR/07/2020, Internal Report, National Centre for Medium Range Weather Forecasting, Ministry of Earth Sciences, Government of India, 18p.
10. Monsoon Monograph -II, India Meteorological Department 2012: https://imetsociety.org/wp-content/pdf/docs/MM2.pdf.

Appendix - I

Table1: Correlations values over Mumbai region w.r.t observations

Forecast day	**Obs vs NCUM-G**	**Obs vs NCUM-R**
Day – 1	**0.31**	**0.56**
Day – 2	**0.26**	**0.63**
Day - 3	**0.21**	**0.54**

Appendix -II

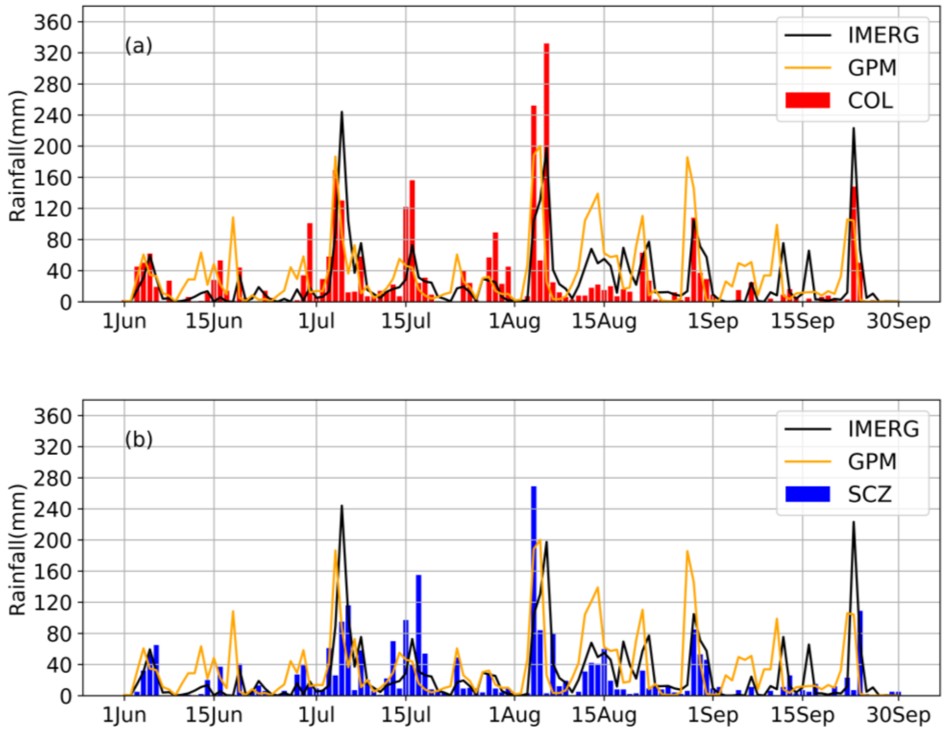

Figure 1: Time series of daily rainfall over (a) Colaba (CLB, red bars) and (b) Santa Cruz (SCZ, blue bars) along with IMERG and GPM rainfall during the monsoon 2020.

