NMRF/RR/01/2021

सत्यमेव जयते

RESEARCH REPORT

# Monitoring the Quality of AWS/ARG Rainfall Observations over India during Monsoon Season

Upal Saha, M. Das Gupta and Ashis K. Mitra

February 2021

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences, Government of India
A-50, Sector-62, NOIDA-201309, INDIA

# Monitoring the Quality of AWS/ARG Rainfall Observations over India during Monsoon Season

**Upal Saha, M. Das Gupta and Ashis K. Mitra**

**February 2021**

**National Centre for Medium Range Weather Forecasting**

**Ministry of Earth Sciences**

**A-50, Sector 62, NOIDA-201309, 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/01/2021 |
| 3 | Date of publication | February 2021 |
| 4 | Title of the document | Monitoring the Quality of AWS/ARG Rainfall Observation over India during Monsoon Season |
| 5 | Type of Document | Research Report |
| 6 | No. of pages & Figures | 52 & 25 |
| 7 | Number of References | 28 |
| 8 | Author (s) | Upal Saha, M. Das Gupta and Ashis K. Mitra |
| 9 | Originating Unit | NCMRWF |
| 10 | Abstract | Deployment and maintenance of the observing systems is one of the most useful component of weather and climate monitoring and prediction system. In-situ observations are considered as important component of the observing system and the same is often used to validate the remotely-sensed counterpart of the same. With the advent of technology and the growing demand of high density observations, various automatic components are being introduced in observing technology of in-situ observations. In India, with the modernisation of India Meteorological Department (IMD) during 2010, apart from IMD's established operational rain gauge network, a new network of in-situ surface observations, comprising of automatic weather stations (AWS) and automatic rain gauges (ARG) have become operational. Many studies have shown the beneficial usage of these observed surface parameters, viz. pressure, temperature etc. from this network in now-casting and monitoring of synoptic scale weather systems. However, the studies on the quality of these observed rainfall from this network is limited.<br>In this study an attempt has been made to validate 24-hourly accumulated rainfall of these stations against available neighbouring in-situ (SYNOP) and gridded (in-situ and satellite–gauge merged) observations for July-August of 2018 to 2020. Based on this procedure, a real-time monitoring method has been developed to assign quality flags to each station based on last 15-days validation for its possible utilisation in operation and research purpose by many end users. |
| 11 | Security classification | Non-Secure |
| 12 | Distribution | Unrestricted Distribution |
| 13 | Key Words | Automatic weather station; automatic rain gauge; monsoon; time-series. |

# Contents

## Abstract

Deployment and maintenance of the observing systems is one of the most useful component of weather and climate monitoring and prediction system. In-situ observations are considered as important component of the observing system and the same is often used to validate the remotely-sensed counterpart of the same. With the advent of technology and the growing demand of high density observations, various automatic components are being introduced in observing technology of in-situ observations. In India, with the modernisation of India Meteorological Department (IMD) during 2010, apart from IMD's established operational rain gauge network, a new network of in-situ surface observations, comprising of automatic weather stations (AWS) and automatic rain gauges (ARG) have become operational. Many studies have shown the beneficial usage of these observed surface parameters, viz. pressure, temperature etc. from this network in now-casting and monitoring of synoptic scale weather systems. However, the studies on the quality of these observed rainfall from this network is limited.

In this study an attempt has been made to validate 24-hourly accumulated rainfall of these stations against available neighbouring in-situ (SYNOP) and gridded (in-situ and satellite–gauge merged) observations for July-August of 2018 to 2020. Based on this procedure, a real-time monitoring method has been developed to assign quality flags to each station based on last 15-days validation for its possible utilisation in operation and research purpose by many end users.

## 1. Introduction

With the continuous progress in science and advent of advanced technology, the quality of weather analysis and forecasts, along its usage in various sectors has grown tremendously in last few decades. At present, continuous efforts and emphasis are on to observe and predict localised weather phenomenon more precisely. Accurate analysis and prediction of spatial rainfall patterns exert a key control on the applications to various sectors namely hydrology, agriculture etc. Meteorological observations are the back bone of the real-time weather analyses, forecasts as well as severe weather caveats. Present day's global observing system comprises of in-situ observations along with space and land based remotely sensed observations. The remotely sensed observations (satellite and radar) have large areal coverage and also provide very high density observations. However, due to its associated uncertainty, still in-situ observations are required to validate and rectify the same (Simpson and Jones, 2014; Haiden et al., 2018; Saha et al., 2020). Although, the present days manned surface observational network is adequate for monitoring synoptic scale weather phenomena but the same is not true for meso-scale systems and their variability. Automatic weather observations from surface stations (AWS/ARG) had promised varied applications in operational meteorology such as agro-meteorology (McNew et al., 1991; Hubbard et al., 1983), flash flood forecasting (McCulloch and Strangeways, 1966) and NWP models, etc.

A network of 125 automatic weather stations (AWS) all over India was established by IMD before 2006, mainly along the coastal region for monitoring tropical cyclones. IMD also has an established network of rain gauge stations which are operationally used for analysing rainfall over the country. Apart from these, during 2010, an impressive new network of 550 automatic weather stations (AWS) and 1350 automatic rain-gauges (ARG) was conceived by India IMD, under its modernisation plan - Phase-I. It was planned that each district of India will have one AWS and two ARGs to monitor localised weather systems and by the end of 2012 a network of 541 AWS and 557 ARGs was installed over various parts of India (Ranalkar et al., 2012, 2014, 2015). The observations from this new network has been monitored and validated in this present study and is referred as "New IMD-AWS/ARG network" hereafter.

These automatic stations are equipped with tipping bucket rain-gauge for measuring rainfall , with collector diameter of rain gauge as 20 cm and bucket is calibrated to tip when 0.5 mm of rain water collected in it, with both hourly (reset at each UTC) and daily rainfall counter set (reset at 03 UTC everyday) (Ranalkar et al., 2012). The observations recorded by each stations are transmitted to receiving earth stations (ES) via satellite (INSAT). At earth stations, these observations are encoded in WMO SYNOP-MOBIL code and transmitted globally via global telecommunication system (GTS) for its further utilisation. However, for the appropriate operation of the network, it is very much necessary for a periodic maintenance of the AWS & ARG stations, including sensor checks and calibrations, as well as validation of collected data (Estévez et al., 2011).

Quality control of rain gauge data has been an important topic since the beginning of manual rainfall data collection through tipping bucket rain gauge (Michaelides, 2008). Ranalkar et al. (2015) developed a quality control procedure for AWS observations, mainly comprising of climatological check and internal consistency check. As suggested by Zahumenský (2004), the gross error checks on raw and processed data, time consistency check involving single parameter and internal consistency check involving two parameters are the QC procedures adopted while processing AWS/ARG data in real time. However, a very important check, the horizontal consistency check, popularly known as "Buddy Check" is not generally applied at the station level quality control method. As an objective analysis of observations is required for Buddy Check, so generally this check is performed at global data-processing centres before utilisation of these observations. As precipitation is a highly discrete phenomenon in space and time scale, horizontal consistency is also not achieved always and so other methods are tried to validate these observations. Lewis et al. (2018) compared the daily accumulated gauge data to the high-resolution gridded daily dataset to estimate the initial quality of the gauge data over Great Britain region.

Several studies (Mohapatra et al., 2011; 2014, 2015; Ray et al., 2015) have shown the beneficial usage of Indian AWS observations, specially wind and pressure in monitoring of synoptic scale tropical cyclones and now-casting of meso-scale systems, however, there are not many studies on the utilisation of observed AWS and ARG rainfall for operational as well as research purpose. This may be due to the limited studies on validation and the quality of observed AWS and ARG rainfall. In this study an attempt has been made to validate AWS, ARG rainfall observations from this "New-IMD AWS/ARG network" against neighbouring in-situ SYNOP observations as well as gridded rainfall data sets. Based on the validation results, a method has been devised to delineate good quality AWS/ARG rainfall observations for its further utilisation in real time operation as well as for validation of NWP model outputs. As in recent years many states of India are developing high-density meso-net observing network comprising of AWS and ARGs, the methodology developed through this study may be useful for judicially using the rainfall observations of these stations for weather monitoring and validation of NWP outputs.

## 2. Data and methodology

Daily rainfall observations (24 hourly accumulated from previous day 0300 UTC to present day 0300 UTC) of stations from the "New-IMD AWS/ARG network" has been validated against neighboring SYNOP observations and two gridded (one in-situ and other satellite-gauge merged) rainfall datasets over Indian region for July-August of 2018, 2019 and 2020 respectively. As rainfall is generally fairly well distributed over major parts of the country during the core monsoon months (July-August), the rainfall validation during core monsoon months has been carried out for this study.

### 2.1 *In-situ Observations*:

Indian SYNOP, AWS and ARG observations are transmitted to NCMRWF through a dedicated FTP system in real-time by Regional Telecommunication Hub (RTH) of IMD. These observations are then decoded and archived in proper data-base for its further utilization in NWP system. Indian SYNOP stations generally reports observations at a frequency of 3hours, eight times a day (00, 03, 06, 09, 12, 15, 18 and 21 UTC), however maximum number of stations (~300 /day) are being received at 0300 and 1200 UTC. The coverage of SYNOP observations on a typical day of August, 2018, received at NCMRWF is shown in Figure 1.

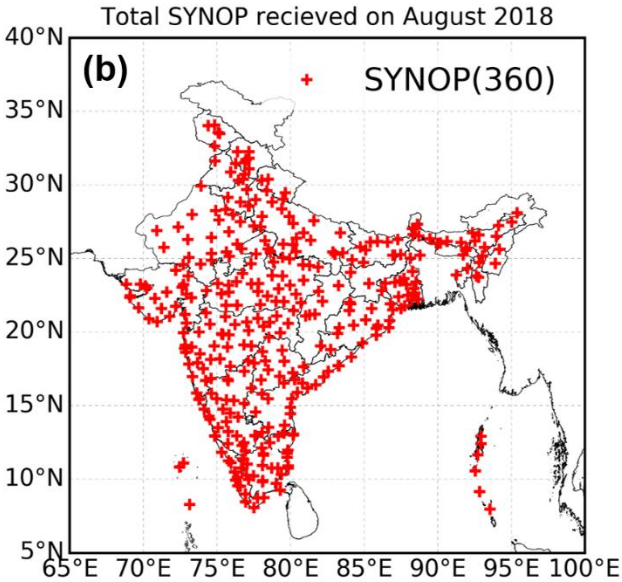


**Fig. 1.** Coverage of SYNOP observations received at NCMRWF on a typical day of August, 2018.

AWS and ARG stations from this new network report observations with a frequency of one hour, however the count of daily reception of these hourly observations are highly variable with average maximum reception during the day time. The reception of all these observations are being monitored at NCMRWF in real-time (Das Gupta and Rani, 2010; Singh et al., 2018). The average number of AWS and ARG 0003 UTC observations are about 344 and 550 in July-August 2018, which was reduced to 220 and 438 respectively in 2019 and further improved to 288 and 348 respectively in 2020.

### *2.2 Gridded Rainfall Datasets:*

Two gridded rainfall datasets based on gauge-only and satellite-based gridded observations are used for validation of AWS/ARG rainfall observations in this study. First one is the daily gauge-only gridded rainfall dataset at 25 km spatial resolution generated by IMD (Pai et al., 2014). However, the temporal density of the station points was not uniform and on average, about 2600 stations per year with a maximum of 6955 stations were available for the preparation of daily grid-point data (Rajeevan et al., 2006, 2008; Pai et al., 2014). Out of 6955 rain gauge station records, 547 records from IMD observatory stations, 494 records from Hydro-meteorological observatories, 74 records from Agro-met observatories and 5845 records from stations maintained by the state governments (Pai et al., 2014; Singh et al., 2021). This data set has been widely used as a reference rainfall data for the evaluation of satellite-derived rainfall, validation of predicted rainfall by NWP models and various hydro-meteorological applications in India.

The other daily gridded set used in this study is NCMRWF merged satellite gauge (MSG) rainfall product, generated jointly by NCMRWF and IMD (Mitra et al., 2003; Mitra et al., 2009). In this method rainfall is analysed using IMERG (GPM) satellite product as first guess and IMD in-situ gauge discussed above as observations, which corrects the satellite-estimated rainfall. This dataset is also extensively used in several studies (Prasad et al., 2016; Sharma et al., 2017; Sridevi et al., 2020 etc.) to validate the model forecast over Indian region. As these gridded rainfall are used as background data for quality control of AWS/ARG rainfall stations from this new network, so an attempt has been made to inter-compare these gridded datasets for monsoon of 2018-2020.

Figure 2 indicates the spatial variability of the rainfall time-averaged over core monsoon seasons (July-August) during 2018 to 2020 from IMD (first panel) and MSG (second panel) sources respectively. Western Ghats including Kerala, Goa, coastal regions of Karnataka and

Maharashtra received a very good amount of rainfall (> 20 mm day$^{-1}$) during the study period. Orissa, Chattisgarh, eastern Madhya Pradesh, sub-Himalayan West Bengal (Sikkim) in the east, Assam and Meghalaya, Mizoram in the north-east and Uttarakhand, parts of Himachal Pradesh, Haryana and west Uttar Pradesh in the northern India also received pretty good amount of rainfall (≥ 20 mm day$^{-1}$) during this period. Other regions including Jammu and Kashmir, Punjab and Haryana in the north, Rajasthan and Gujarat in the west and Tamil Nadu and Andhra Pradesh in the south experiences rainfall in the range ~2-6 mm day$^{-1}$ during 2018-2020 core monsoon months (Figure 2a-c).

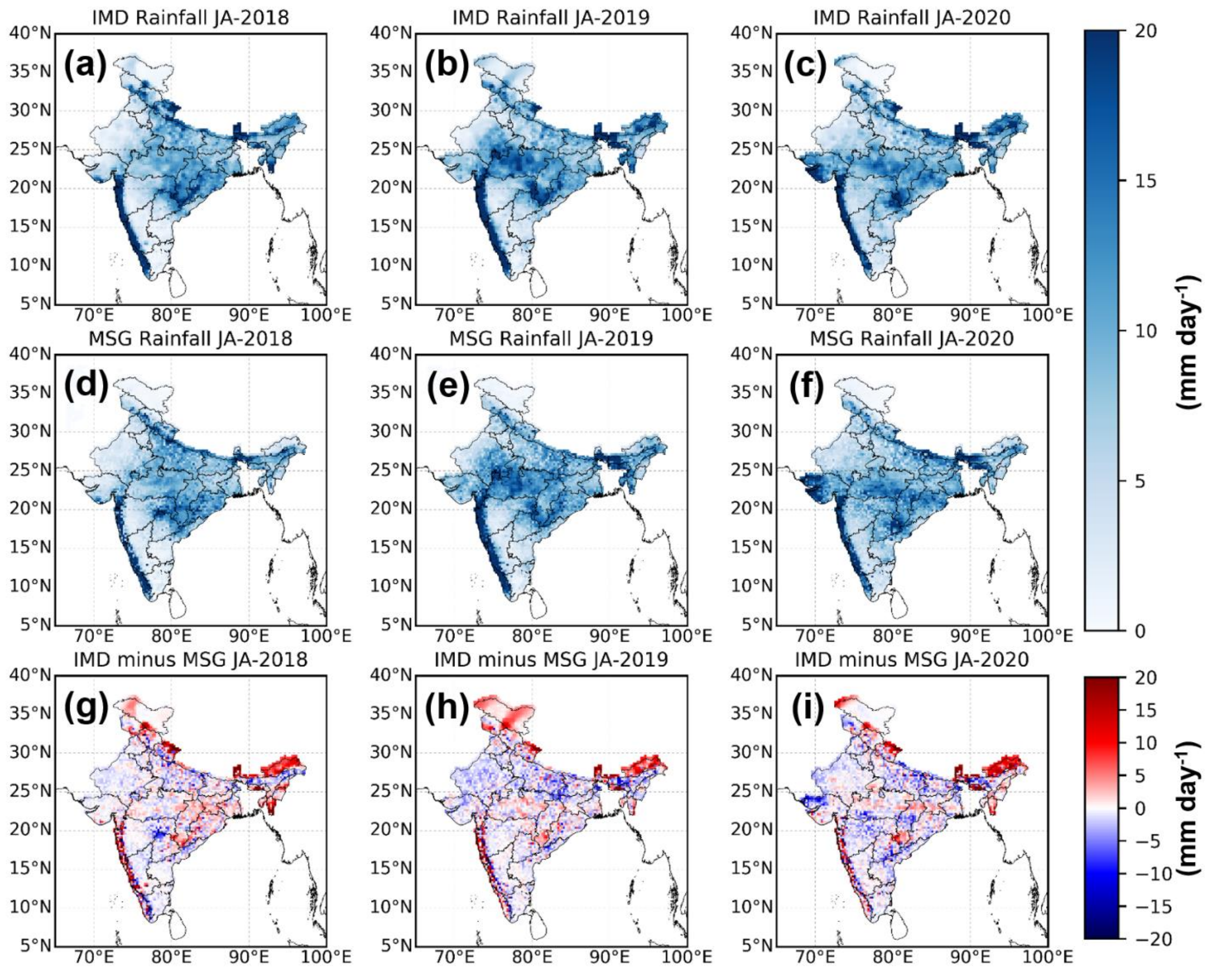


**Fig. 2.** First panel: (a)-(c) Spatial distribution of daily IMD rainfall time-averaged over core monsoon months (July-August) during 2018 to 2020 respectively over Indian landmass. Second panel: (d)-(f) same as before but for MSG rainfall. Third panel: (g)-(i) same as before but for the difference between IMD and MSG rainfall respectively.

Similar to the spatial variability of IMD gridded rainfall, during 2018-2020, MSG too have captured the similar patterns in rainfall distribution over the Indian landmass; however with under-estimation in Jammu and Kashmir, Madhya Pradesh, Assam, and Arunachal Pradesh and over-estimations in Rajasthan, Tamil Nadu and Andhra Pradesh (Figure 2d-f). The difference plots between IMD and MSG rainfall indicates that over the hilly regions (Jammu, Kashmir, Arunachal Pradesh, etc.), Western Ghats and Central North-east homogeneous rainfall regions are having positive biases and rest of Indian sub-continent are having negative biases (Figure 2g-i).

### 2.3. *Methodology*:

Validation of the AWS /ARG 24-hourly accumulated rainfall reported at 03 UTC was carried out against neighbouring SYNOP as well as IMD and MSG gridded rainfall data following the steps discussed below.

(i) As the daily reception of AWS/ARG observations at NCMRWF is highly variable, only those stations with frequency of reported observations at least 75% in a month are used for the validation. These stations are termed as "Regular Station ".

(ii) Out of these regular AWS & ARG stations, some of the stations are having neighbouring SYNOP stations within 50km of it. In case, such neighbouring collocated SYNOP stations exists and those SYNOP stations also report 03 UTC observations for at least 75% of 31 days for individual months of July and August (24 days) in which AWS/ARG observations are available, then AWS & ARG rainfall is validated against that SYNOP observation.