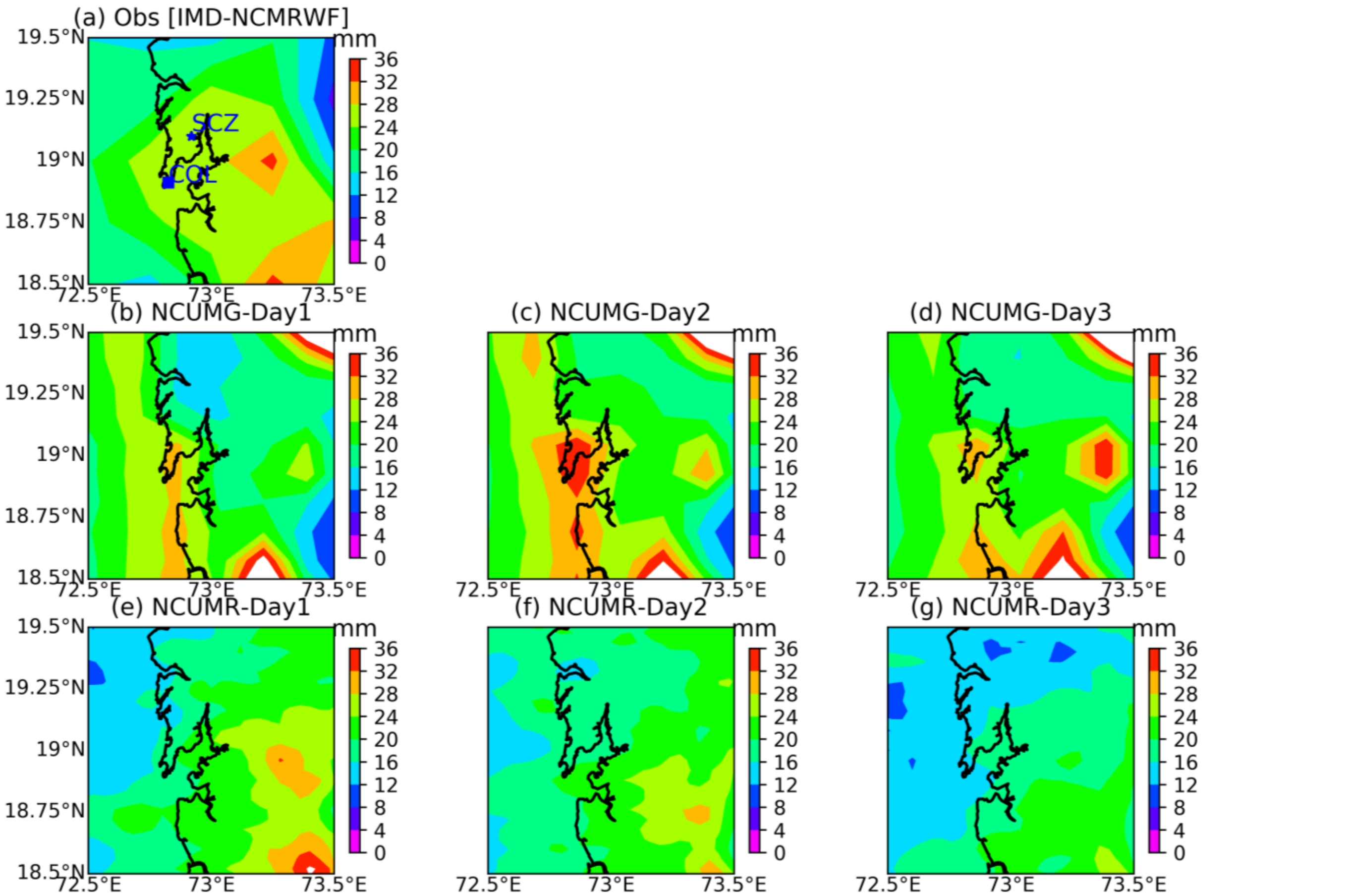

Figure 2: Spatial map of JJAS mean rainfall (mm/day) from IMD (gauge+satellite) merged daily rainfall (top panel). NCUM Global (middle panel), and Regional (bottom panel) upto Day-3 forecasts.

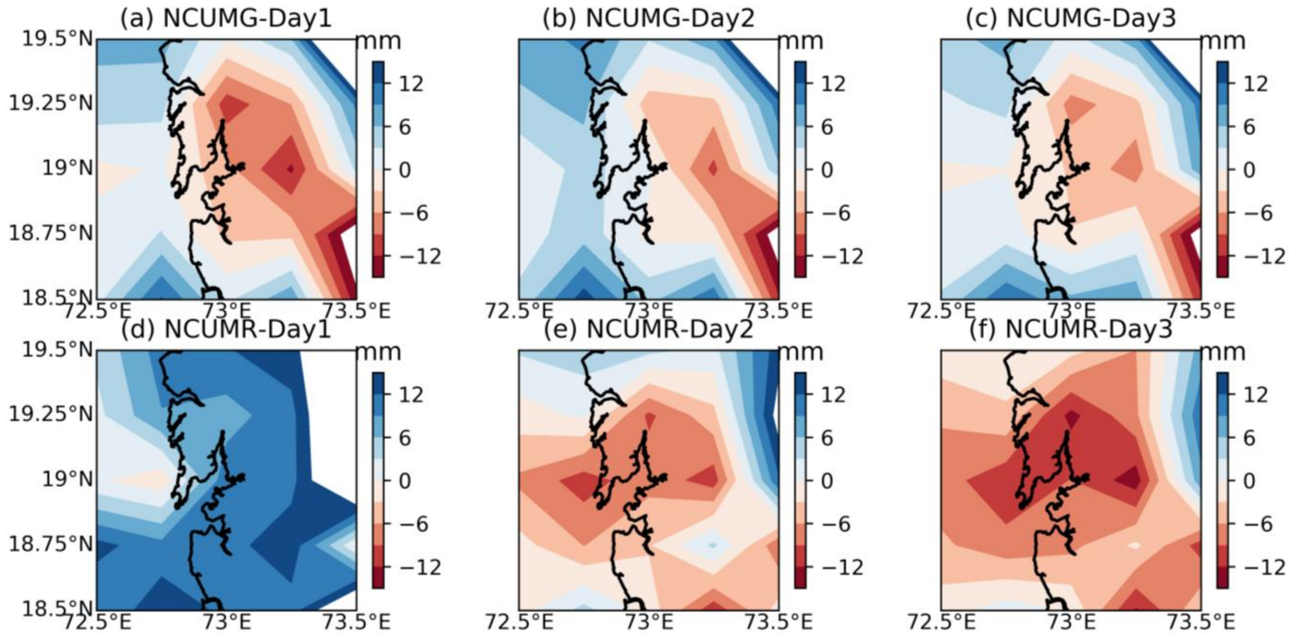

Figure 3: Mean error (ME) for JJAS 2020 season computed from the gridded observations over Mumbai region for NCUMG (top panel) and NCUMR (bottom panel) for the rainfall forecasts upto Day-3.