(iii) When collocated SYNOP station does not exist for any AWS/ARG location, then the rainfall reported by the station is validated against IMD and MSG gridded rainfall only. IMD and MSG rainfall is interpolated on the location of AWS/ARG.

(iv) Various statistics e.g. correlation, root mean square error and biases for AWS/ARG rainfall observations are computed against collocated SYNOP, IMD and MSG rainfall data for month.

(v) Monthly time series and scatter coverage plots for AWS/ARG rainfall observations along with collocated SYNOP, IMD and MSG also generated for all regular AWS/ARG stations.

(vi) AWS and ARG stations with collocated SYNOP is said to be “matching” if the correlation of the same with SYNOP $\geq 0.7$.

(vii) For AWS/ARG stations with no collocated SYNOP is said to be “matching” if the correlation of the same with either of IMD or MSG $\geq 0.7$.

Based on the validation results a real time monitoring procedure has been developed for flagging each individual observations based on computed scores over previous 15 days. For flagging purpose a rather strict criteria has been used. Rainfall reported by each AWS/ARG stations is attached with flags ranging from 0 to 9, with 0 as the best. Flag 7-9 are assigned to irregular stations. Detail discussion on flagging criteria will be presented in the later part (Section 4). Depending upon the assigned flags, end user can decide upon which stations to be used for a particular application.

## 3. Validation of AWS and ARG rainfall during monsoon:

In most of the cases, the surface observations like point rainfall from SYNOP/AWS/ARG are typically used as the ‘ground truth’ for evaluating model simulations, weather and climate research, etc since these observations are considered absolutely precise in nature. Since, these observations, are used in NWP for assimilation and verification, quality control (QC) of rain gauge data has been an important topic. Still, the QC of rainfall is particularly a challenging task as it is highly variable in space and time. In several reports, valid observations erroneously identified as invalid “false positives” and its detection and correction is very much needed before it is applied for any NWP verification purposes. Thus, it is important to investigate the coverage (total, regular, matching and non-matching) of AWS and ARG stations over Indian landmass during the core monsoon months of 2018-2020 that are actually reporting for assessing and monitoring the quality of AWS/ARG rainfall data and thereby NWP verification via WMO GTS.

### 3.1. *Coverage plots*:

Out of the functional AWS and ARG stations over India operated by IMD, NCMRWF received about 285 AWS and 446 ARG 03 UTC rainfall observations in average for the study period.

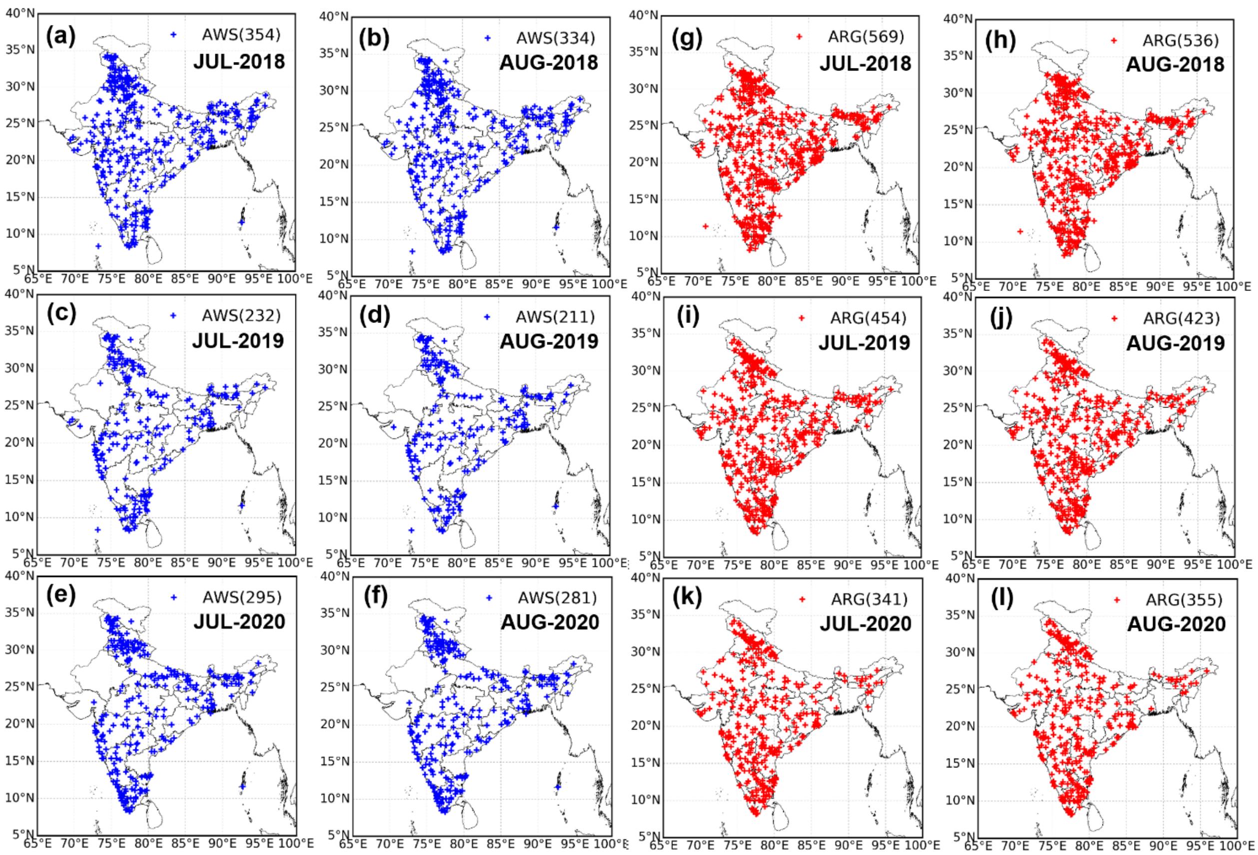


**Fig. 3.** Total coverage of AWS and ARG stations reporting 03Z rainfall received at NCMRWF for July and August 2018- 2020. [Left panel: (a)-(f) AWS stations and Right panel: (g)-(l) ARG stations.]

Figure 3 depicts the coverage of AWS and ARG stations that reported 03 UTC rainfall (accumulated last 24-hours rainfall) for July and August during 2018-2020. The number of stations received for each month are also displayed on the plot. As seen from plots (Figure 3a-f), reception of AWS observations deteriorated in 2019 (~35%) with respect to that of 2018 and however again slightly improved (~ 28%) in 2020. Spatial distribution of AWS in 2018 shows fairly uniform coverage throughout India with good number of stations in the north, over Punjab and adjoining region. However, in 2019 there was less number of stations over the central and north-east India reporting 03 UTC observation. In 2020, marginal increase in the coverage is noticed especially over Kerala. The coverage of ARG stations reduced gradually from 2018 to 2020 with relatively more decrease over north-east and central India.

The daily reception of these observations at NCMRWF are not regular. As the validation of these observations are based on statistical scores, hence the observations only those stations which are reported observations at least 75 % occasions in a month are validated here. The AWS/ARG stations for which observations are received at NCMRWF for more than 24 days (75%) in a month are termed as 'Regular Stations'. The coverage of regular stations for the different months for the study period is depicted in Figure 4. Thus, Figure 4 represents the coverage plot for the counts of AWS and ARG stations that has reported at least 75% of rainfall observations (i.e. 24 days) at 03Z out of 31 days during July and August of 2018-2020 respectively over the Indian landmass. As seen from Figure 4 (a-b), spatial distribution of total AWS count in 2018 (July and August) are showing good coverage of stations over north India and southern peninsular region. The total count of regular AWS observations in July, 2018 was 213 while in August, 2018, the count somewhat increased to 219. Figure 4 (c-d) indicates that the regularly reporting AWS rainfall observations has been reduced in both July and August of 2019 than the previous year. In July, 2019 there are lesser number of AWS regular stations over north, west and south-eastern peninsular Indian region, which is reflecting similar in August, 2019 too.

It was almost 36.1% and 39.3% reduction in regular observations in July and August of 2019 respectively than July and August of 2018. But the regular AWS count in July 2020 has been increased over north, central and southern peninsular region than in July, 2019 (17.6% increased) whereas the regular count of AWS in August 2020 increased to 37.6% than in August, 2019 (Figure 4 e-f).

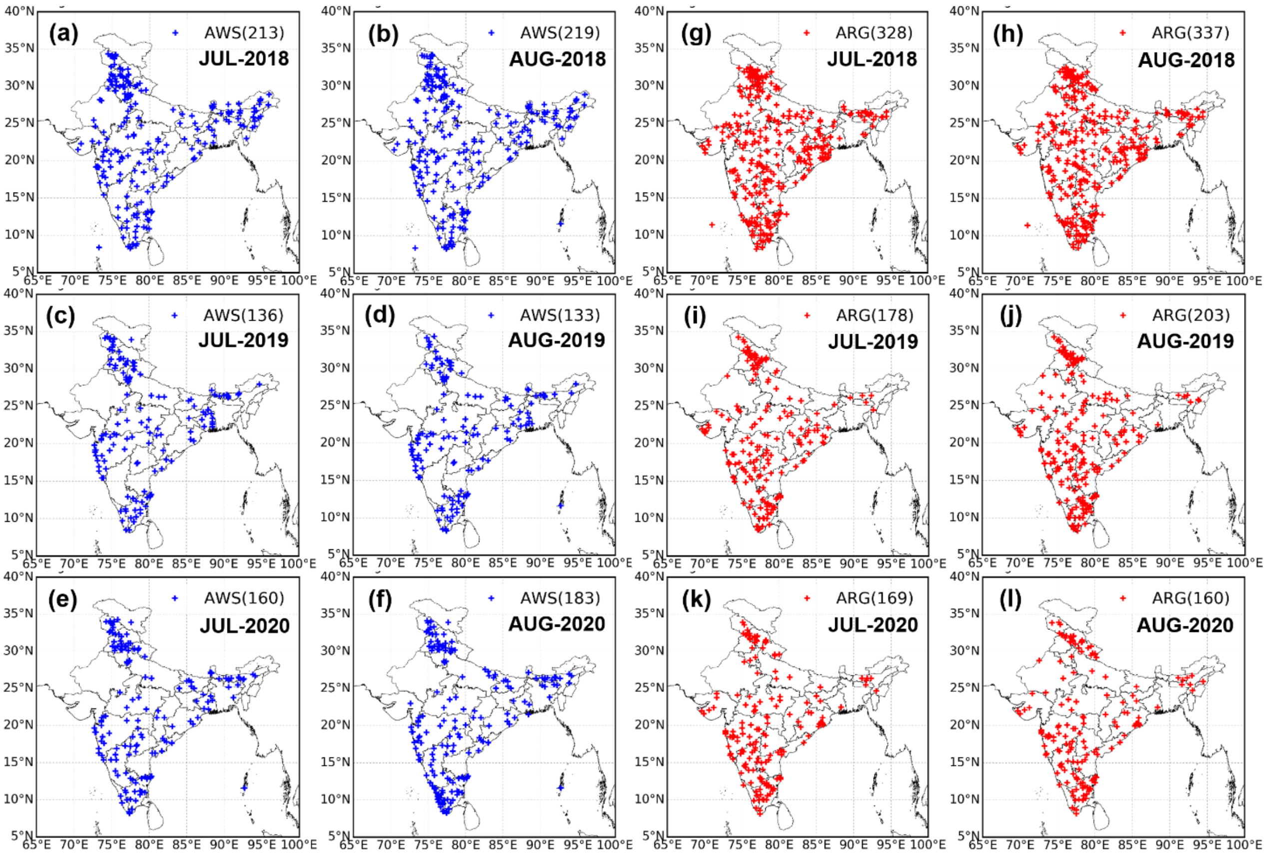


**Fig. 4.** Regular coverage count of AWS and ARG stations reported during (a, g) July 2018, (b, h) August 2018, (c, i) July 2019, (d, j) August 2019, (e, k) July 2020 and (f, l) August 2020 respectively. [Criterion used: 2018-2020: AWS ≥ 24 and ARG ≥ 24 for an individual month of July or August]

Similarly, spatial distribution of regular ARG count in 2018 (July and August) shows good observation coverage in the north, east and southern peninsular region and relatively less clustering in the count over central Indian region (Figure 4 g-h). Figure 4 (i-j) indicates that the regular count of ARG has also been decreased in July 2019 (45.7% than July 2018) and August 2019 (39.8% than August 2018) similar to the regular count reduction in AWS count. In July, 2020 the regular ARG count has been further decreased in small percentage over southern peninsular region, while in August, 2020, the regular ARG count has been decreased than the previous years of 2018 and 2019 (Figure 4 k-l).

In the present analysis, some strict criteria based on correlation have been followed in order to sort out the performance of each regular AWS/ARG stations. Thus, statistics based on time-series analysis of July-August, 2018-2020 were carried out and AWS/ARG stations were sorted thereafter.

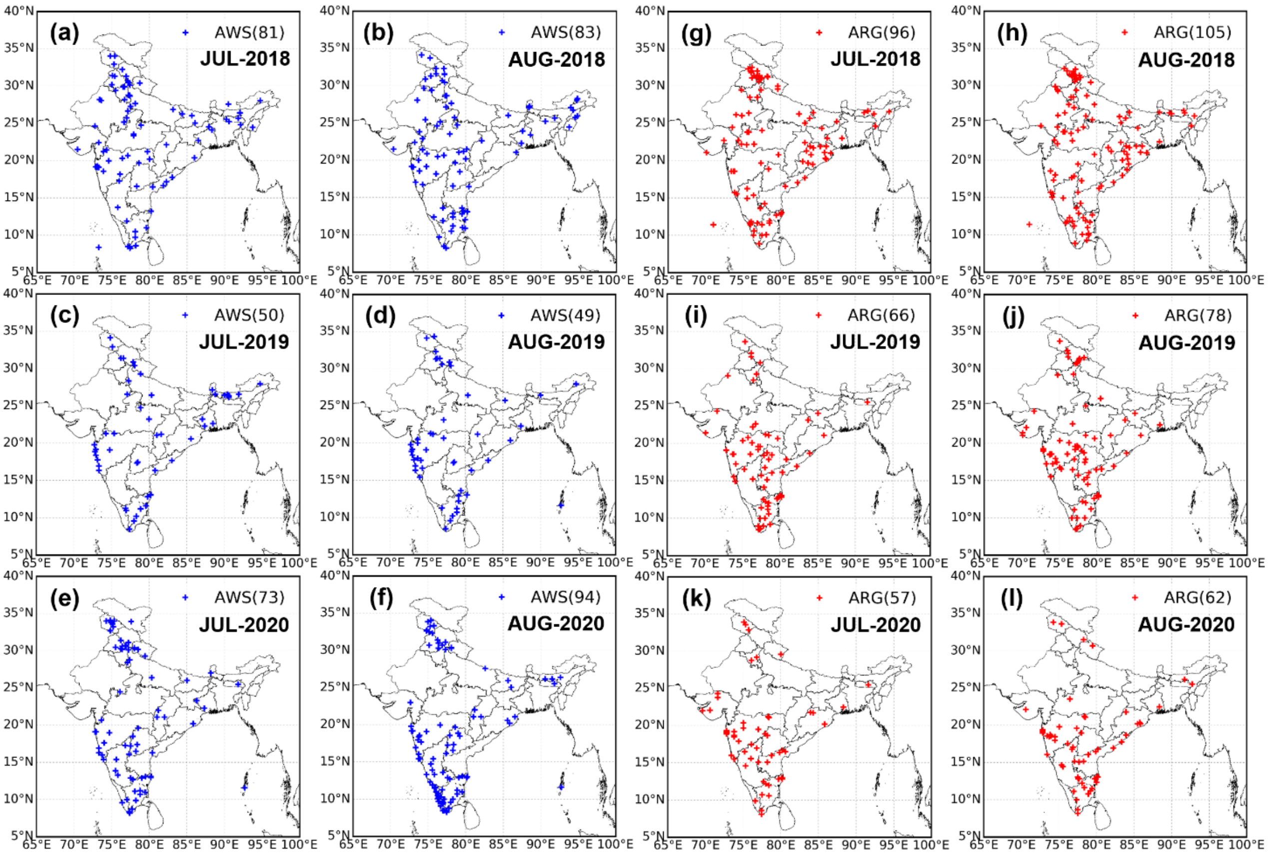


**Fig. 5A.** Same as Figure 4, but for matching AWS/ARG rainfall with other in-situ/gridded/merged gauge-satellite gridded observations. [Criterion used: $r \geq 0.7$ for AWS/ARG with an individual month of July or August]

Figure 5A represents the coverage plot of AWS and ARG stations with correlation of observed rainfall against SYNOP or gridded rainfall computed for the month (July and August) is greater than equal to 0.7. This particular type is defined as “Matching stations with other observations” in this particular study. On an average, in 2018, 82 and 101 AWS and ARG stations respectively, which is basically ~ 23.8% and ~ 18.3% of total AWS and ARG stations are recommended as good stations. In 2019, the average number was 48 (~ 21.7%) and 72 (~ 16.4%) respectively while in 2020, the number was 84 (~ 29.2%) and 60 (~ 17.2%) during July-August respectively. The percentage values within parentheses indicate the change in reported AWS/ARG out of total observations received at NCMRWF of July-August during 2018-2020 respectively.