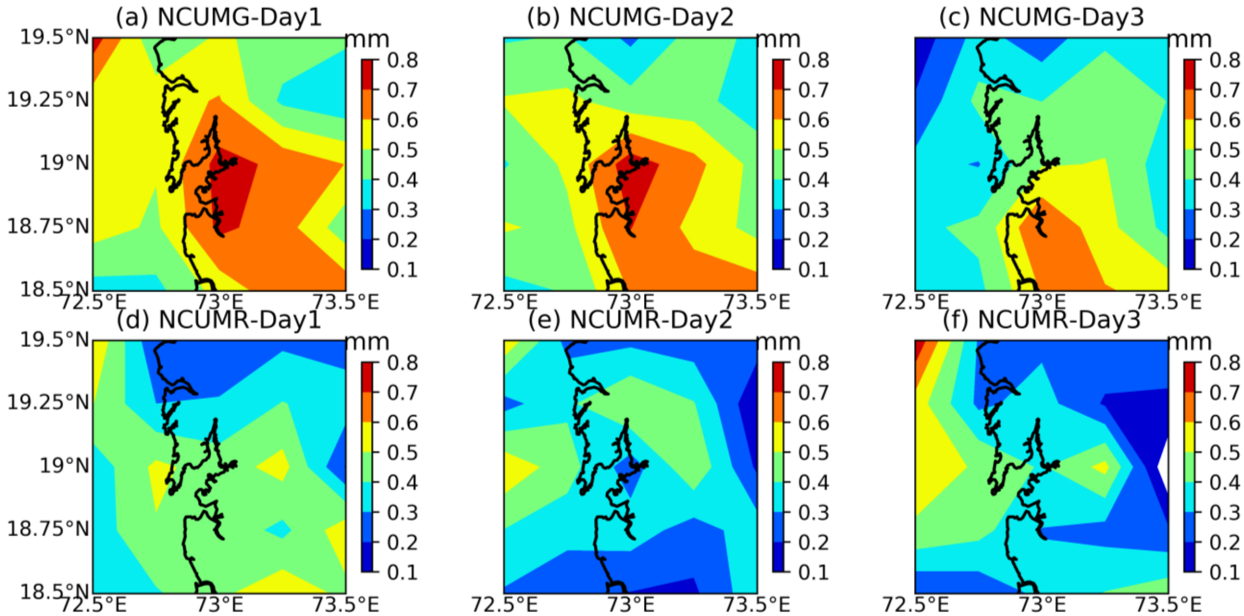

Figure 4: Same as Figure 2, but for the correlation during JJAS 2020 season.

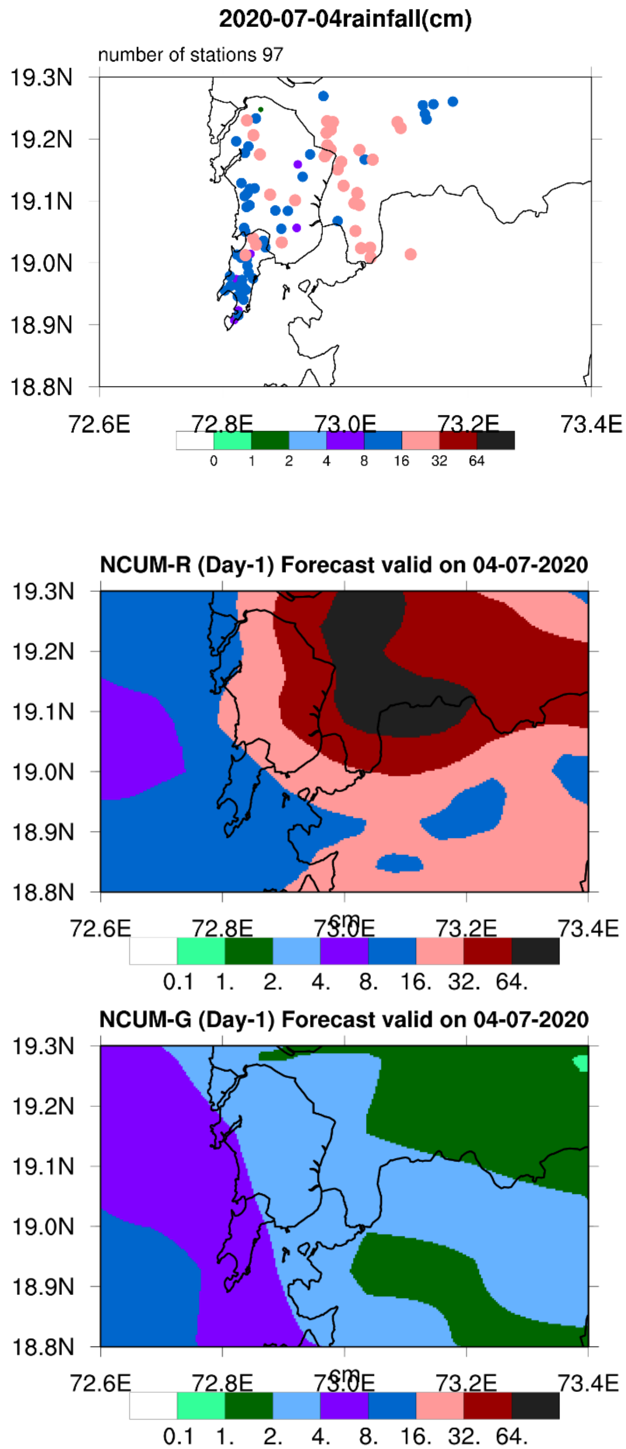

Figure 5: Typical spatial distribution of rainfall event that occurred over Mumbai region on 4th July 2020 from observations (Top panel) and Day-1 forecasts of NCUMR (middle panel) and NCUMG (bottom panel).