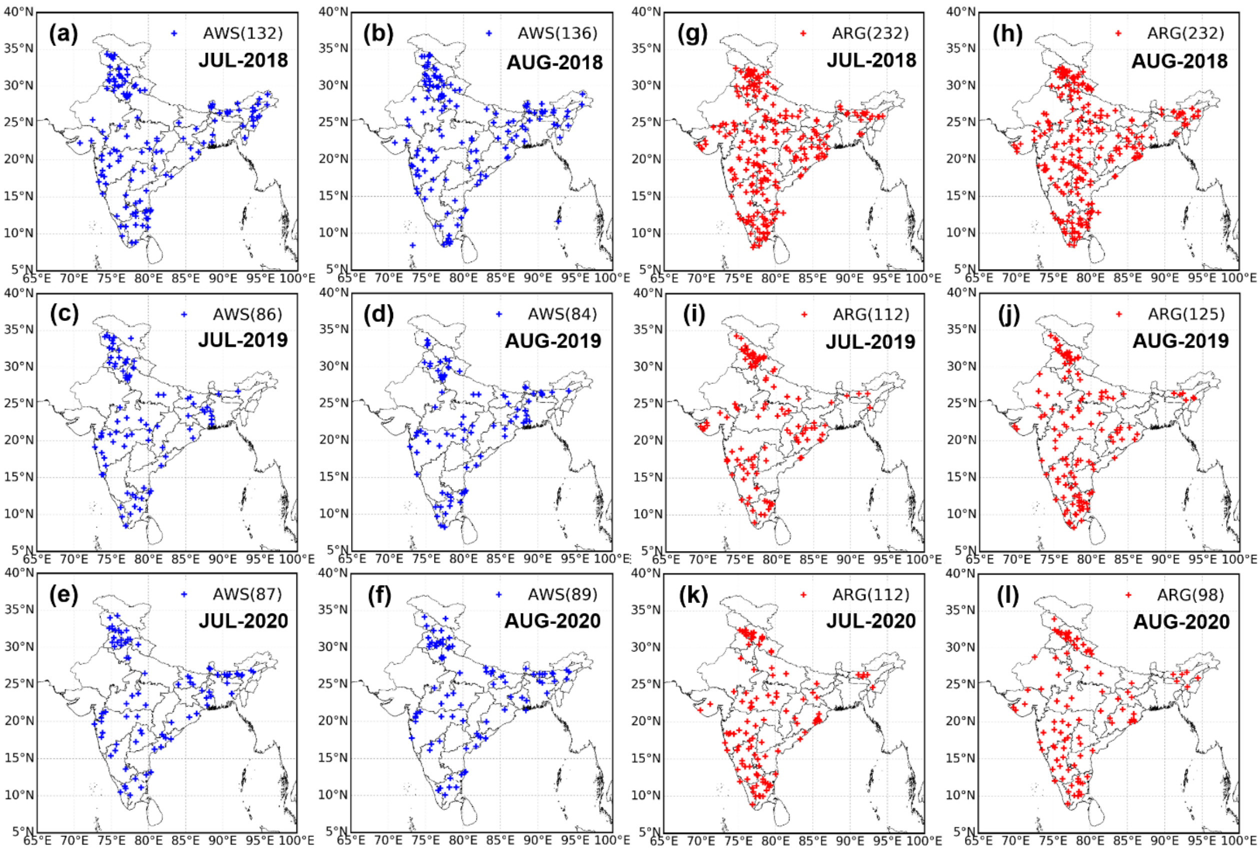


**Fig. 5B.** Same as Figure 4, but for non-matching AWS/ARG rainfall with other in-situ/gridded/merged gauge-satellite gridded observations. [Criterion used: r < 0.7 for AWS/ARG with an individual month of July or August]

On the contrary, there were also a large number of AWS/ARG stations which have poor correlation with individual rainfall time series. These stations are not recommended to be taken for any application purpose. Figure 5B represents the coverage plot for the counts of AWS and ARG stations whose correlation between the time-series of each stations for each individual month and the time series of each station observations as obtained by gridded observation (IMD rainfall) or merged satellite-gauge rainfall (MSG) being less than 0.7. This particular type is defined as "Non-Matching stations with other observations" in this particular study.

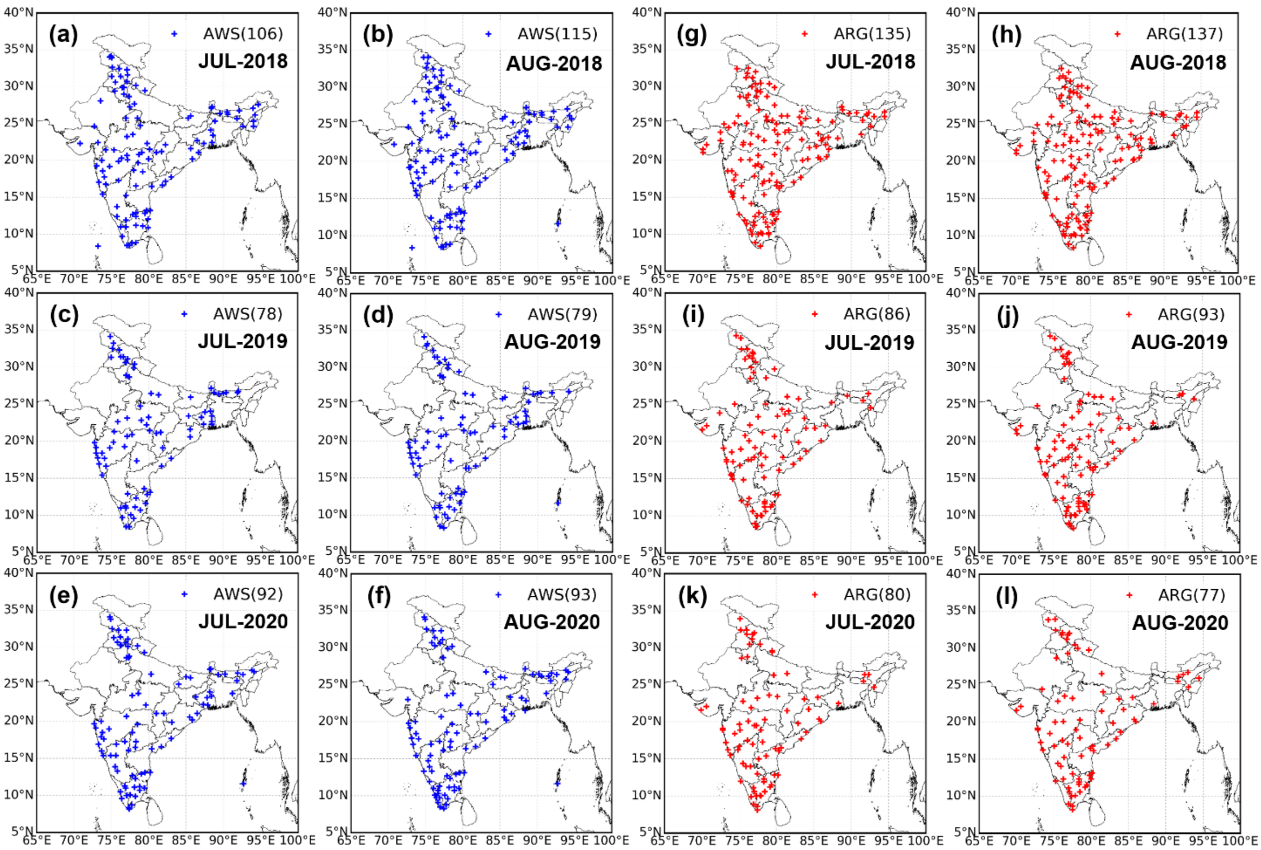


**Fig. 6.** Same as Figure 4, but for collocated AWS/ARG stations with land surface in-situ (SYNOP) observations. [Criterion used: SYNOP stations present within the radius of 50 km for regular AWS and ARG stations]

On an average, in 2018, 134 (~ 39%) and 232 (~ 42%) AWS and ARG stations respectively are not recommended as good stations. In 2019, the average number was 85 (~ 38.4%) and 119 (~

27.2%) respectively while in 2020, the number was 88 (~ 30.6%) and 105 (~ 30.2%) during July-August respectively.

There are a large number of AWS/ARG stations which have no collocated land surface in-situ observations (SYNOP) stations nearby. For validation purpose, rainfall recorded by SYNOP stations are always taken as ground truth. Hence, it is necessary to decipher the coverage of collocated SYNOP with AWS/ARG stations during the study period. Figure 6 represents the coverage plot for the counts of AWS and ARG stations which have collocated SYNOP within a radius of 50 km to that particular AWS or ARG station. On an average, 111 and 136 AWS and ARG stations respectively have atleast one SYNOP station within 50 km radius during July-August 2018, which is reduced to 78 and 90 respectively in 2019 while it increased to 93 and 79 respectively in 2020.

### 3.2. *Time Series plots*:

#### *3.2.1 Matching with in-situ/gridded (IMD/merged satellite) observations*

Figures 7 (A-C) represents the time-series plots for the AWS/ARG rainfall observations matching with land surface in-situ (SYNOP) or gridded rainfall observations (IMD/MSG) for July and August of the year 2018-2020 respectively. Figure 7A (a-d) indicates the time-series of AWS with SYNOP/IMD/MSG rainfall during July and August of 2018 while for ARG rainfall time-series are indicated in Figure 7A (e-h). In most of the cases, the correlation co-efficient of AWS/ARG rainfall with SYNOP rainfall observations during July and August 2018 is greater than 0.75 and the maximum correlation value being 0.99 for the best matching station. In any case, if SYNOP observation is missing or the correlation co-efficient value between SYNOP and AWS/ARG rainfall observations is less than 0.7, then extracted station data from IMD-gridded or merged satellite-gauge rainfall (MSG) comes into account. In that case, the correlation co-efficient of AWS/ARG are checked with IMD/MSG rainfall observations and if the correlation value shows greater than 0.7 while SYNOP is not showing, still the station can be treated as good station.

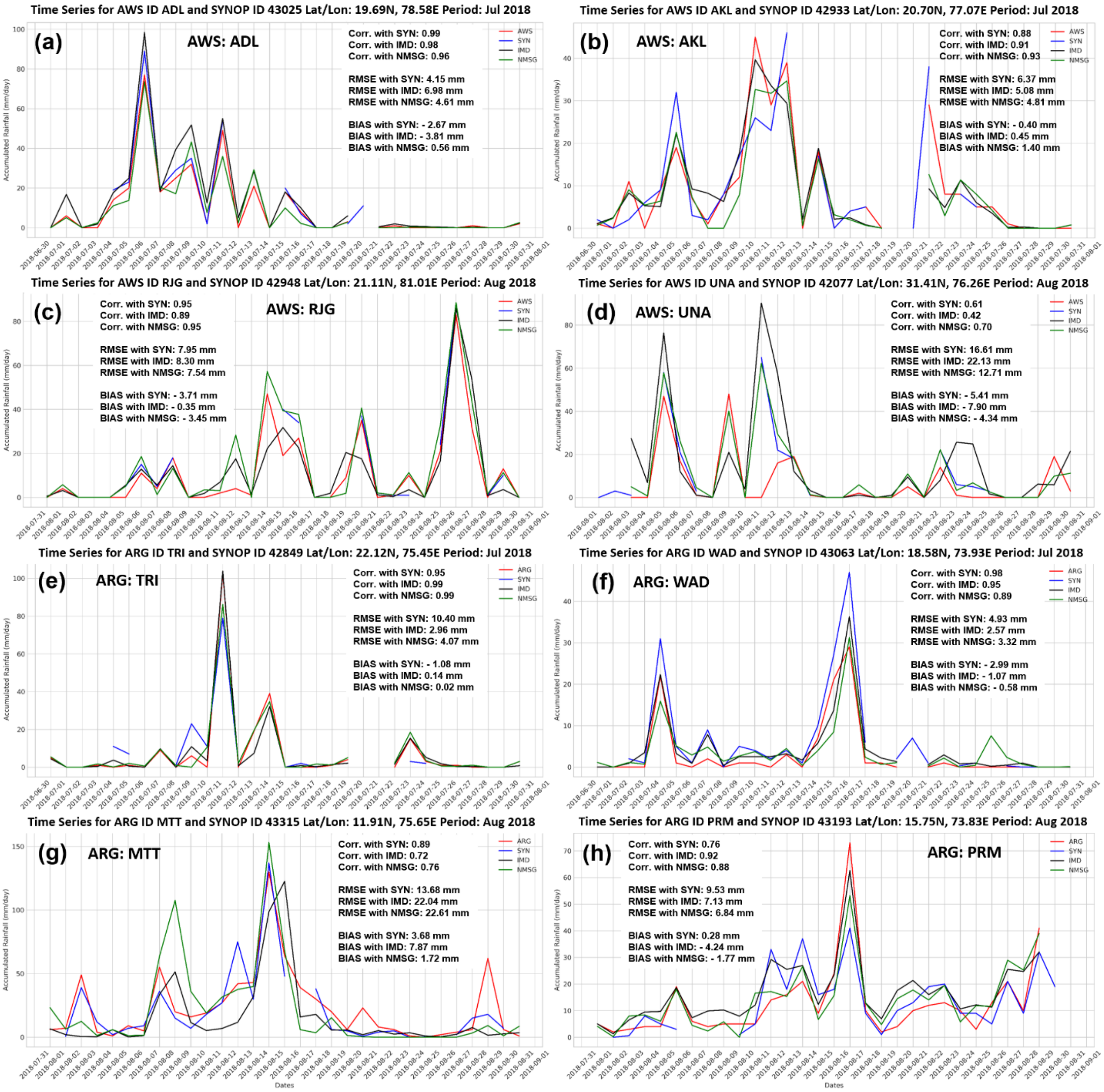


**Fig. 7A.** Time-series plots for in-situ (SYNOP) or gridded (IMD/MSG) rainfall completely matching with (a-d) AWS rainfall and (e-h) ARG rainfall respectively during July and August of 2018. Statistical details within figure can be found more clearly in Appendix Table 1a.

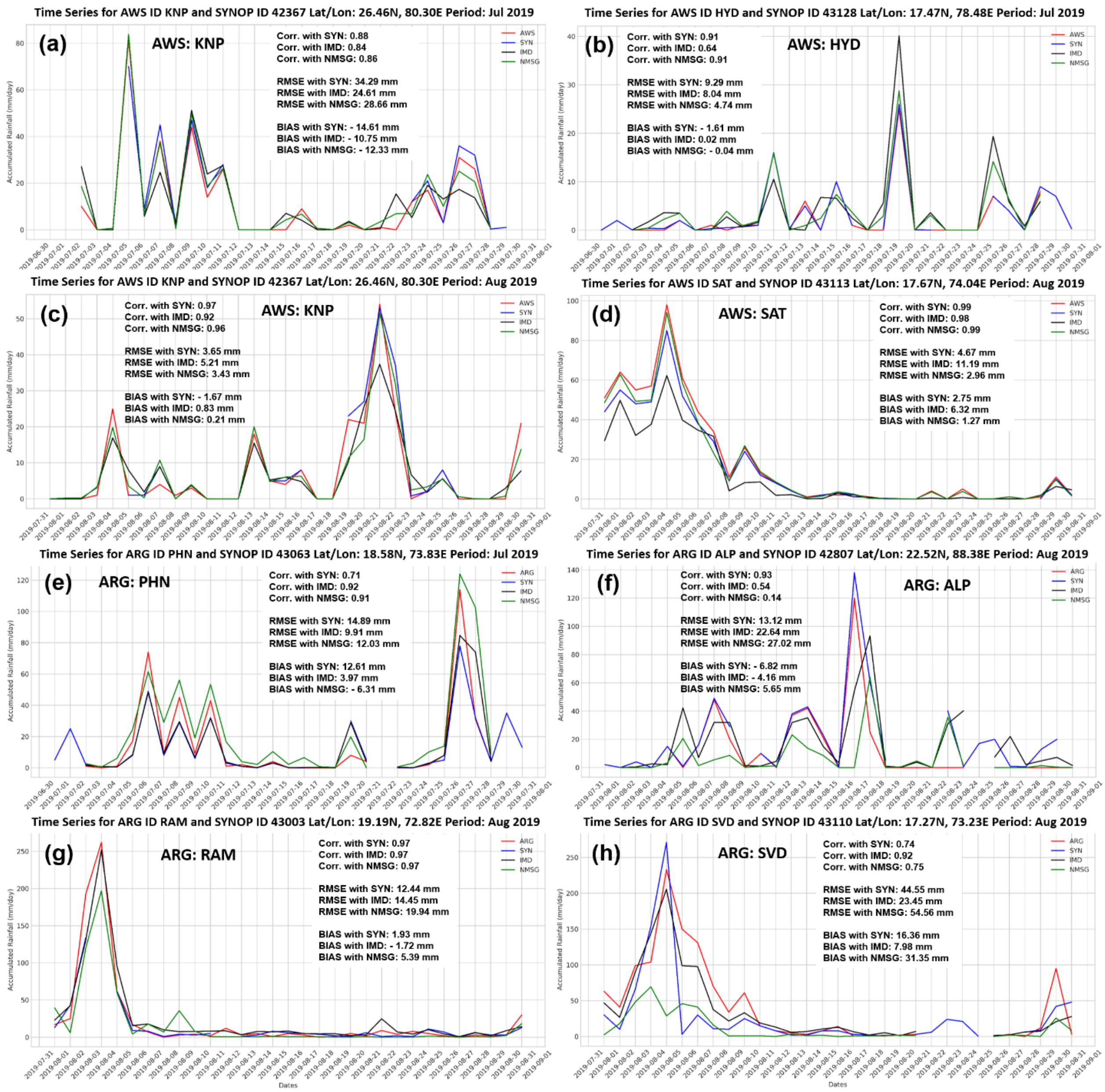


**Fig. 7B.** Time-series plots for in-situ (SYNOP) or gridded (IMD/MSG) rainfall completely matching with (a-d) AWS rainfall and (e-h) ARG rainfall respectively during July and August of 2019. Statistical details within figure can be found more clearly in Appendix Table 1b.

It is hereby noted that, the stations with significant correlations have less root mean square error (RMSE) value and BIAS between the two time series is also significantly less (Figure 7A a-h). Similarly, Figure 7B (a-d) indicates the time-series of AWS with SYNOP/IMD/MSG rainfall during July and August of 2019 while for ARG rainfall time-series are indicated in Figure 7B (e-h).

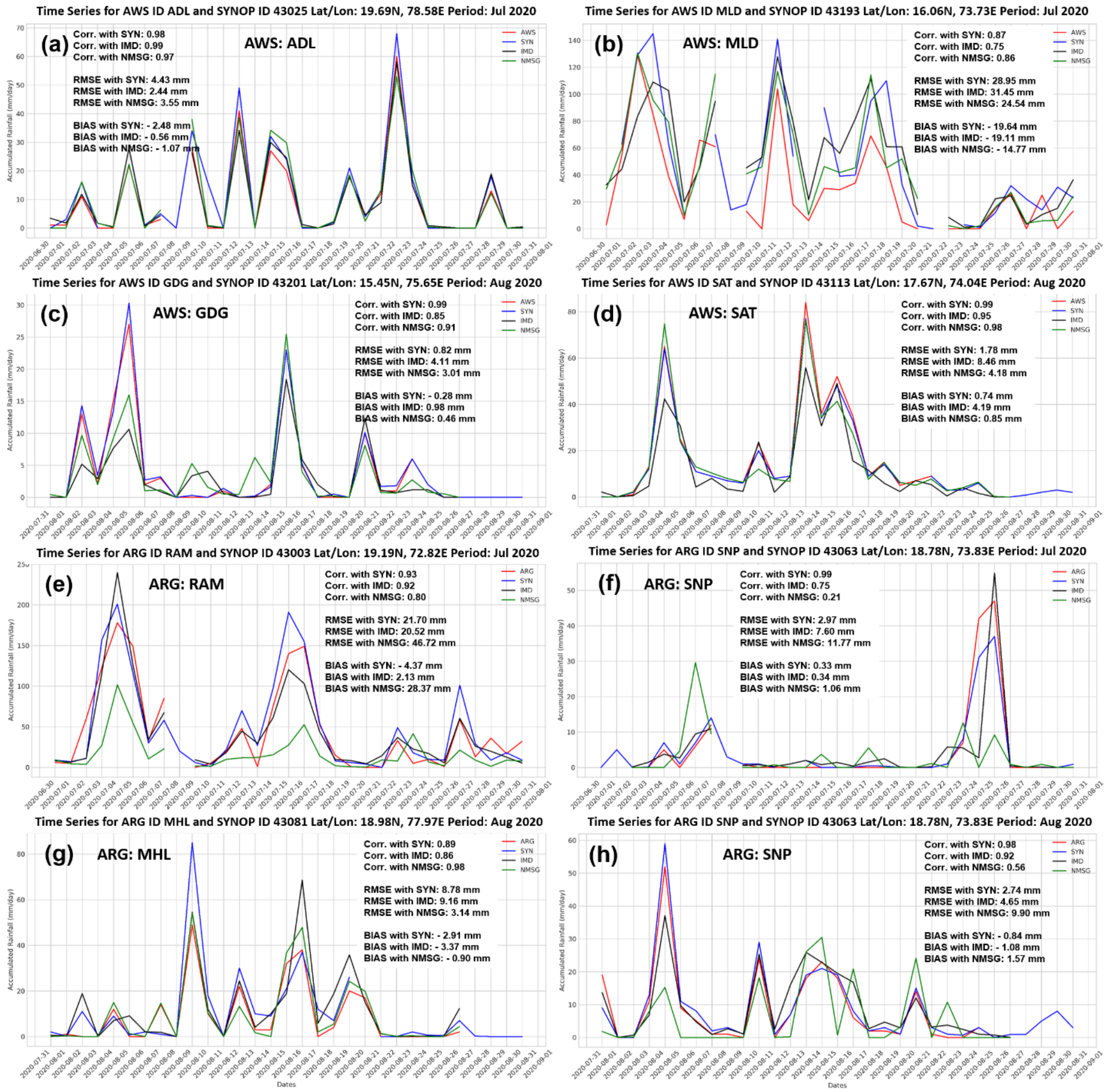


**Fig. 7C.** Time-series plots for in-situ (SYNOP) or gridded (IMD/MSG) rainfall completely matching with (a-d) AWS rainfall and (e-h) ARG rainfall respectively during July and August of 2020. Statistical details within figure can be found more clearly in Appendix Table 1c.