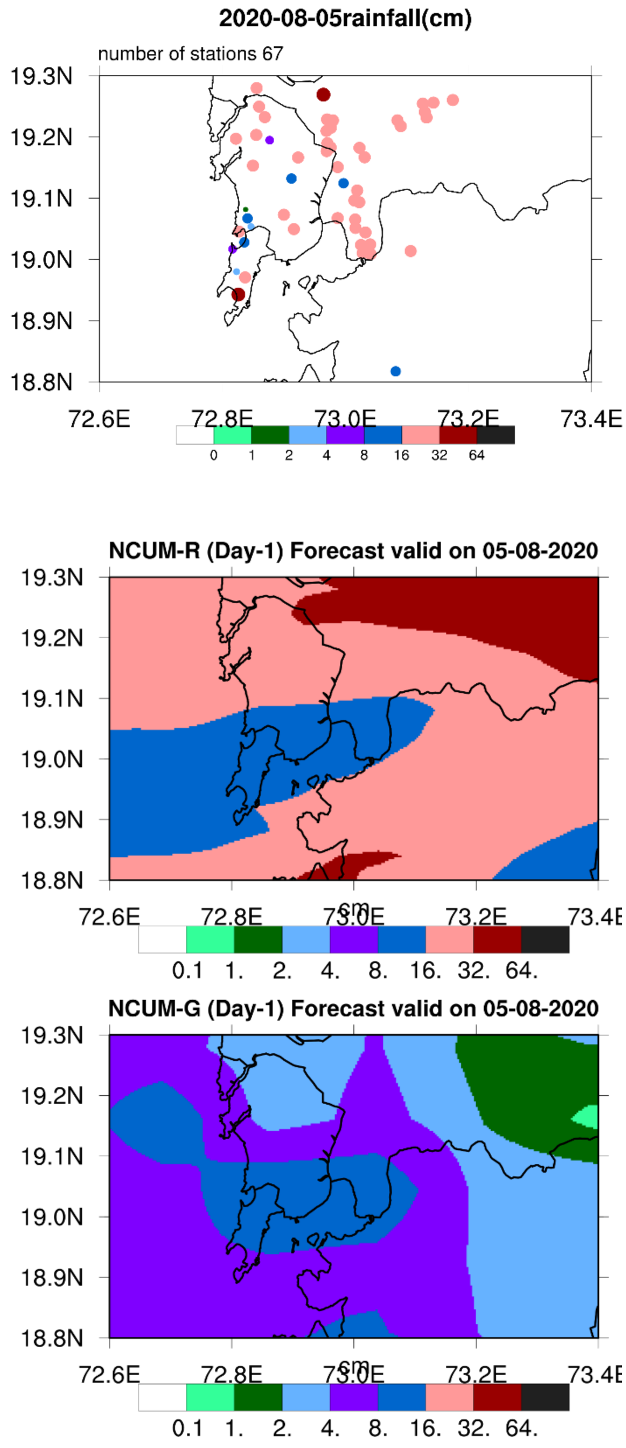

Figure 6: Same as Figure 5, but for the extreme rainfall event during 5th August 2020

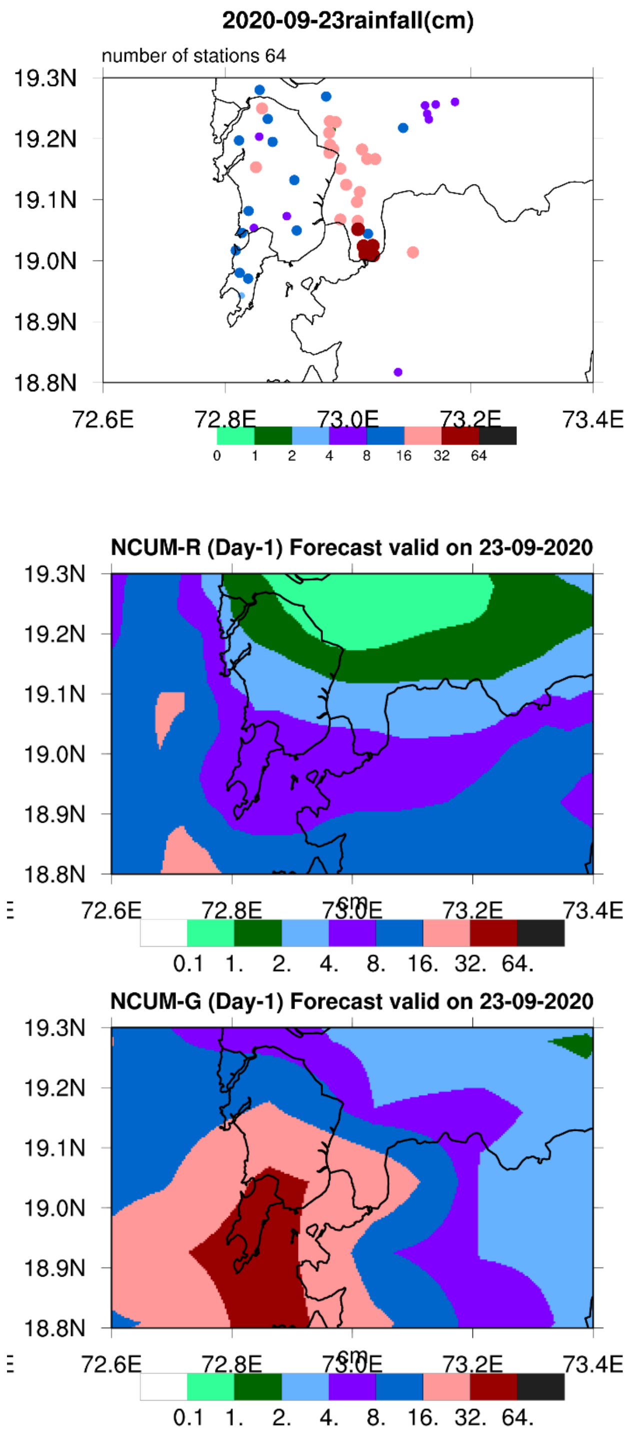

Figure 7: Same as Figure 5, but for the event on 23rd September 2020

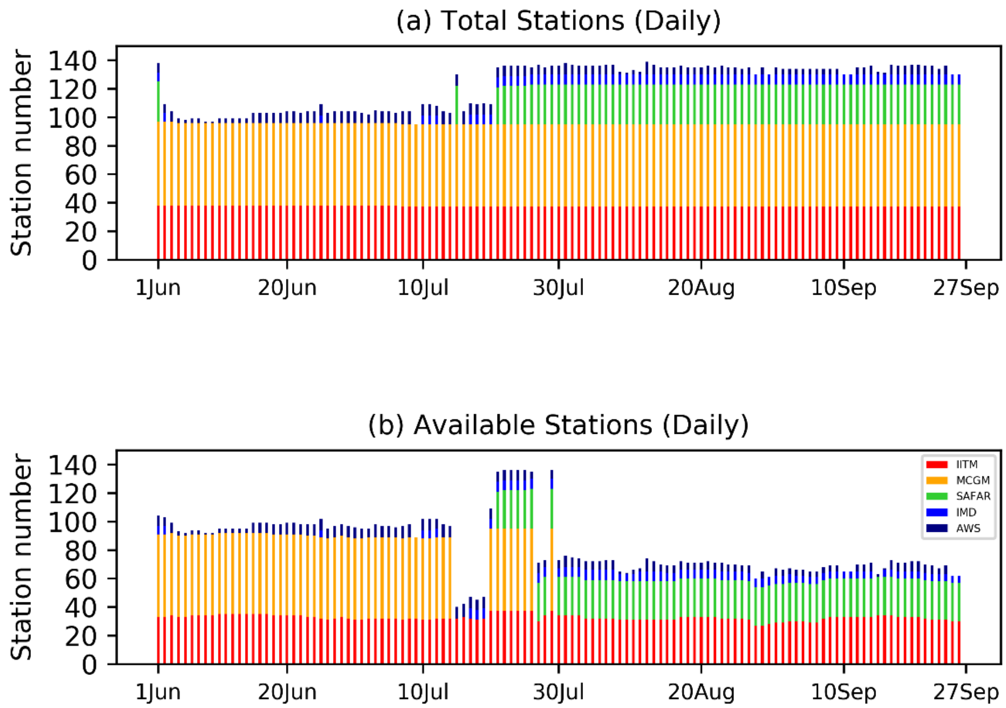

Figure 8: Time series of (a) total number of raingauge stations and the available stations during the JJAS 2020. Different colors on the stacked bar indicates the various organizations or sources maintaining the surface observations.

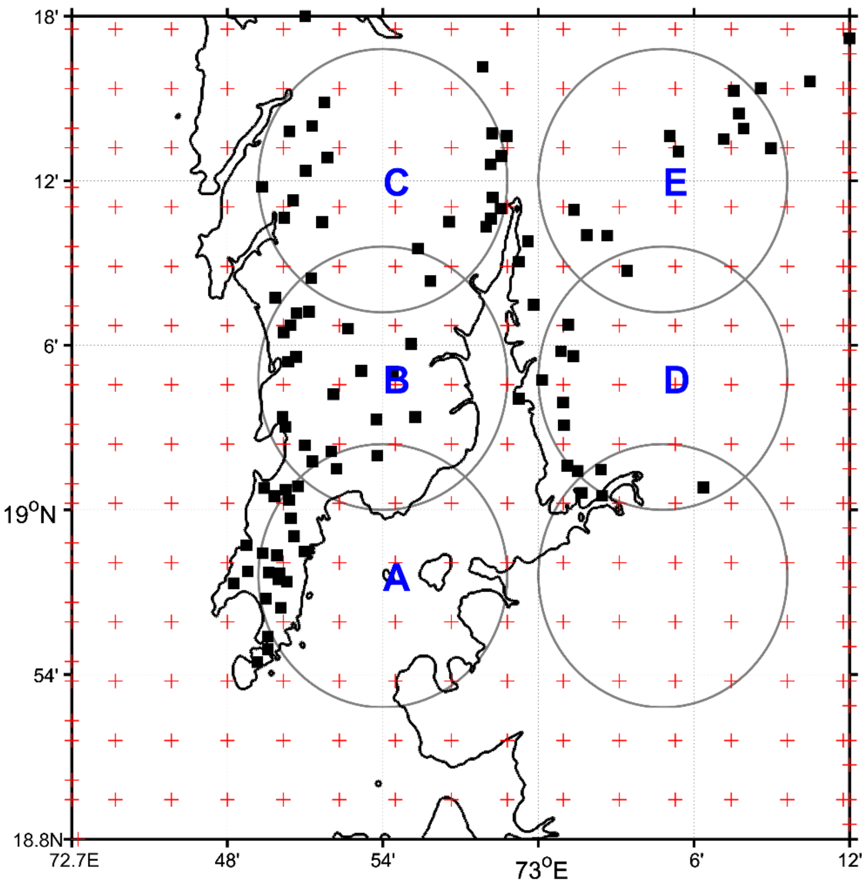

Figure 9: A typical spatial distribution of observations (black solid squares) along with NCUMR model grids (red "+"sign) and NCUMG (blue colored alphabets). The gray circles in the figure indicate the areas considered for model verification. The circle has radius of ~9km.