Moreover, Figure 7C (a-d) indicates the time-series of AWS with SYNOP/IMD/MSG rainfall during July and August of 2020 while for ARG rainfall time-series are indicated in Figure 7C (e-h). The peak rainfall by SYNOP/IMD/MSG are captured very well in all the cases of all the years and the continuous monthly time series of AWS/ARG rainfall are also significantly matched with the other observations.

### *3.2.2 Partially matching with in-situ observations*

Figures 8 (A-B) represents the time-series plots for the AWS/ARG rainfall observations partially matching with in-situ (SYNOP) or gridded rainfall observations (IMD/MSG) for July and August during 2018-2020 respectively. Figure 8A (a-b) indicates that during the starting period of the month, when SYNOP/IMD/MSG shows peak rainfall, AWS failed to capture the rainfall. But in the mid/last phase of the month, AWS rainfall perfectly matches with SYNOP/IMD/MSG rainfall and even AWS also captured peak rainfall activity.

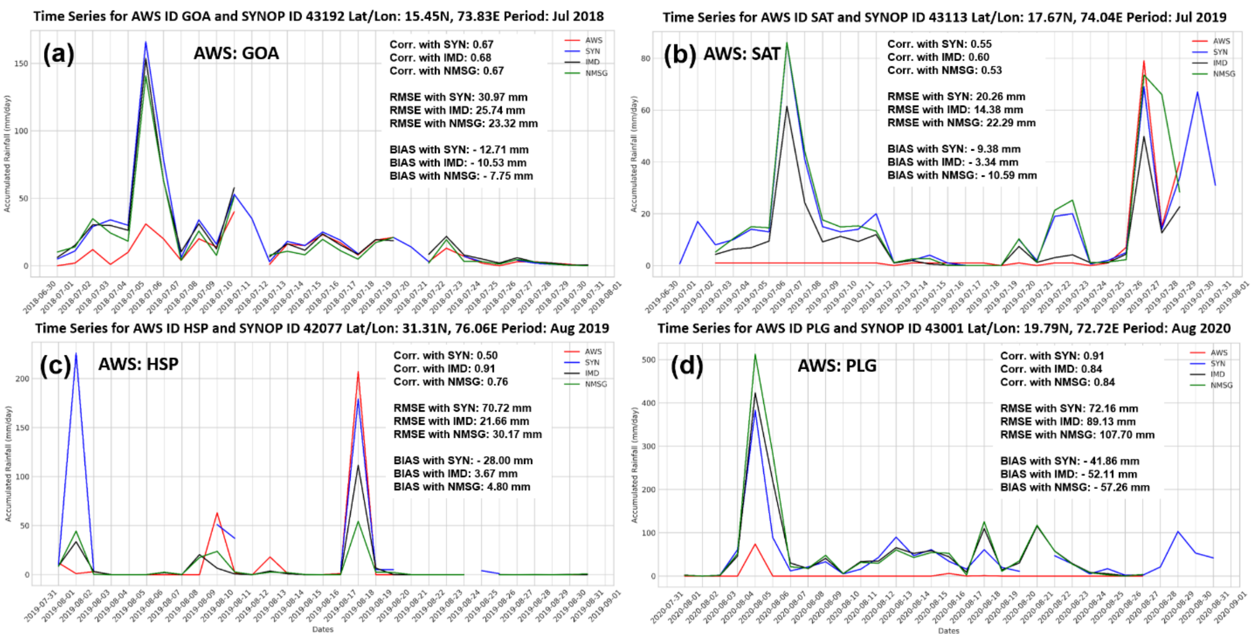


**Fig. 8A** (a-d) Time-series plots for in-situ (SYNOP) or gridded (IMD/MSG) rainfall partially matching with AWS rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 2a.

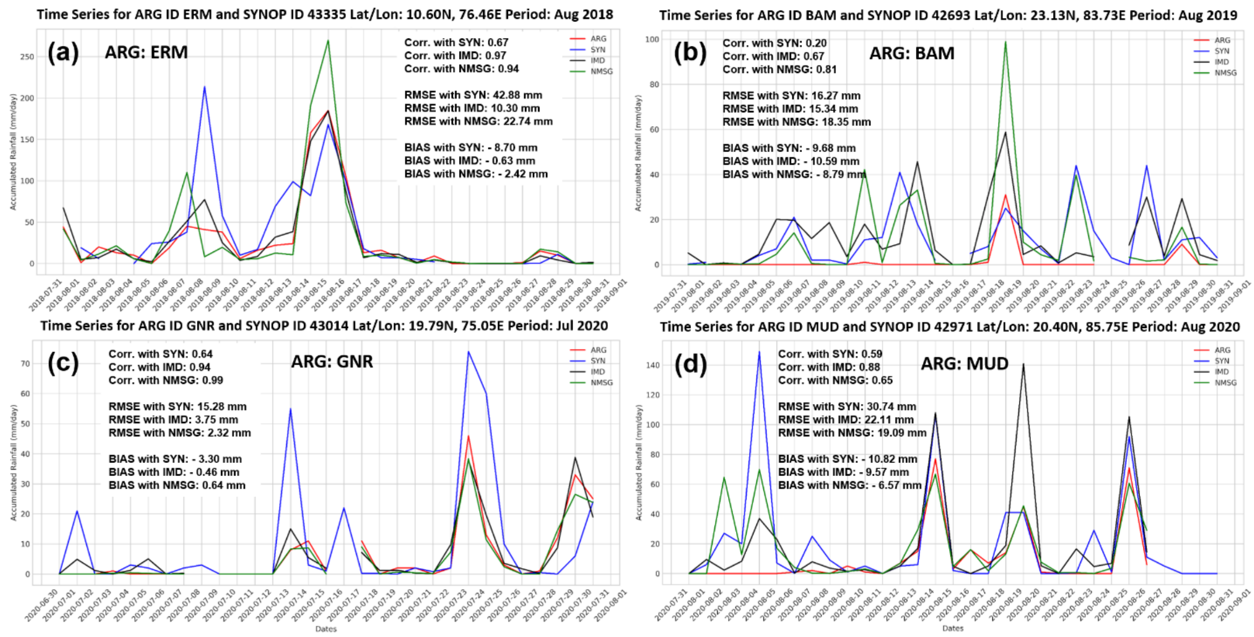


**Fig. 8B** (a-d) Time-series plots for in-situ (SYNOP) or gridded (IMD/MSG) rainfall partially matching with ARG rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 2b.

Similarly, Figure 8A (c-d) indicates that AWS partially captured the other rainfall observations, either in the first or mid phase of that particular month. Hence, it can be found that the correlation co-efficient between the time series of SYNOP/IMD/MSG rainfall with AWS rainfall is particularly low (in some cases, it may be high due to the best fitting in any of the phase of the month) and their RMSE and BIASes are quite high indicating partial matching of different rainfall observations. Similarly, Figures 8B (a-d) represents the time-series plots for the ARG rainfall observations partially matching with land surface in-situ (SYNOP) or gridded rainfall observations (IMD/MSG) for July and August during 2018-2020 respectively. It is found to be comparable to that of Figure 8A (a-d) as mentioned above and is self-explanatory.

### *3.2.3 Matching with satellite/gridded observations*

As noted before, there are a large number of AWS and ARG stations which are not collocated with any SYNOP station within the radius of 50 km. For those stations, which have no collocated SYNOP, IMD and MSG rainfall remains the only option to check the quality of the station through correlation co-efficient, RMSE or BIAS during an individual month of a particular year. Figure 9 (A-B) represents the time-series plots for the AWS/ARG rainfall observations matching with gridded rainfall observations (IMD/MSG) while SYNOP rainfall is absent or missing for July and August during 2018-2020 respectively.

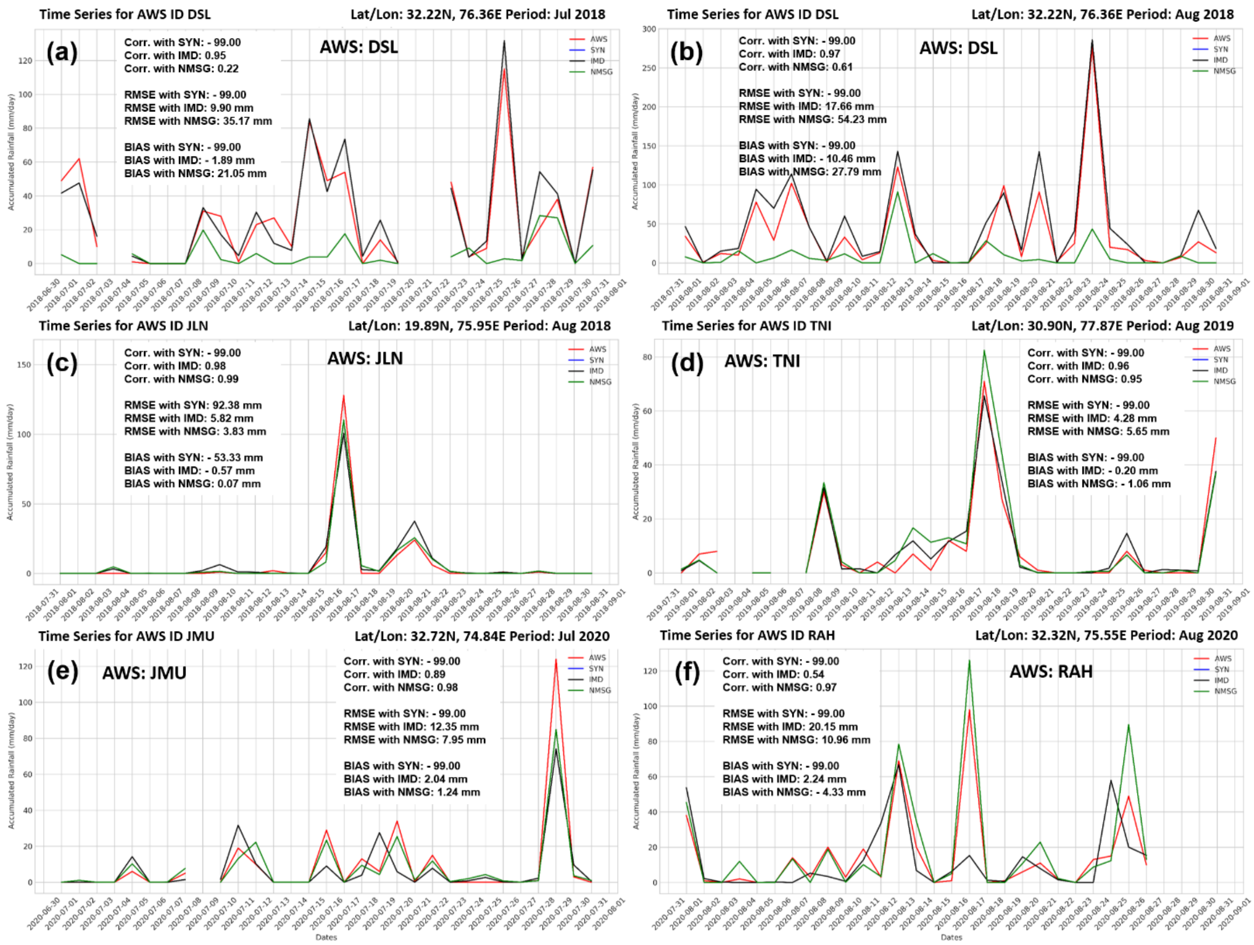


**Fig. 9A** (a-f) Time-series plots for gridded (IMD/MSG) rainfall (while for missing SYNOP) completely matching with AWS rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 3a.

It is indicative from Figure 9A (a-f) as well as Figure 9B (a-f) that the correlation co-efficient, RMSE and BIAS between AWS and SYNOP rainfall is – 99.00 since these stations have no collocated SYNOP stations. All the figures in Figure 9 (A-B) deciphers that the time series of both AWS and ARG rainfall respectively either matches pretty well with the extracted station rainfall from IMD-gridded data or MSG rainfall data in an individual month of a particular year.

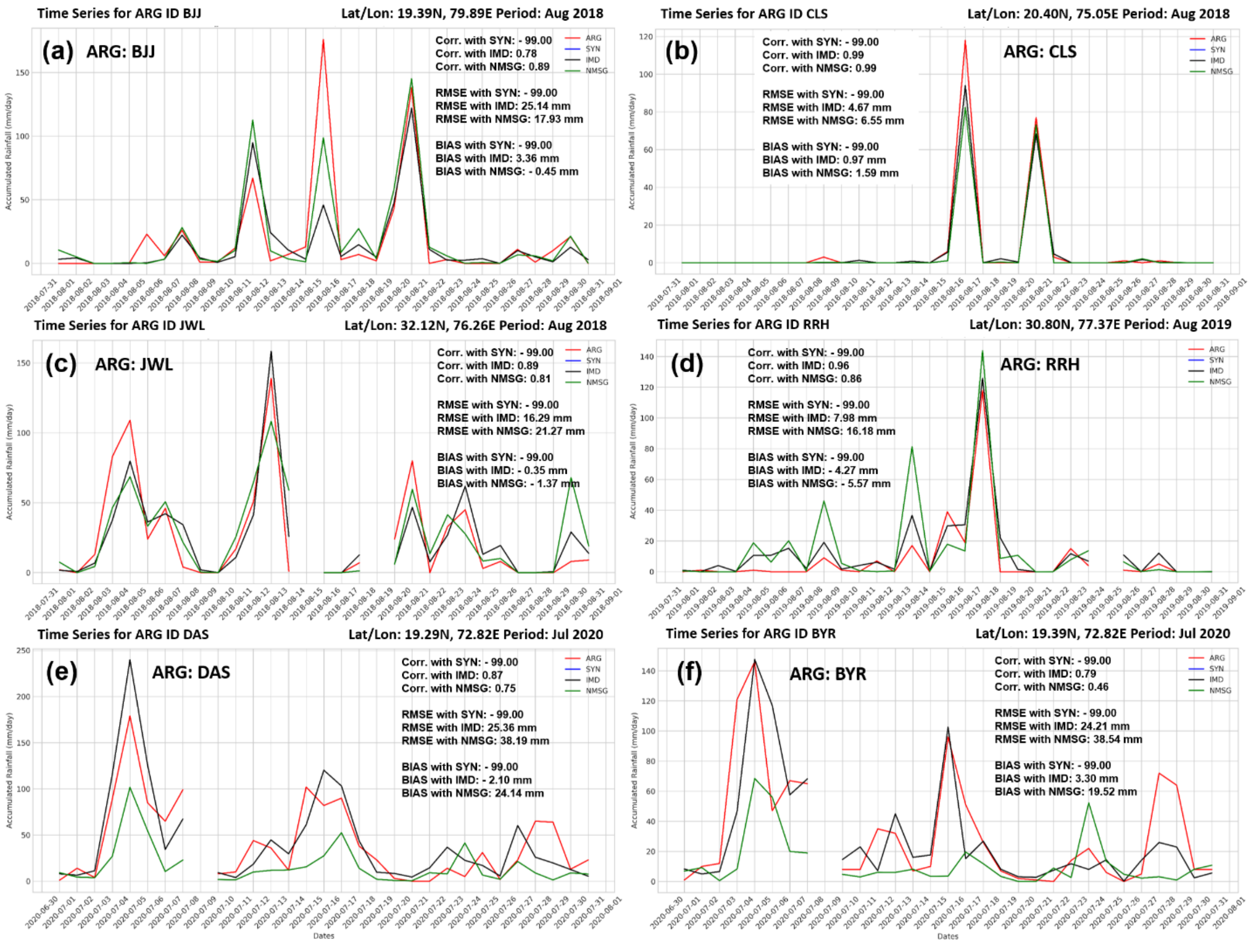


**Fig. 9B** (a-f) Time-series plots for gridded (IMD/MSG) rainfall (while for missing SYNOP) completely matching with ARG rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 3b.

### *3.2.4 Partially matching with satellite/gridded observations*

Figures 10 (A-B) represents the time-series plots for the AWS/ARG rainfall observations partially matching with gridded rainfall observations (IMD/MSG) for July and August during

2018-2020 respectively. These stations basically indicate that either the rainfall measuring sensor stopped (or started) working in the first or mid or last phase of the month due to various reasons.

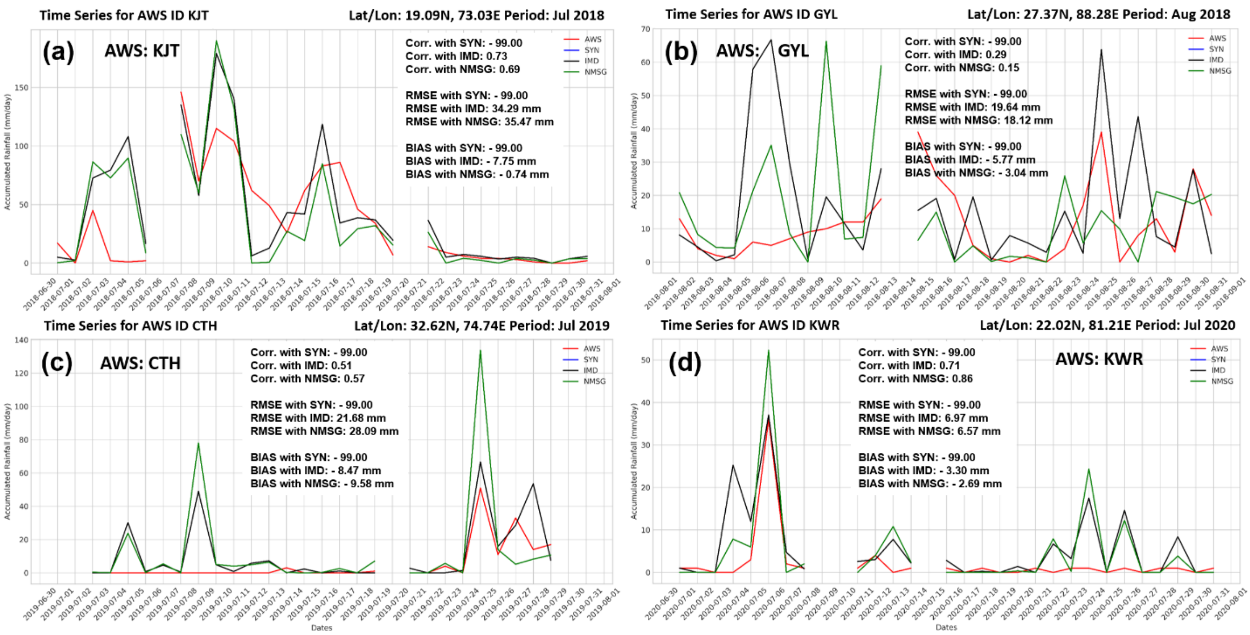


**Fig. 10A** (a-d) Time-series plots for gridded (IMD/MSG) rainfall data (while for missing SYNOP) partially matching with AWS rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 4a.

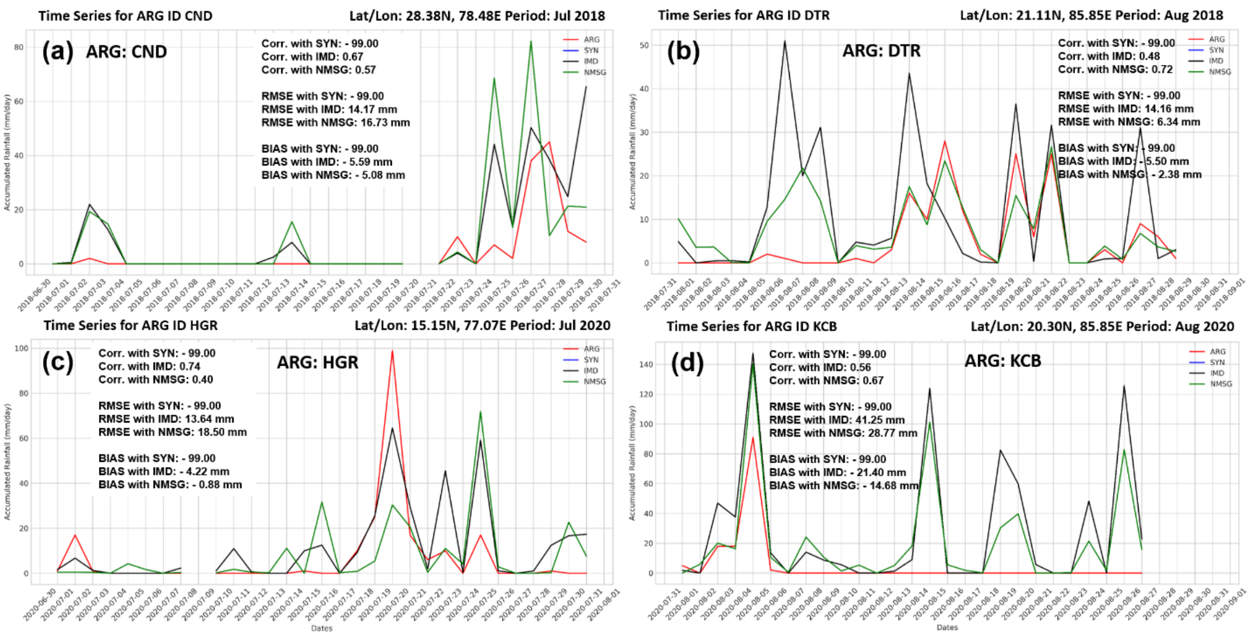


**Fig. 10B** (a-d) Time-series plots for gridded (IMD/MSG) rainfall data (while for missing SYNOP) partially matching with ARG rainfall during July and August of 2018-2020 respectively. Statistical details within figure can be found more clearly in Appendix Table 4b.

Similar to Figure 8 (A-B), the time- series of rainfall from AWS or ARG stations partially matches (either in the first, mid or last phase of the month) with IMD or MSG rainfall observations during an individual month of a particular year.

### *3.2.5 Not matching with in-situ/gridded (IMD/merged satellite) observations*

It is found that there are several AWS and ARG stations for which observed rainfall is not matching with either SYNOP or any gridded rainfall data. For some stations, the observed rainfall from AWS/ARG is negligible throughout the period, whereas nearby SYNOP and gridded rainfall shows good amount of rainfall at that particular location.

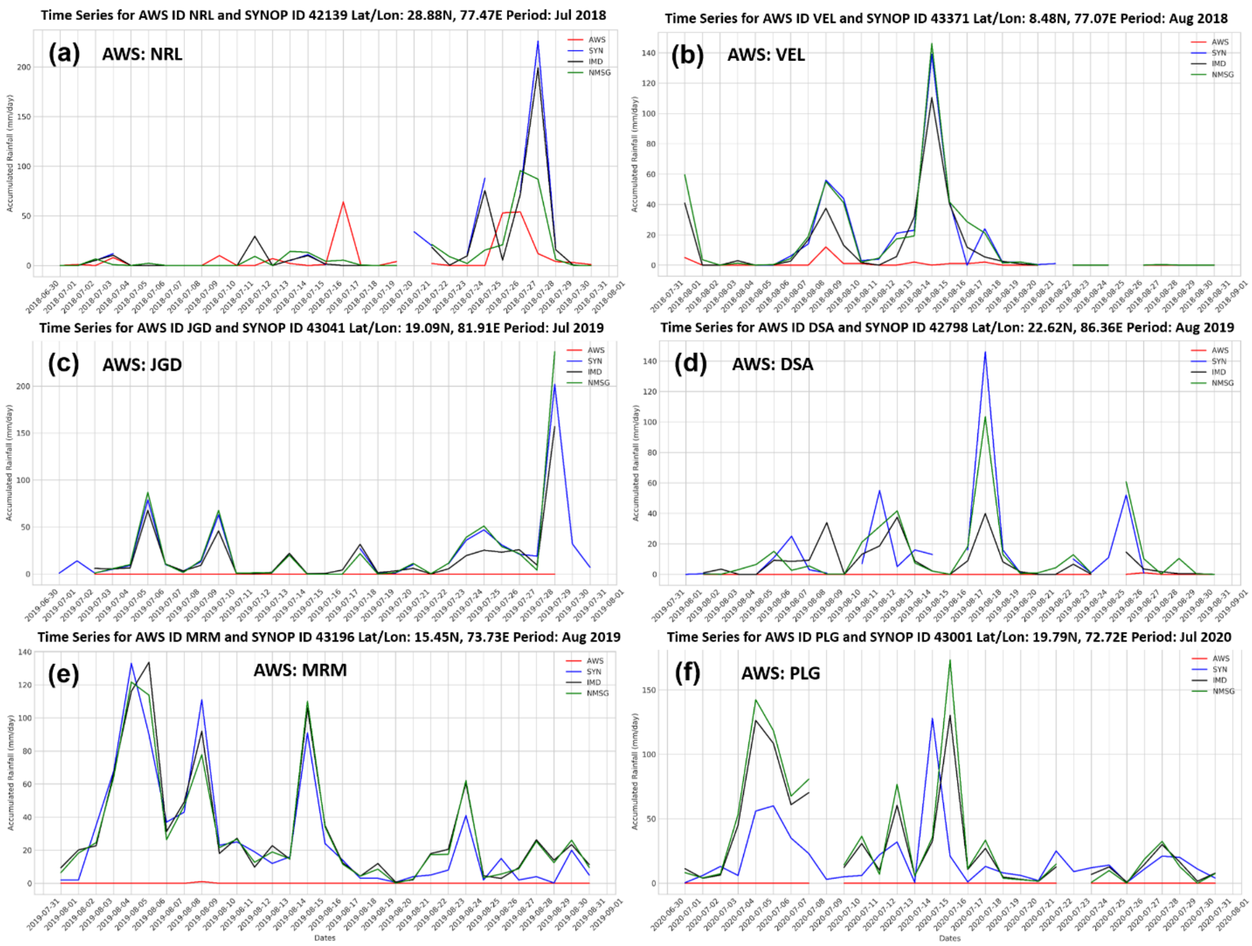


**Fig. 11A** (a-f) Time-series plots for land surface in-situ (SYNOP) or gridded (IMD/MSG) rainfall not matching with AWS rainfall during July and August of 2018-2020 respectively.

Time series plots for some of the locations are shown in Figure 11. Figure 11 (A-B) represents the time-series plots for the AWS and ARG rainfall observations respectively which do not capture the SYNOP or gridded rainfall observations (IMD/MSG) at all for July and August during 2018-2020. It can be found from the figures that in most of the cases, AWS/ARG rainfall is showing zero rainfall and failed to capture the time series of SYNOP/IMD/MSG rainfall showing heavy or very heavy or extremely heavy rainfall in some months of a particular year.

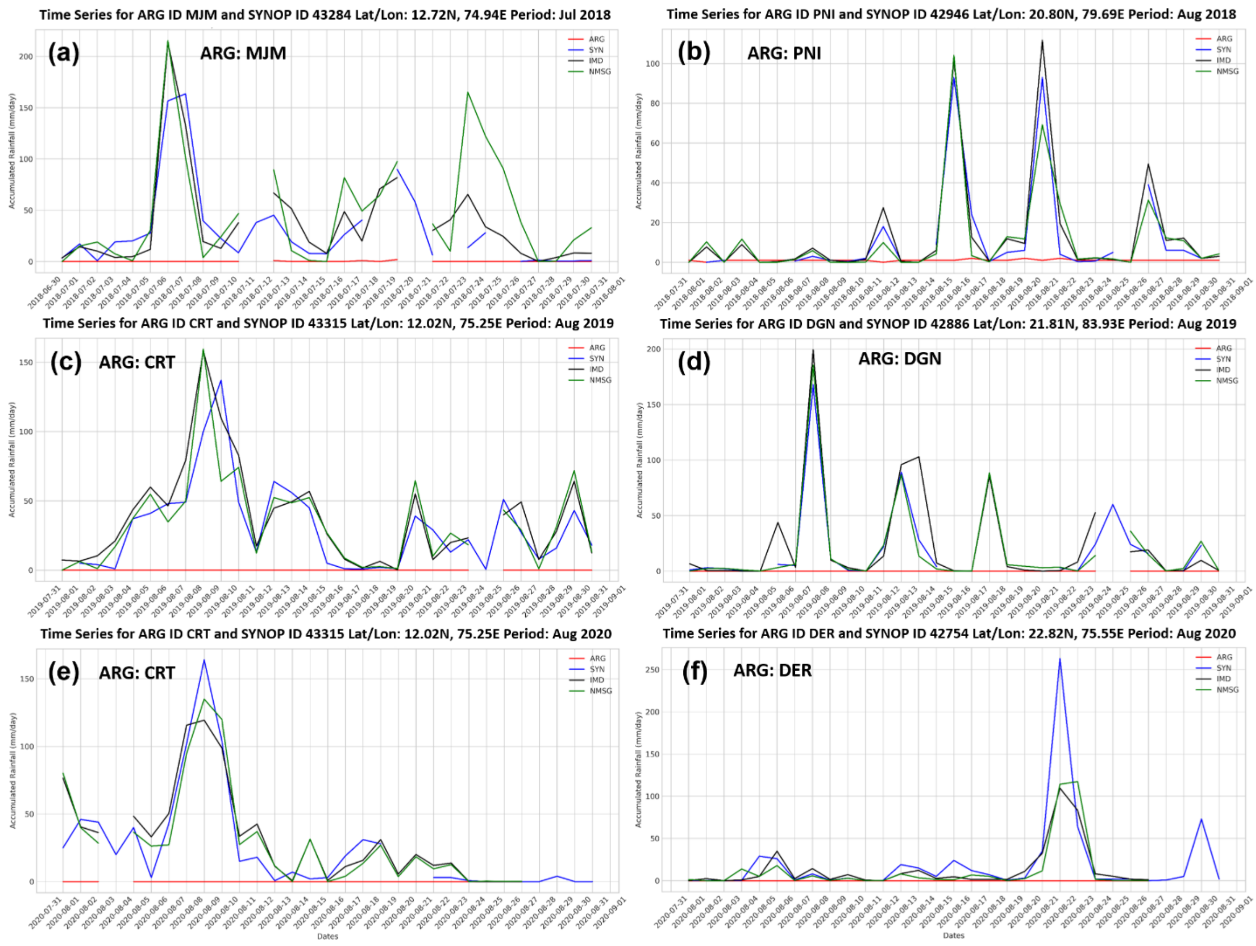


**Fig. 11B** (a-f) Time-series plots for land surface in-situ (SYNOP) or gridded (IMD/MSG) rainfall not matching with ARG rainfall during July and August of 2018-2020 respectively.

These type of observations for capturing or missing different intensity of rainfall by AWS/ARG stations are described in the upcoming section (Section 3.3). Since, AWS or ARG rainfall do not capture the other observations, statistics table is not shown within the figure.

There are some AWS and ARG stations reporting very high amount of rainfall (spurious) on some days, which are not at all observed either in SYNOP or gridded rainfall data. Einfalt et al. (2006) studied the sampling error during the quality control procedure and has found that mechanical complications of the sensor could be associated with the same. However, Estévez et al. (2015) found that verification of rainfall records for being true correspond to spurious rainfall signals can only be approached by testing the consistency of rainfall data with other meteorological parameters measured at the same meteorological station.

Figure 11C (a-h) represents the time series plots for the spurious values of AWS/ARG rainfall observations along with other rainfall observations (SYNOP/IMD/MSG) for July and August during 2018-2020 respectively. Figure 11C (a) indicates that AWS station 'DRS' is showing rainfall in the range of 300-350 mm day$^{-1}$ for the month of July 2018 while other observations do not show any such extreme rainfall. Figure 11C (b) shows that AWS station 'SLP' has given a peak of ~ 500 mm day$^{-1}$ rainfall in August, 2018 and AWS station 'KEJ' has given three extreme rainfall peaks (minimum: ~ 620 mm day$^{-1}$ and maximum: ~ 850 mm day$^{-1}$) during July, 2019 (Figure 11C c). Figure 11C (d) indicates that AWS station 'KEJ' has given two rainfall peaks (~ 700 mm day$^{-1}$) in August, 2019 similar to the previous month of this year. AWS station 'ANR' has indicated two rainfall peaks of ~ 850 mm day$^{-1}$ and ~ 980 mm day$^{-1}$ respectively in July, 2020 (Figure 11C e). Figure 11C (f) shows that AWS station 'AKL' has given spurious rainfall peak ~ 700 mm day$^{-1}$ in August, 2020. Figure 11C (g) indicates that ARG station 'MTZ' has given rainfall peak ~ 140 mm day$^{-1}$ in July, 2018 while in July, 2020, ARG station 'AAD' indicated spurious rainfall peak of ~ 500 mm day$^{-1}$ (Figure 11C h). Thus, it may be noted that, AWS stations are much more vulnerable to indicate spurious rainfall peaks or ‘false positives’ than the ARG stations during July and August of 2018-2020 respectively.

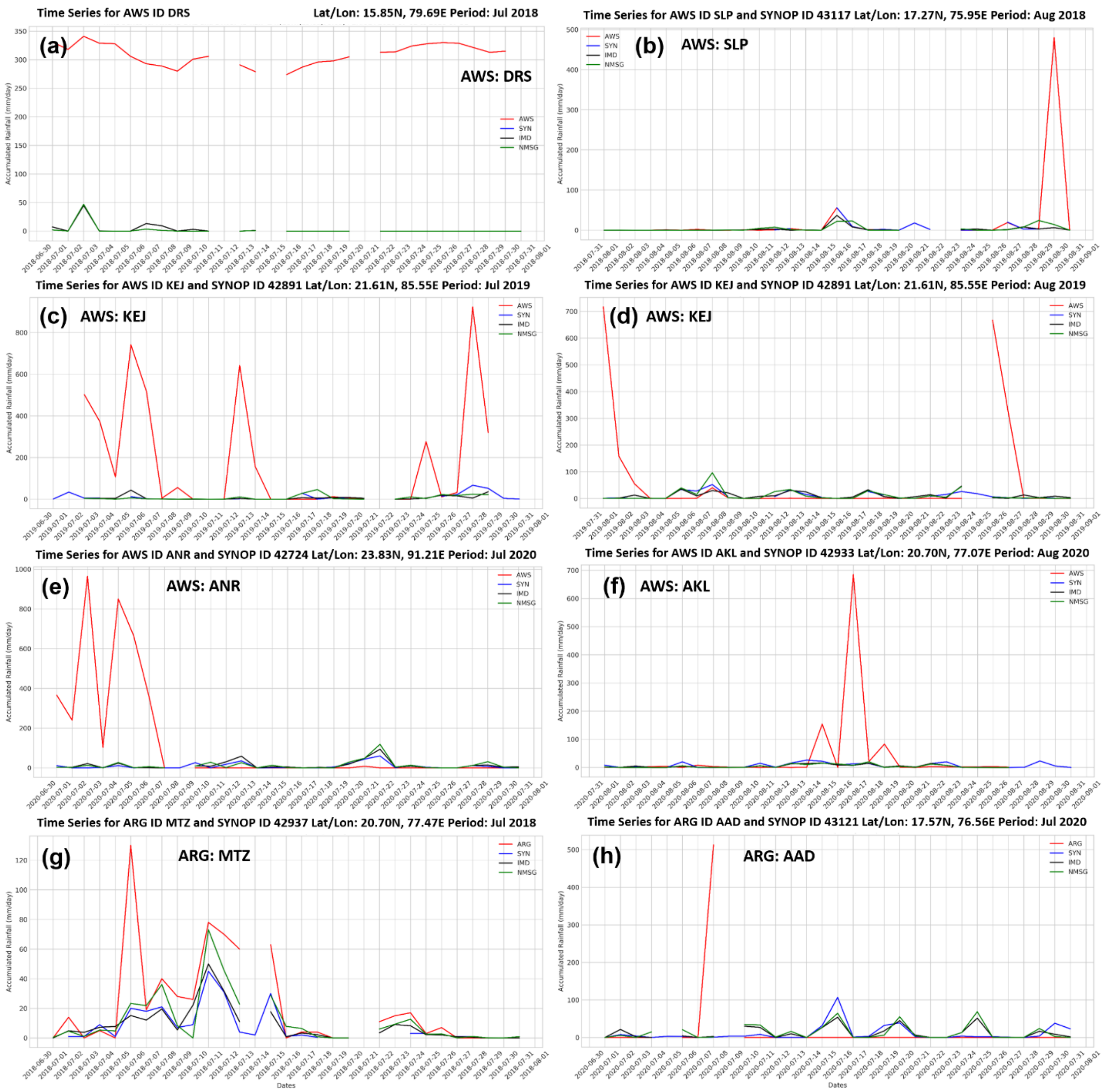


**Fig. 11C** (a-h) Time-series plots for spurious AWS/ARG rainfall with land surface in-situ (SYNOP) or gridded (IMD/MSG) rainfall during July and August of 2018-2020 respectively.