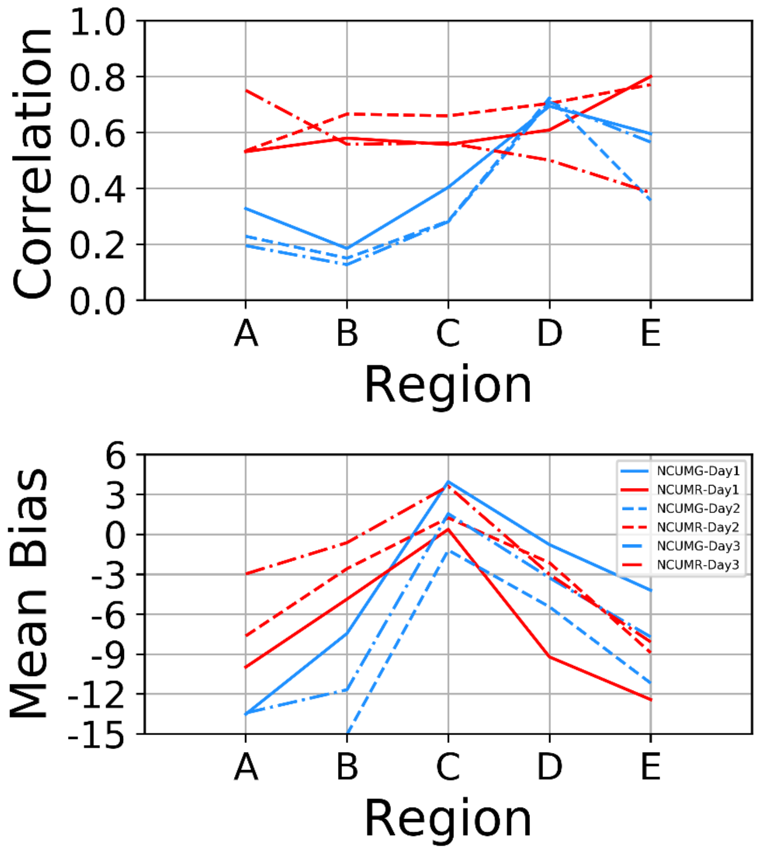

Figure 10: Region wise Correlation and mean bias (mm) of the rainfall forecast against the station observations during June to September 2020 over Mumbai. X-axis is the region names (see section 2).

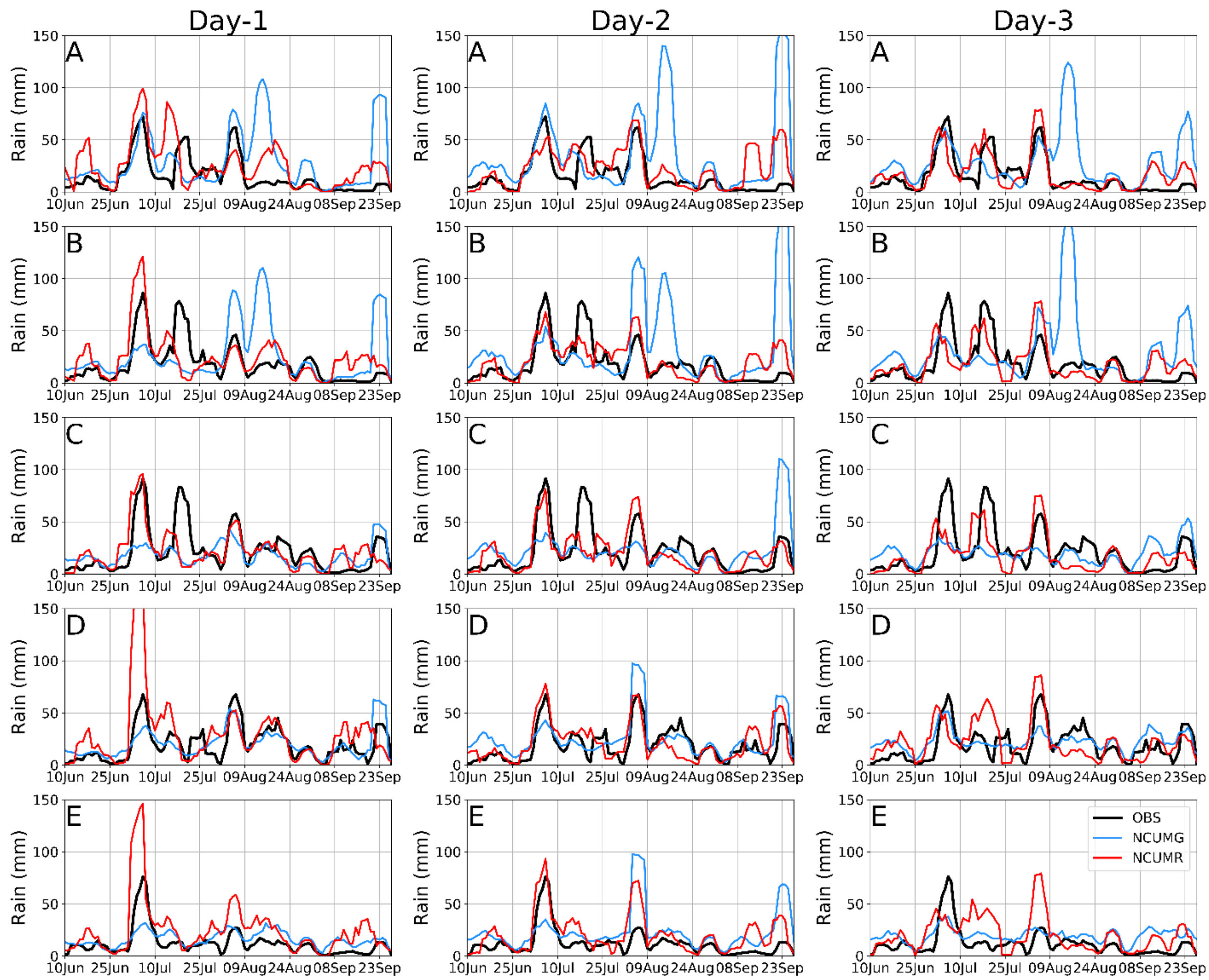

Figure 11: Time series of NCUMG (blue) and NCUMR (magenta) forecasted rainfall along with observations (gray) during the Day-1 (left panel), day-2 (middle) and Day-3 (right panel) during June through September 2020. The alphabets indicated in the top left corner of the each subplot indicates the region names in Figure (9).

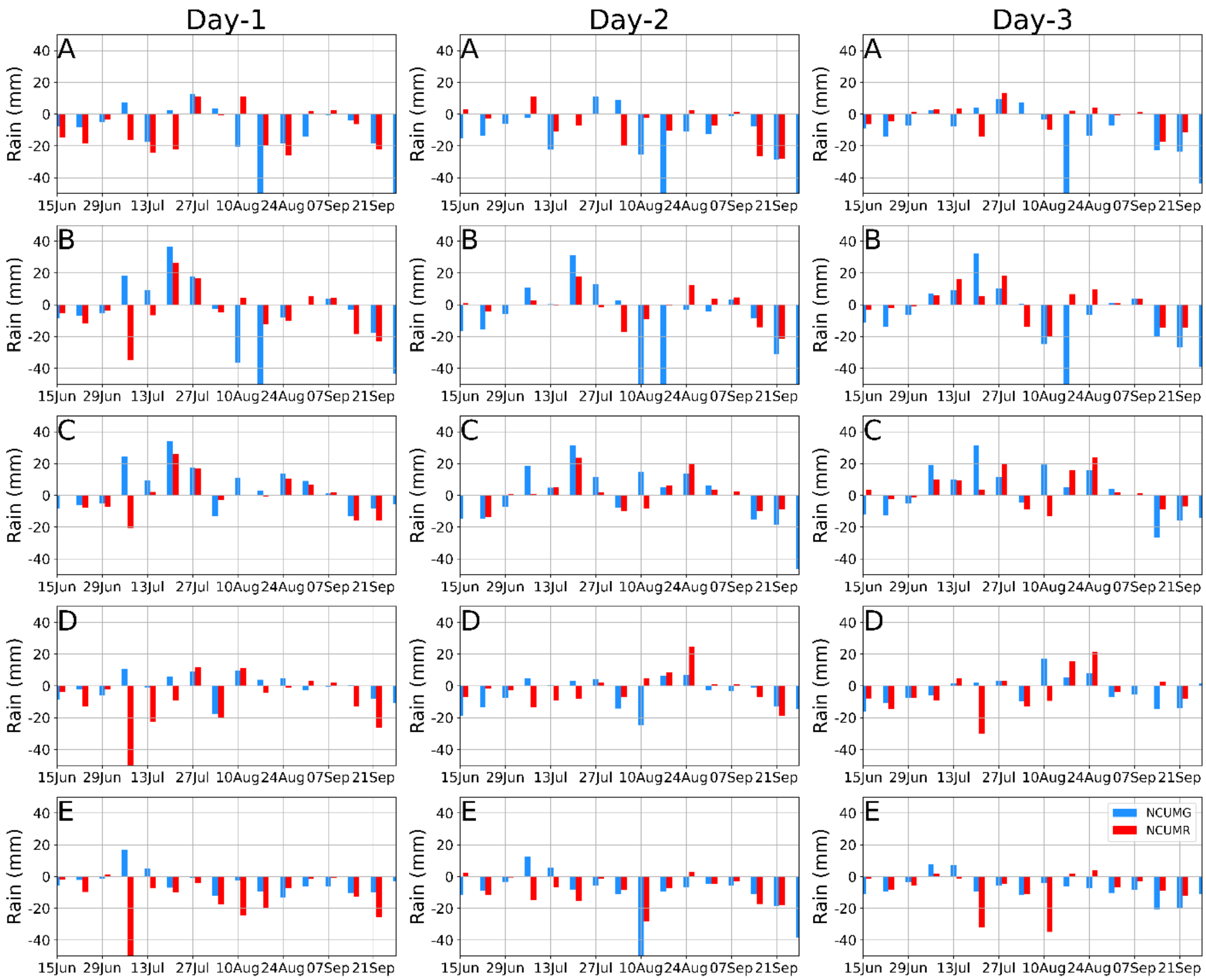

Figure 12: Time series of NCUMG (blue) and NCUMR (magenta) rainfall bias (in mm) w.r.t surface raingauge observations during the Day-1 (left panel), Day-2 (middle) and Day-3 (right panel) during June through September 2020. The alphabets showed in the top left corner of the each subplot indicates the region names. (as of Figure (9)).