## 3.3. Extreme Rainfall Cases: Time Series plots

Based on the amount of accumulated rainfall in a day for extreme event studies, India Meteorological Department (IMD) has classified rainfall into three broad categories, viz. i) Heavy Rainfall (HR, 64.5 mm < R ≤ 115.4 mm), ii) Very Heavy Rainfall (VHR, 115.5 mm < R ≤ 204.4 mm) and iii) Extremely Heavy Rainfall (EHR, R ≥ 204.5 mm), where R indicates Rainfall.

An attempt has been made to verify the cases where AWS/ARG stations could capture this extreme rainfall events.

### *3.3.1* AWS/ARG capturing Extremely Heavy Rainfall observations

Figure 12A (a-d) represents the time-series plots for the AWS/ARG rainfall observations from this new network that could capture the EHR events and matching well with SYNOP or gridded rainfall observations (IMD/MSG) during the core monsoon of 2018-2020.

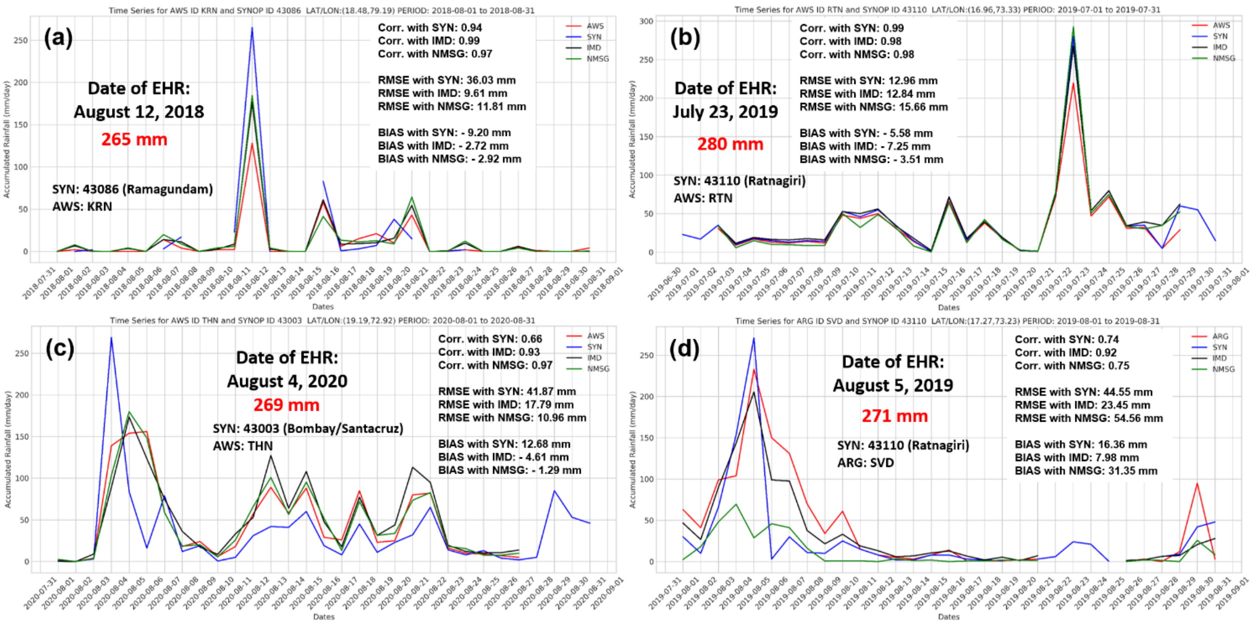


**Fig. 12A** (a-d) Time-series plots for AWS and ARG rainfall along with SYNOP or gridded (IMD/MSG) rainfall for EHR matching cases.

It is to be noted that all the EHR hotspot locations do not either have any collocated AWS/ARG stations or if any AWS/ARG stations are present nearby hotspot locations, they could not have captured the event (details described in the next section). Hence, in these figures only those stations are shown which have collocated AWS/ARG stations that has excellently captured the EHR event. Figure 12A (a) indicates that the extremely heavy rainfall event occurred in August 12, 2018 over Ramagundam (with nearest AWS station ‘KRN’) with an accumulated rainfall amount to 265 mm according to SYNOP rainfall observations. The AWS station captured the EHR event but is underestimated, however, the station excellently follows the SYNOP/IMD/MSG rainfall all throughout the month. IMD/MSG too have not captured the intensity of the SYNOP-observed extreme rainfall event. Statistical details are given within the figure itself. Figure 12A (b) indicates that the EHR event occurred in July 23, 2019 over Ratnagiri (with nearest AWS station ‘RTN’) with an accumulated rainfall amount to 280 mm according to SYNOP rainfall observations. The time series of AWS rainfall excellently captured the whole month rainfall observations from SYNOP/IMD/MSG, with special emphasis on the EHR event. Figure 12A (c) indicates that the EHR event occurred in August 4, 2020 over Bombay/Santacruz (with nearest AWS station ‘THN’) with an accumulated rainfall amount to 269 mm according to SYNOP rainfall observations. The station failed to capture the peak intensity of the EHR event in the first phase of the month (August 4, 2020), however, it has slightly overestimated in the mid/last phase of the month. Figure 12A (d) indicates that the EHR event occurred in August 5, 2019 over Ratnagiri (with nearest ARG station ‘SVD’) with an accumulated rainfall amount to 271 mm according to SYNOP rainfall observations. In this case, the AWS station ‘RTN’ have not provided the rainfall data for the whole month of August, 2019. The ARG station captured the EHR event excellently as indicated by SYNOP and have followed the SYNOP/IMD/MSG rainfall all throughout the month. Statistical details are given within each of the figure itself.

### *3.3.2* AWS/ARG not capturing Extremely Heavy Rainfall observations

Figure 12B (a-d) represents the time-series plots for the AWS/ARG rainfall observations that has failed to capture EHR observations reported by SYNOP or gridded rainfall observations (IMD/MSG) during the core monsoon of 2018-2020.

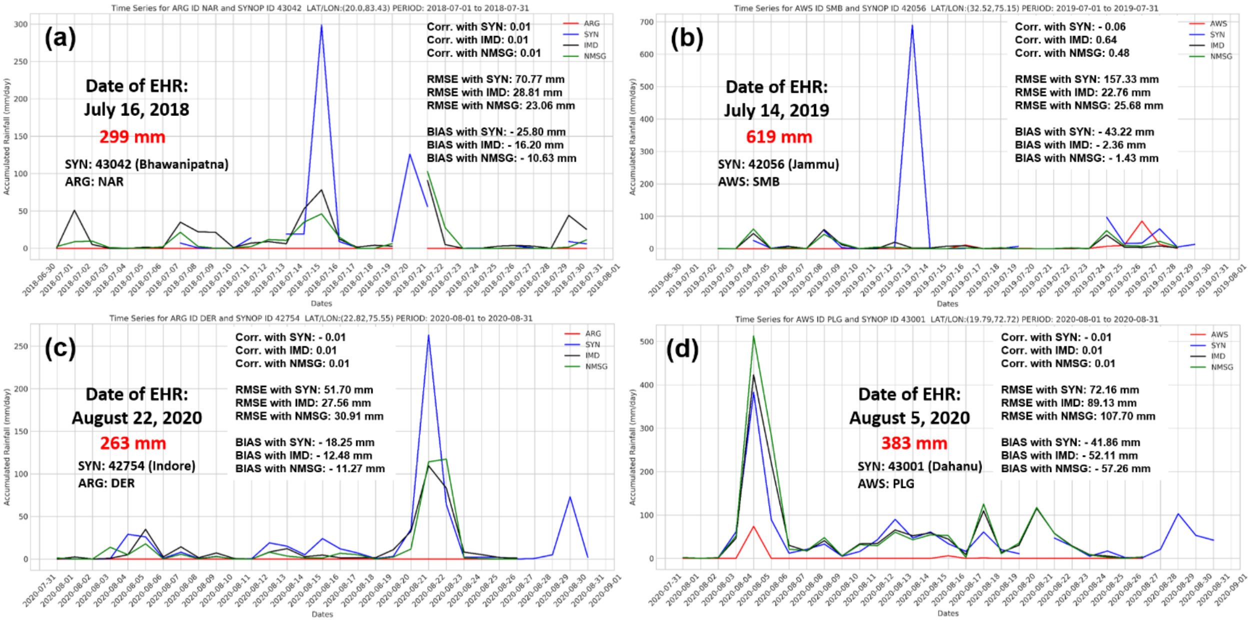


**Fig. 12B** (a-d) Time-series plots for AWS and ARG rainfall along with SYNOP or gridded (IMD/MSG) rainfall for non-matching EHR matching cases.

In all the cases of these EHR events during July and August of 2018-2020, both AWS and ARG stations has completely failed to capture the extreme event and not even they have captured any rainfall occurrences as shown by SYNOP/IMD/MSG observations. Statistical details are given within each of the figure itself and is self-explanatory. In most of the cases, AWS and ARG stations have provided zero rainfall data which indicates that either mechanical errors occurred during extreme events or sensors are non-functional due to non-calibration or non-maintenance of the AWS/ARG site. Thus, a periodic maintenance of the stations, including sensor checks and calibrations, as well as validation of data collected are need to be properly assessed and monitored.

### *3.3.3* AWS/ARG capturing Very Heavy Rainfall observations

Similar to Figure 12A, Figure 12C represents the time-series plots for the AWS/ARG rainfall observations that excellently captured the land surface in-situ (SYNOP) or gridded rainfall observations (IMD/MSG) during 2018-2020 containing VHR events.

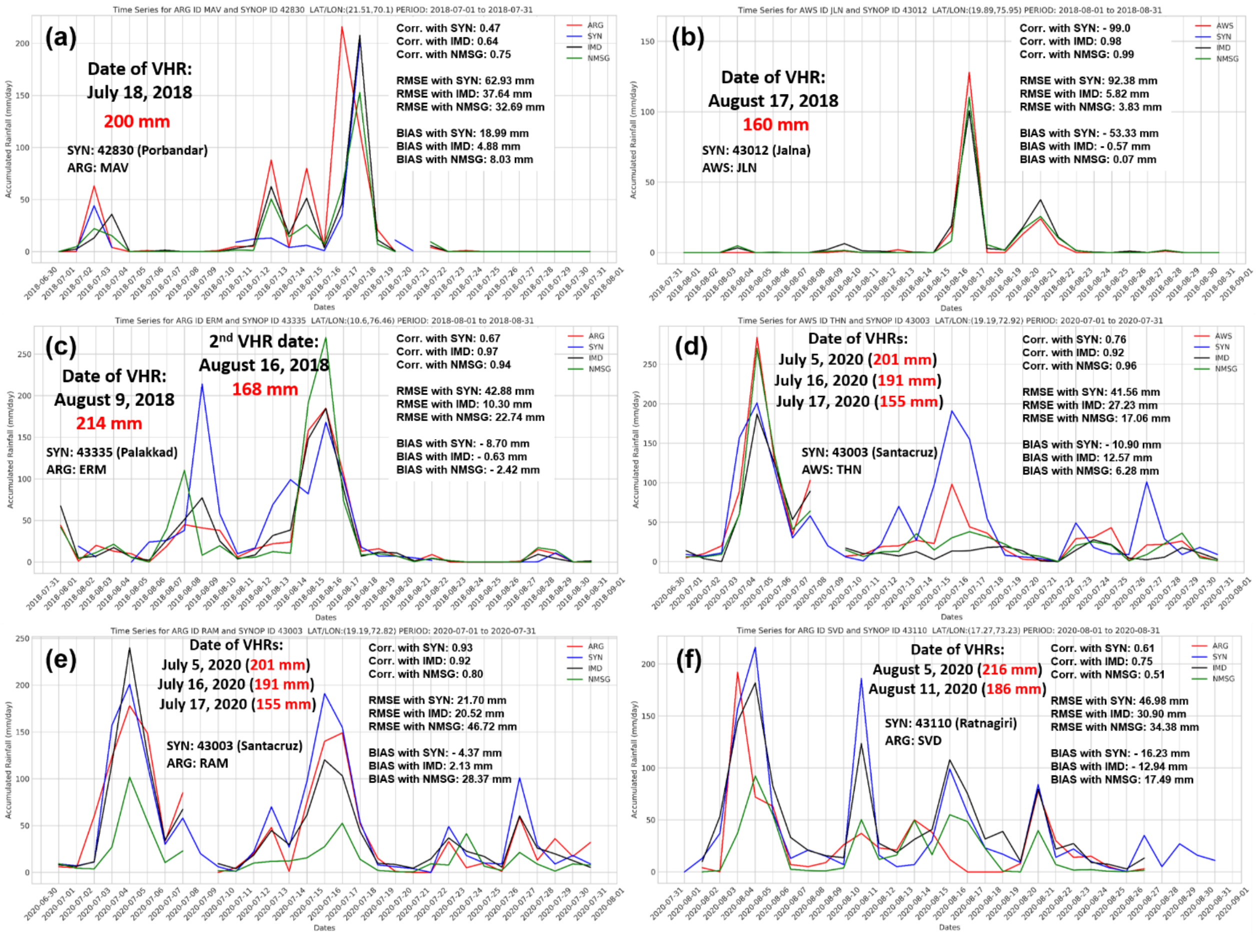


**Fig. 12C** (a-f) Time-series plots for AWS and ARG rainfall along with SYNOP or gridded (IMD/MSG) rainfall for matching VHR matching cases.

In most of the VHR events, the AWS/ARG stations have captured the event (as indicated by SYNOP rainfall) pretty well but in some cases, if there are two or more events have occurred in a month, some AWS/ARG failed to capture all the VHR events. Figure 12C (a-f) is self-explanatory and statistical details are given within each of the figure itself.

### *3.3.4* AWS/ARG not capturing Very Heavy Rainfall observations

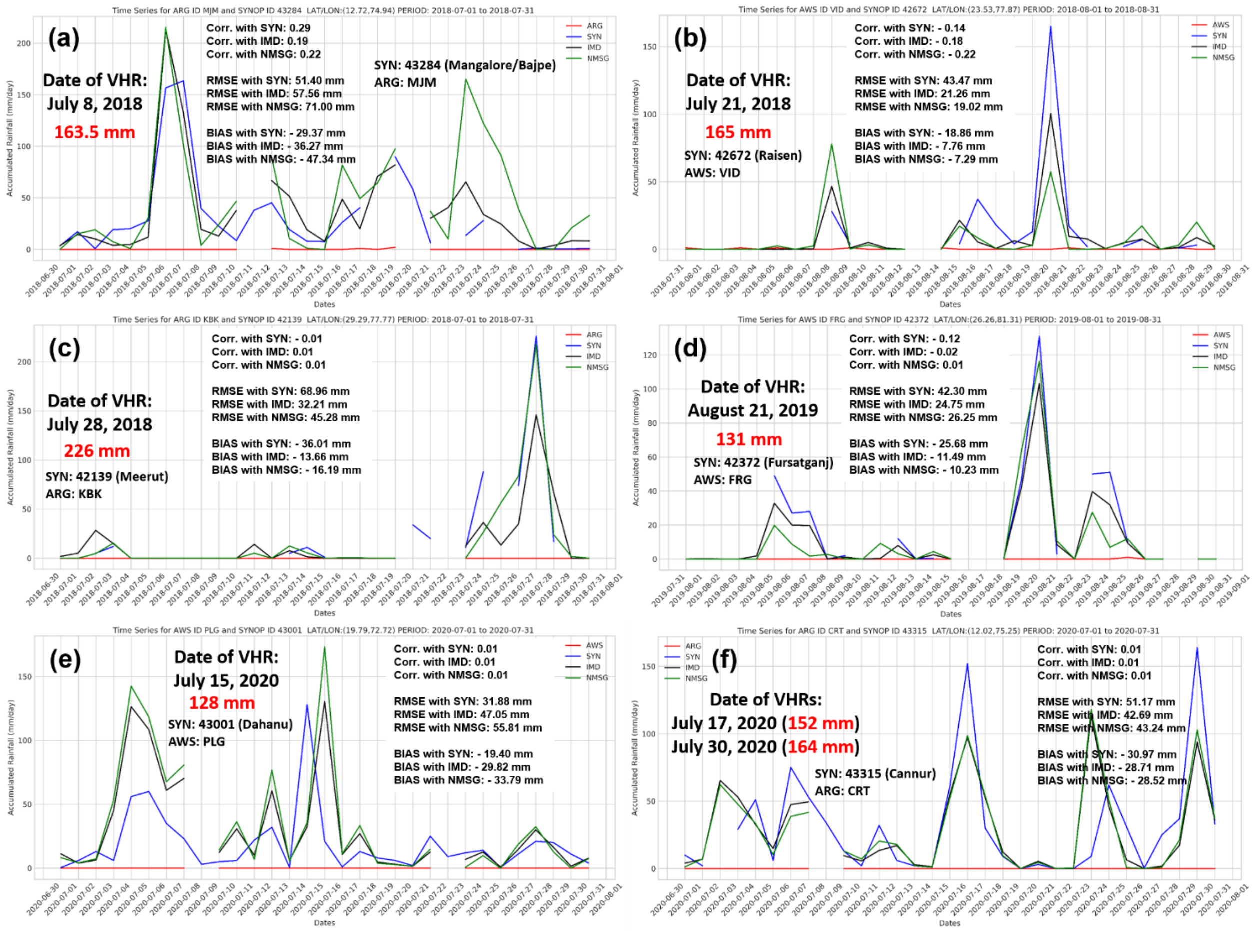


**Fig. 12D** (a-f) Time-series plots for AWS and ARG rainfall along with SYNOP or gridded (IMD/MSG) rainfall for non-matching VHR matching cases.

Similar to Figure 12B, Figure 12D represents the time-series plots for the AWS/ARG rainfall observations that has completely failed to capture the in-situ (SYNOP) or gridded rainfall observations (IMD/MSG) during July and August of 2018-2020 containing VHR events. In all the cases of these VHR events during core-monsoon of 2018-2020, both AWS and ARG stations has completely failed to capture the extreme event and not even they have captured any rainfall occurrences as shown by SYNOP/IMD/MSG observations. Statistical details are given within each of the figure itself and is self-explanatory. AWS or ARG stations have reported consistently zero rainfall occurrences in all of these cases while very heavy rainfall events have occurred. Thus, this is one of the primary aim of our study to assess and report the non-functional stations so that they get rectified after proper maintenance of sensor checks and calibration.