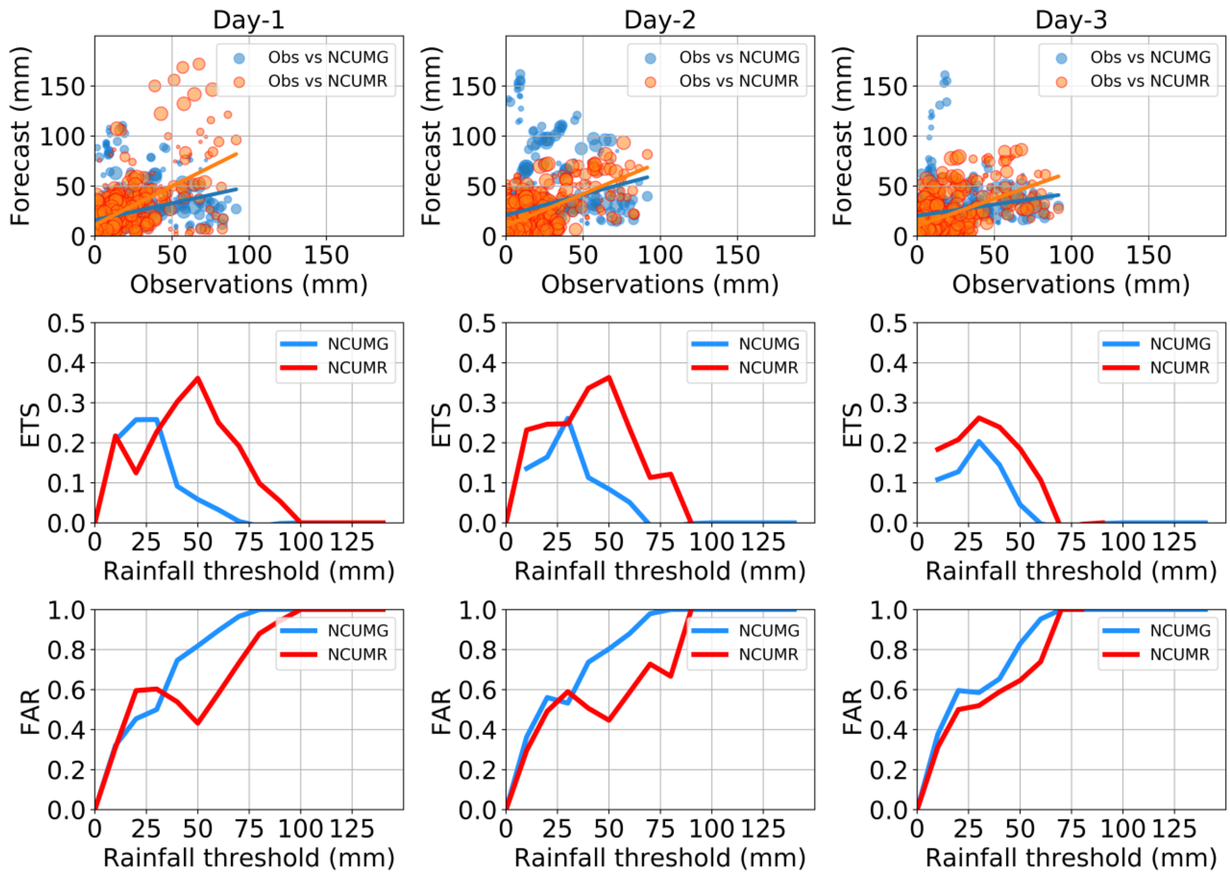

Figure 13: Top panel indicate the correlations of NCUMG and NCUMR with observations. Solid lines in the top panel indicate the best fits of global and regional models forecasts. Middle and bottom panel shows the ETS and FAR values computed for different rainfall thresholds for global and regional rainfall forecasts.

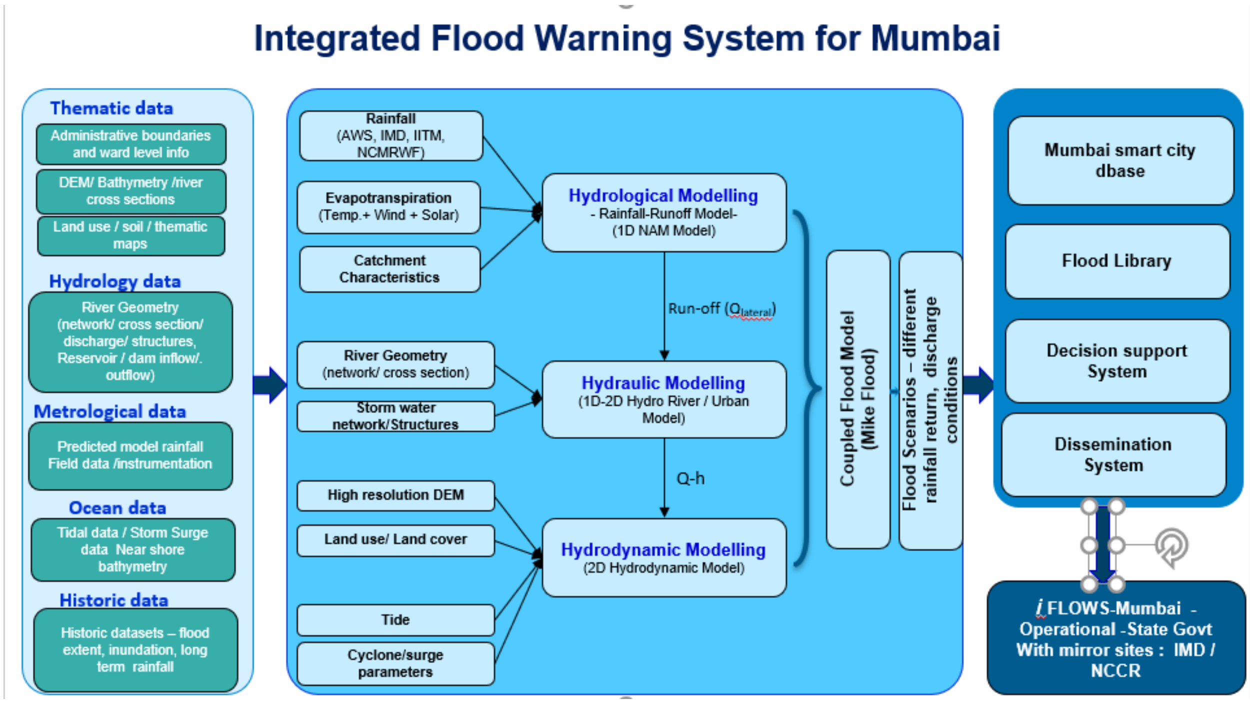

Figure S1: flow chart describing various components for the iFlows flood warning system for Mumbai