## 4. Monitoring of AWS and ARG rainfall during monsoon

It is important to monitor the quality or performance of AWS and ARG rainfall data over this vast network in both real time and non-real time activity. As discussed in the validation part, there are a huge number of AWS/ARG stations which are either not reporting daily or reporting erroneous data. It is of utmost necessary to identify those stations through proper real-time monitoring of AWS and ARG rainfall network in order to calibrate the bad sensors or repair those to utilizable condition. Extreme rainfall over the globe is increasing, particularly Indian region (Roxy et al., 2017; Saha et al., 2020). Hence, it is required to monitor the performance of the AWS and ARG stations reporting rainfall observations in real time in order to issue forecasts for the upcoming extreme rainfall events occurring over the Indian landmass. Also, the monitoring of the AWS/ARG stations are also very much important to agro-meteorology to aviation sectors. For this purpose, a real-time quality monitoring of AWS and ARG stations over the Indian region has been developed in NCMRWF, based on the validation results for flagging each individual observations centred on computed scores over previous 15 days. Rainfall reported by each AWS/ARG stations is attached with a flags ranging from 0 to 9, with 0 as the best. Flag 7-9 are assigned to irregular stations. Flags 1-6 indicates different qualities and the rainfall data with these flags can be chosen by end-user for particular applications.

**FLAG DETERMINATION: MINIMUM BIAS PROCESS**

Corr_SYN = C1, Corr_IMD = C2, Corr_NMSG = C3,
BIAS_SYN = B1, BIAS_IMD = B2, BIAS_NMSG = B3,
Mean_AWS/ARG = M0, Mean_SYN = M1, Mean_IMD = M2, Mean_NMSG = M3
R = 30% of the mean AWS/ARG rainfall
Flag = F

**STEP 1: When SYNOP is PRESENT,**

**Corollary 1:** If C1 ≥ 0.7:
- (a) If min(B1, B2, B3) ≤ R, then F = 0
- (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 3

**Corollary 2:** If 0.6 ≤ C1 < 0.7:
- (A) If max (C2, C3) ≥ 0.7:
  - (a) If min(B1, B2, B3) ≤ R, then F = 1
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 4
- (B) If 0.6 ≤ max (C2, C3) < 0.7:
  - (a) If min(B1, B2, B3) ≤ R, then F = 2
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 5
- (C) If 0.5 ≤ max (C2, C3) < 0.6:
  - (a) If min(B1, B2, B3) ≤ R, then F = 3
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 6
- (D) If max (C2, C3) < 0.5: then, F = 6 [NO OTHER CHECKS ARE DONE]

**Corollary 3:** If C1 < 0.6:
- (A) If max (C2, C3) ≥ 0.7:
  - (a) If min(B1, B2, B3) ≤ R, then F = 1
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 4
- (B) If 0.6 ≤ max (C2, C3) < 0.7:
  - (a) If min(B1, B2, B3) ≤ R, then F = 2
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 5
- (C) If 0.5 ≤ max (C2, C3) < 0.6:
  - (a) If min(B1, B2, B3) ≤ R, then F = 3
  - (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 6
- (D) If max (C2, C3) < 0.5: then, F = 6 [NO OTHER CHECKS ARE DONE]

**STEP 2: When SYNOP is NOT PRESENT,**

**Corollary 1:** If C2 ≥ 0.7 and C3 ≥ 0.7:
- (a) If min(B2, B3) ≤ R, then F = 0
- (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 3

**Corollary 2:** If max (C2, C3) ≥ 0.7:
- (a) If min(B2, B3) ≤ R, then F = 1
- (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 4

**Corollary 3:** If 0.5 ≤ max (C2, C3) < 0.7:
- (a) If min(B2, B3) ≤ R, then F = 2
- (b) If 30% ≤ min(B1, B2, B3) < 100% , then F = 5

**Corollary 4:** If max (C2, C3) < 0.5: then, F =6 [NO OTHER CHECKS ARE DONE]

**STEP 3: For IRREGULAR STATIONS,**

**Corollary 1:** If 9 ≤ Day_Count < 11: then F = 7

**Corollary 2:** If 7 ≤ Day_Count < 9: then F = 8

**Corollary 3:** If Day_Count < 7: then F = 9

During validation of rainfall data of individual AWS and ARG stations with in-situ/satellite/merged-satellite data, it has been obtained that many of the stations have more biases between two collocated stations taken for validation purpose. Hence, it is decided to proceed the flagging of individual stations which have minimum biases. The methodology adopted for assigning flags to individual stations is described above in a flow diagram.

For monitoring purpose, figures with assignment of flags from 0 to 9 is plotted daily based on the locus of each AWS and ARG stations over the Indian landmass as shown in Figure 13. Figure 13 (A-B) represents the distribution of quality flags on 23 August 2020 and 27 August 2020 respectively over the Indian map. Stations with “Flag 0” to “Flag 5” are treated as considerable stations to be taken for research as well as operational purpose since these stations are sorted based on high correlation and minimum bias criteria. On 23 August 2020, out of 270 AWS stations, a totality of 124 stations are flagged as “Flag 0” to “Flag 5”, while 74 stations are flagged as “Flag 6” following the corollaries of STEP-1 and STEP-2 of Quality Flagging and these stations are not considered to be taken for further research or operational purpose (Figure 13A a-b). 72 stations out of total 270 AWS stations are reporting irregular rainfall observations and are flagged as “Flag 7” to “Flag 9” following the STEP-3 criteria of Quality Flagging (Figure 13A c). Similarly, out of 335 ARG stations, a totality of 96 stations are flagged as “Flag 0” to “Flag 5”, while 101 stations are flagged as “Flag 6” (Figure 13A d-e). 138 stations out of total 335 ARG stations are reporting irregular rainfall observations and are flagged as “Flag 7” to “Flag 9” and are reported as irregular stations (Figure 13A f). Similar to Figure 13A, a similar plot is shown for another date (27 August 2020) where 109 stations out of 268 AWS stations and 88 stations out of 334 ARG stations are flagged as “Flag 0” to “Flag 5”, while 68 AWS and 87 ARG stations respectively are flagged as “Flag 6”. 91 AWS and 159 ARG stations respectively as flagged as “Flag 7” to “Flag 9” based on the Quality Flagging criteria (Figure 13B a-f). Daily quality flagging is done for monitoring purpose in order to identify the considerable and not-recommended AWS/ARG stations for further research or operational purposes. A snapshot of the table containing the assigned Quality Flags for 23 August 2020 is given in Appendix Table 5.

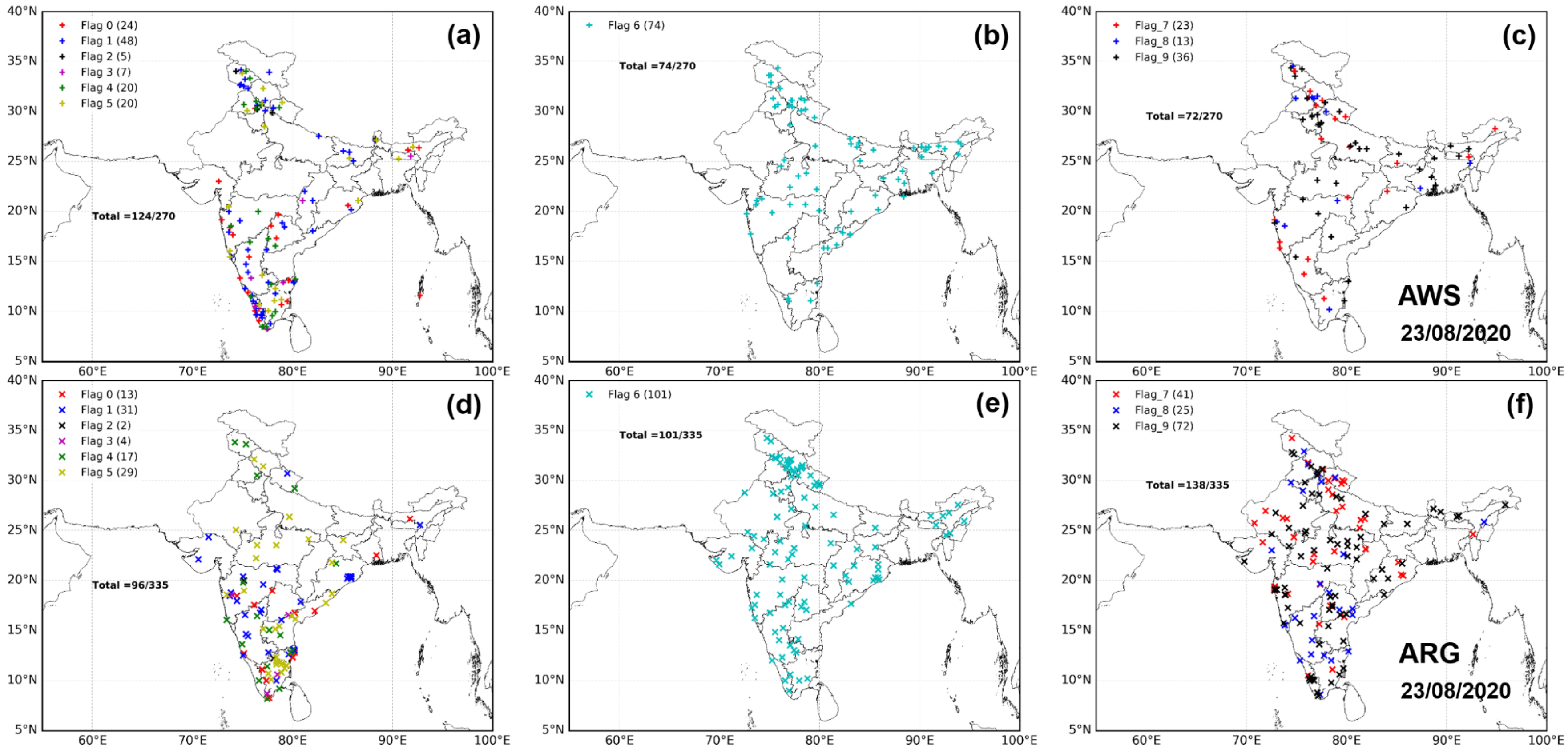


**Fig. 13A.** Distribution of Quality Flags on 23 August 2020 over the Indian landmass indicating (a) “Flag 0” to “Flag 5”, (b) “Flag 6”, and (c) “Flag 7” to “Flag 9” for AWS stations reporting rainfall observations. (d)-(f) Same as before but for ARG stations.

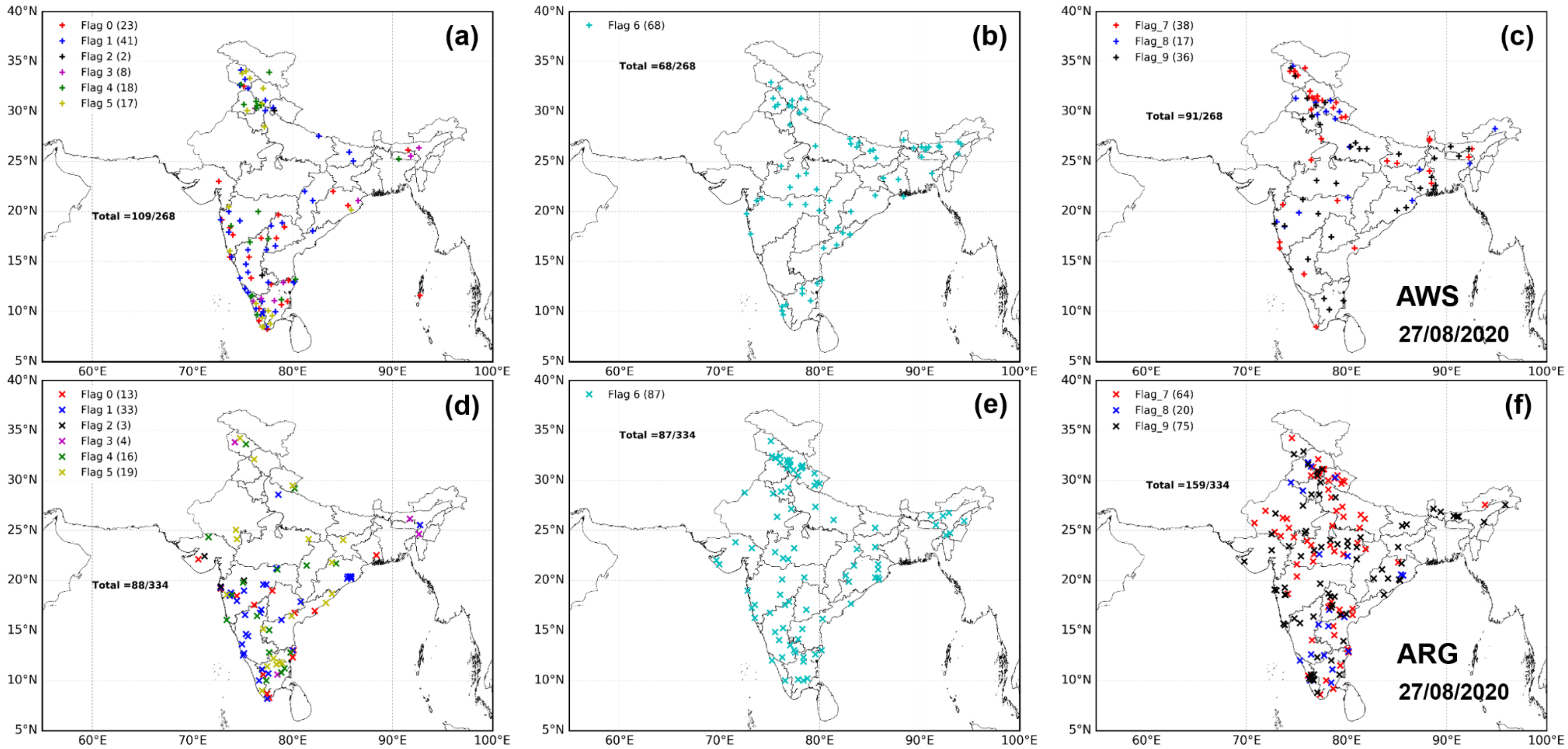


**Fig. 13B.** Distribution of Quality Flags on 27 August 2020 over the Indian landmass indicating (a) “Flag 0” to “Flag 5”, (b) “Flag 6”, and (c) “Flag 7” to “Flag 9” for AWS stations reporting rainfall observations. (d)-(f) Same as before but for ARG stations.

## 5. Summary and Conclusions

In-situ observational networks, in spite of their limited coverage continue to be the backbone of the observing system. Automatic weather stations such as AWSs and ARGs are complementing the SYNOP Observations and are strong means to provide the data from remote locations. Under the IMD Modernization Programme Phase-I, the automated network expanded and a totality of around 675 AWS and 1350 ARG stations (New IMD-AWS/ARG network) were established for operational use. The main aim of this study is to monitor and assess the quality of rainfall observed by these AWS and ARG stations during core-monsoon season (July and August) of 2018-2020. The main conclusions drawn from this study are as follows:

1. Though, AWS and ARG stations are reporting hourly rainfall observations over Indian region, however, for many stations report for all hours of a day are not transmitted.

2. In this study, we have considered 03 UTC observations for all stations as at 03 UTC, stations are reporting last 24-hourly accumulated rainfall. However, it has been found that many stations on several days have not reported 03 UTC observation also. Thus, the stations which have reported 03 UTC observation atleast 24 days in a month (75%) are termed as regularly reporting stations and considered for validation purpose.

3. All these regularly reporting stations have to be used with quality control for any research application, since many of the regularly reporting stations are reporting unrealistic rainfall observations. These stations are validated against collocated SYNOP/gridded rainfall data.

4. In most of the cases, AWS or ARG stations are not collocated with LAND SYNOP stations, hence, the matching of rainfall observations by AWS/ARG stations need to be done with the in-situ gridded or merged satellite-gauge rainfall observations. However, it is well known that satellite estimated rainfall always underestimated the actual rainfall. Thus AWS/ARG needs to be

properly assessed and monitored periodically, especially during extreme weather seasons. This imposes the requirement of monitoring and validate AWS/ARG observations in real-time for assessing its quality.

5. Daily monitoring of AWS and ARG stations reporting rainfall observations and assigning quality flags from "Flag 0" to "Flag 9" to individual stations based on the Quality Flagging criteria will identify the "good and regularly reporting" stations as well as "non-considerable and irregular reporting" stations for research applications and NWP validation purpose.

**Acknowledgments**

Authors are thankful to IMD, Pune for providing the dataset of IMD gridded rainfall data as well as merged rainfall product through the webpages:

1. http://www.imdpune.gov.in/Seasons/Temperature/Rainfall/Rain_Download.html, and
2. http://www.imdpune.gov.in/Seasons/Temperature/gpm/Rain_Download.html respectively.

The GPM satellite estimated rainfall data were provided by the NASA/GSFC and JAXA. We also thankfully acknowledge the use of GPM data as an input for MSG rainfall data in our study.

**References**

Das Gupta, M. and Rani, S.I., 2010. Online monitoring of Indian observations and their impact on NWP system. Vayu Mandal, 35-36 (1-4), 39-45.

Einfalt T, Jessen M, Quirmbach M. 2006. Can we check raingauge data automatically? In Proceedings 7th International Workshop on Precipitation in Urban Areas, St. Moritz, Switzerland, 7–10 December.

Estévez, J., Gavilán, P. and Giráldez, J.V., 2011. Guidelines on validation procedures for meteorological data from automatic weather stations. Journal of Hydrology, 402 (1-2), 144-154.

Estévez, J., Gavilán, P., García-Marín, A.P. and Zardi, D., 2015. Detection of spurious precipitation signals from automatic weather stations in irrigated areas. International Journal of Climatology, 35 (7), 1556-1568.

Haiden, T., Dahoui, M., Ingleby, B., de Rosnay, P., Prates, C., Kuscu, E., Hewson, T., Isaksen, L., Richardson, D., Zuo, H. and Jones, L., 2018. *Use of in situ surface observations at ECMWF*. European Centre for Medium Range Weather Forecasts.

Hubbard, K.G., Rosenberg, N.J. and Nielsen, D.C., 1983. Automated weather data network for agriculture. Journal of Water Resources Planning and Management, 109 (3), 213-222.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H.J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P. and Woods, R., 2018. A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. Journal of Hydrology, 564, 930-943.

McCulloch, J.S.G. and Strangeways, I.C., 1966. Automatic weather stations for hydrology. In *Proc. WMO Technical Conf. on Automatic Weather Stations, WMO Technical Note*, 82, pp. 263-264.

McNew, K.P., Mapp, H.P., Duchon, C.E. and Merritt, E.S., 1991. Sources and uses of weather information for agricultural decision makers. Bulletin of the American Meteorological Society, 72 (4), 491-498.

Michaelides, S.C., 2008. Precipitation: Advances in Measurement, Estimation and Prediction. Springer: Vienna, 540 p.

Mitra, A.K., Gupta, M.D., Singh, S.V. and Krishnamurti, T.N., 2003. Daily rainfall for the Indian monsoon region from merged satellite and rain gauge values: Large-scale analysis from real-time data. Journal of Hydrometeorology, 4 (5), 769-781.

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

Mohapatra, M., Kumar, N. and Ranalkar, M., 2011. Utility of automatic weather station (AWS) data for monitoring and prediction of cyclonic disturbances during 2010. IMD Met Monograph Synop. Met, (10), 189-203.

Mohapatra, M., 2014. Tropical cyclone forecast verification by India Meteorological Department for north Indian Ocean: A review. Tropical Cyclone Research and Review, 3 (4), 229-242.

Mohapatra, M., Nayak, D.P., Sharma, M., Sharma, R.P. and Bandyopadhyay, B.K., 2015. Evaluation of official tropical cyclone landfall forecast issued by India Meteorological Department. Journal of Earth System Science, 124 (4), 861-874.

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

Prasad, V.S., Johny, C.J. and Sodhi, J.S., 2016. Impact of 3DVar GSI-ENKF hybrid data assimilation system. Journal of Earth System Science, 125 (8), 1509-1521.

Ranalkar, M.R., Mishra, R.P., Anjan, A. and Krishnaiah, S., 2012. Network of automatic weather stations: Pseudo random burst sequence type. Mausam, 63 (4), 587-606.

Ranalkar, M.R., Gupta, M.K., Mishra, R.P., Anjan, A. and Krishnaiah, S., 2014. Network of automatic weather stations: time division multiple access type. Mausam, 65 (3), 393-406.

Ranalkar, M.R., Anjan, A., Mishra, R.P., Mali, R.R. and Krishnaiah, S., 2015. Development of operational near real-time network monitoring and quality control system for implementation at AWS data receiving earth station. Mausam, 66 (1), 93 –106.

Roxy, M.K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., Terray, P. and Rajeevan, M., 2017. A threefold rise in widespread extreme rain events over central India. Nature Communications, 8 (1), 1-11.

Saha, U., Singh, T., Sharma, P., Gupta, M.D. and Prasad, V.S., 2020. Deciphering the extreme rainfall scenario over Indian landmass using satellite observations, reanalysis and model forecast: Case studies. Atmospheric Research, 240, 104943.

Sharma, K., Ashrit, R., Bhatla, R., Mitra, A.K., Iyengar, G.R. and Rajagopal, E.N., 2017. Skill of predicting heavy rainfall over India: improvement in recent years using UKMO global model. Pure and Applied Geophysics, 174 (11), 4241-4250.

Simpson, I.R. and Jones, P.D., 2014. Analysis of UK precipitation extremes derived from Met Office gridded data. International Journal of Climatology, 34 (7), 2438-2449.

Singh, J., Sharma, P., Verma, A., Choudhary, S.K. and Gupta, M.D., 2018. Monitoring of In-situ Observations over Indian Region. Vayu Mandal, 44 (1), 31-37.

Singh, T., Saha, U., Prasad, V.S. and Gupta, M.D., 2021. Assessment of newly-developed high resolution reanalyses (IMDAA, NGFS and ERA5) against rainfall observations for Indian region. Atmospheric Research, 259, 105679.

Sridevi, C., Singh, K.K., Suneetha, P., Durai, V.R. and Kumar, A., 2020. Rainfall forecasting skill of GFS model at T1534 and T574 resolution over India during the monsoon season. Meteorology and Atmospheric Physics, 132 (1), 35-52.

Zahumenský, I., 2004. Guidelines on quality control procedures for data from automatic weather stations. *World Meteorological Organization*, *Switzerland*, 955, 2-6.

## APPENDIX

### Table 1a. *Statistical details from Figure 7A*:

| Statistics from Figure No. 7A | | (a) AWS: ADL SYN: 43025 | (b) AWS: AKL SYN: 42933 | (c) AWS: RJG SYN: 42948 | (d) AWS: UNA SYN: 42077 | (e) ARG: TRI SYN: 42849 | (f) ARG: WAD SYN: 43063 | (g) ARG: MTT SYN: 43315 | (h) ARG: PRM SYN: 43193 |
|---|---|---|---|---|---|---|---|---|---|
| Correlation with | SYNOP | 0.99 | 0.88 | 0.95 | 0.61 | 0.95 | 0.98 | 0.89 | 0.76 |
| | IMD | 0.98 | 0.91 | 0.89 | 0.42 | 0.99 | 0.95 | 0.72 | 0.92 |
| | NMSG | 0.96 | 0.93 | 0.95 | 0.70 | 0.99 | 0.89 | 0.76 | 0.88 |
| RMSE (mm) with | SYNOP | 4.15 | 6.37 | 7.95 | 16.61 | 10.40 | 4.93 | 13.68 | 9.53 |
| | IMD | 6.98 | 5.08 | 8.30 | 22.13 | 2.96 | 2.57 | 22.04 | 7.13 |
| | NMSG | 4.61 | 4.81 | 7.54 | 12.71 | 4.07 | 3.32 | 22.61 | 6.84 |
| BIAS (mm) with | SYNOP | - 2.67 | - 0.40 | - 3.71 | - 5.41 | - 1.08 | - 2.99 | 3.68 | 0.28 |
| | IMD | - 3.81 | 0.45 | - 0.35 | - 7.90 | 0.14 | - 1.07 | 7.87 | - 4.24 |
| | NMSG | 0.56 | 1.40 | - 3.45 | - 4.34 | 0.02 | - 0.58 | 1.72 | - 1.77 |

### Table 1b. *Statistical details from Figure 7B*:

| Statistics from Figure No. 7B | | (a) AWS: KNP SYN: 42367 | (b) AWS: HYD SYN: 43128 | (c) AWS: KNP SYN: 42367 | (d) AWS: SAT SYN: 43113 | (e) ARG: PHN SYN: 43063 | (f) ARG: ALP SYN: 42807 | (g) ARG: RAM SYN: 43003 | (h) ARG: SVD SYN: 43110 |
|---|---|---|---|---|---|---|---|---|---|
| Correlation with | SYNOP | 0.88 | 0.91 | 0.97 | 0.99 | 0.71 | 0.93 | 0.97 | 0.74 |
| | IMD | 0.84 | 0.64 | 0.92 | 0.98 | 0.92 | 0.54 | 0.97 | 0.92 |
| | NMSG | 0.86 | 0.91 | 0.96 | 0.99 | 0.91 | 0.14 | 0.97 | 0.75 |
| RMSE (mm) with | SYNOP | 34.29 | 9.29 | 3.65 | 4.67 | 14.89 | 13.12 | 12.44 | 44.55 |
| | IMD | 24.61 | 8.04 | 5.21 | 11.19 | 9.91 | 22.64 | 14.45 | 23.45 |
| | NMSG | 28.66 | 4.74 | 3.43 | 2.96 | 12.03 | 27.02 | 19.94 | 54.56 |
| BIAS (mm) with | SYNOP | - 14.61 | - 1.61 | - 1.67 | 2.75 | 12.61 | - 6.82 | 1.93 | 16.36 |
| | IMD | - 10.75 | 0.02 | 0.83 | 6.32 | 3.97 | - 4.16 | - 1.72 | 7.98 |
| | NMSG | - 12.33 | - 0.04 | 0.21 | 1.27 | - 6.31 | 5.65 | 5.39 | 31.35 |

### Table 1c. *Statistical details from Figure 7C*:

| Statistics from Figure No. 7C | | (a) AWS: ADL SYN: 43025 | (b) AWS: MLD SYN: 43193 | (c) AWS: GDG SYN: 43201 | (d) AWS: SAT SYN: 43113 | (e) ARG: RAM SYN: 43003 | (f) ARG: SNP SYN: 43063 | (g) ARG: MHL SYN: 43081 | (h) ARG: SNP SYN: 43063 |
|---|---|---|---|---|---|---|---|---|---|
| Correlation with | SYNOP | 0.98 | 0.87 | 0.99 | 0.99 | 0.93 | 0.99 | 0.89 | 0.98 |
| | IMD | 0.99 | 0.75 | 0.85 | 0.95 | 0.92 | 0.75 | 0.86 | 0.92 |
| | NMSG | 0.97 | 0.86 | 0.91 | 0.98 | 0.80 | 0.21 | 0.98 | 0.56 |
| RMSE (mm) with | SYNOP | 4.43 | 28.95 | 0.82 | 1.78 | 21.70 | 2.97 | 8.78 | 2.74 |
| | IMD | 2.44 | 31.45 | 4.11 | 8.46 | 20.52 | 7.60 | 9.16 | 4.65 |
| | NMSG | 3.55 | 24.54 | 3.01 | 4.18 | 46.72 | 11.77 | 3.14 | 9.90 |
| BIAS (mm) with | SYNOP | - 2.48 | - 19.64 | - 0.28 | 0.74 | - 4.37 | 0.33 | - 2.91 | - 0.84 |
| | IMD | - 0.56 | - 19.11 | 0.98 | 4.19 | 2.13 | 0.34 | - 3.37 | - 1.08 |
| | NMSG | - 1.07 | - 14.77 | 0.46 | 0.85 | 28.37 | 1.06 | - 0.90 | 1.57 |

Table 2a. *Statistical details from Figure 8A*:

| Statistics from Figure No. 8A | | (a) AWS: GOA SYN: 43192 | (b) AWS: SAT SYN: 43113 | (c) AWS: HSP SYN: 42077 | (d) AWS: PLG SYN: 43001 |
|---|---|---|---|---|---|
| Correlation with | SYNOP | 0.67 | 0.55 | 0.50 | 0.91 |
| | IMD | 0.68 | 0.60 | 0.91 | 0.84 |
| | NMSG | 0.67 | 0.53 | 0.76 | 0.84 |
| RMSE (mm) with | SYNOP | 30.97 | 20.26 | 70.72 | 72.16 |
| | IMD | 25.74 | 14.38 | 21.66 | 89.13 |
| | NMSG | 23.32 | 22.29 | 30.17 | 107.70 |
| BIAS (mm) with | SYNOP | - 12.71 | - 9.38 | - 28.00 | - 41.86 |
| | IMD | - 10.53 | - 3.34 | 3.67 | - 52.11 |
| | NMSG | - 7.75 | - 10.59 | 4.80 | - 57.26 |

Table 2b. *Statistical details from Figure 8B*:

| Statistics from Figure No. 8B | | (a) ARG: ERM SYN: 43335 | (b) ARG: BAM SYN: 42693 | (c) ARG: GNR SYN: 43014 | (d) ARG: MUD SYN: 42971 |
|---|---|---|---|---|---|
| Correlation with | SYNOP | 0.67 | 0.20 | 0.64 | 0.59 |
| | IMD | 0.97 | 0.67 | 0.94 | 0.88 |
| | NMSG | 0.94 | 0.81 | 0.99 | 0.65 |
| RMSE (mm) with | SYNOP | 42.88 | 16.27 | 15.28 | 30.74 |
| | IMD | 10.30 | 15.34 | 3.75 | 22.11 |
| | NMSG | 22.74 | 18.35 | 2.32 | 19.09 |
| BIAS (mm) with | SYNOP | - 8.70 | - 9.68 | - 3.30 | - 10.82 |
| | IMD | - 0.63 | - 10.59 | - 0.46 | - 9.57 |
| | NMSG | - 2.42 | - 8.79 | 0.64 | - 6.57 |

Table 3a. *Statistical details from Figure 9A*:

| Statistics from Figure No. 9A | | (a) AWS: DSL | (b) AWS: DSL | (c) AWS: JLN | (d) AWS: TNI | (e) AWS: JMU | (f) AWS: RAH |
|---|---|---|---|---|---|---|---|
| Correlation with | IMD | 0.95 | 0.97 | 0.98 | 0.96 | 0.89 | 0.54 |
| | NMSG | 0.22 | 0.61 | 0.99 | 0.95 | 0.98 | 0.97 |
| RMSE (mm) with | IMD | 9.90 | 17.66 | 5.82 | 4.28 | 12.35 | 20.15 |
| | NMSG | 35.17 | 54.23 | 3.83 | 5.65 | 7.95 | 10.96 |
| BIAS (mm) with | IMD | - 1.89 | - 10.46 | - 0.57 | - 0.20 | 2.04 | 2.24 |
| | NMSG | 21.05 | 27.79 | 0.07 | - 1.06 | 1.24 | - 4.33 |

Table 3b. *Statistical details from Figure 9B*:

| Statistics from Figure No. 9B | | (a) ARG: BJJ | (b) ARG: CLS | (c) ARG: JWL | (d) ARG: RRH | (e) ARG: DAS | (f) ARG: BYR |
|---|---|---|---|---|---|---|---|
| Correlation with | IMD | 0.78 | 0.99 | 0.89 | 0.96 | 0.87 | 0.79 |
| | NMSG | 0.89 | 0.99 | 0.81 | 0.86 | 0.75 | 0.46 |
| RMSE (mm) with | IMD | 25.14 | 4.67 | 16.29 | 7.98 | 25.36 | 24.21 |
| | NMSG | 17.93 | 6.55 | 21.27 | 16.18 | 38.19 | 38.54 |
| BIAS (mm) with | IMD | 3.36 | 0.97 | - 0.35 | - 4.27 | - 2.10 | 3.30 |
| | NMSG | - 0.45 | 1.59 | - 1.37 | - 5.57 | 24.14 | 19.52 |

**Table 4a.** ***Statistical details from Figure 10A*:**

| Statistics from Figure No. 10A | | (a) AWS: KJT | (b) AWS: GYL | (c) AWS: CTH | (d) AWS: KWR |
|---|---|---|---|---|---|
| Correlation with | IMD | 0.73 | 0.29 | 0.51 | 0.71 |
| | NMSG | 0.69 | 0.15 | 0.57 | 0.86 |
| RMSE (mm) with | IMD | 34.29 | 19.64 | 21.68 | 6.97 |
| | NMSG | 35.47 | 18.12 | 28.09 | 6.57 |
| BIAS (mm) with | IMD | - 7.75 | - 5.77 | - 8.47 | - 3.30 |
| | NMSG | - 0.74 | - 3.04 | - 9.58 | - 2.69 |

**Table 4b.** ***Statistical details from Figure 10B*:**

| Statistics from Figure No. 10B | | (a) ARG: CND | (b) ARG: DTR | (c) ARG: HGR | (d) ARG: KCB |
|---|---|---|---|---|---|
| Correlation with | IMD | 0.67 | 0.48 | 0.74 | 0.56 |
| | NMSG | 0.57 | 0.72 | 0.40 | 0.67 |
| RMSE (mm) with | IMD | 14.17 | 14.16 | 13.64 | 41.25 |
| | NMSG | 16.73 | 6.34 | 18.50 | 28.77 |
| BIAS (mm) with | IMD | - 5.59 | - 5.50 | - 4.22 | - 21.40 |
| | NMSG | - 5.08 | - 2.38 | - 0.88 | - 14.68 |

**Table 5.** ***A snapshot of the details of Quality Flags for a particular date (23 August 2020)*****:**

| AWS_STN | LAT | LON | FLAGS | | ARG_STN | LAT | LON | FLAGS |
|---|---|---|---|---|---|---|---|---|
| ABL | 30.2 | 76.46 | 2 | | AAD | 17.57 | 76.56 | 6 |
| ADA | 10.8 | 76.36 | 1 | | ABG | 21.51 | 83.43 | 6 |
| ADL | 19.69 | 78.58 | 0 | | ABM | 31.61 | 76.16 | 8 |
| ADT | 11.01 | 79.49 | 0 | | ADO | 15.65 | 77.27 | 7 |
| AGD | 19.89 | 75.25 | 6 | | AFT | 30.5 | 76.46 | 4 |
| AHC | 34.24 | 75.55 | 9 | | AHI | 16.76 | 75.05 | 6 |
| AHM | 23.03 | 72.62 | 0 | | AKD | 14.54 | 78.78 | 4 |
| AHW | 20.7 | 73.63 | 6 | | AKK | 17.57 | 76.16 | 0 |
| AKL | 20.7 | 77.07 | 6 | | ALP | 22.52 | 88.38 | 0 |
| ALG | 28.28 | 94.84 | 7 | | ALV | 15.15 | 78.28 | 5 |
| AMG | 19.09 | 74.74 | 1 | | AMM | 18.78 | 78.28 | 8 |
| ANK | 17.67 | 83.03 | 6 | | ANC | 8.98 | 76.96 | 6 |
| ANR | 23.83 | 91.21 | 6 | | ANN | 13.03 | 80.2 | 1 |
| ANT | 17.77 | 83.03 | 6 | | ANU | 31.11 | 77.77 | 6 |
| ARY | 11.11 | 79.79 | 9 | | ARV | 16.26 | 74.84 | 8 |
| ASW | 11.11 | 79.09 | 6 | | ASH | 23.03 | 76.76 | 9 |
| ATN | 33.63 | 75.15 | 6 | | ATG | 20.5 | 85.65 | 7 |
| AUR | 26.56 | 79.59 | 6 | | ATI | 26.16 | 81.81 | 7 |
| AYN | 28.48 | 77.17 | 5 | | AUN | 19.59 | 77.07 | 1 |
| BBS | 20.2 | 85.85 | 6 | | AWG | 16.06 | 73.43 | 4 |
| BDA | 34.54 | 74.64 | 8 | | BAB | 27.17 | 77.17 | 6 |
| BDK | 30.7 | 76.86 | 4 | | BAC | 29.19 | 80.2 | 4 |
| BEP | 24.04 | 88.28 | 6 | | BAM | 23.13 | 83.73 | 6 |
| BGG | 26.36 | 90.6 | 6 | | BBD | 19.09 | 72.82 | 9 |
| BGK | 16.16 | 75.55 | 1 | | BBN | 20.2 | 85.85 | 1 |
| BGM | 25.25 | 90.6 | 5 | | BDH | 32.92 | 75.75 | 8 |
| BHK | 21.11 | 86.56 | 5 | | BDM | 27.57 | 95.85 | 9 |
| BHT | 27.27 | 77.47 | 7 | | BDU | 23.43 | 74.24 | 9 |
| BIH | 26.96 | 93.83 | 6 | | BEE | 32.02 | 76.76 | 6 |
| BJR | 31.51 | 77.07 | 8 | | BEI | 28.68 | 76.56 | 9 |
| BKE | 13.73 | 75.75 | 7 | | BET | 26.36 | 91.01 | 9 |
| BLH | 31.01 | 76.36 | 4 | | BGN | 20.2 | 85.45 | 1 |
| BLP | 31.31 | 76.76 | 8 | | BGS | 29.89 | 79.79 | 7 |
| BMH | 34.34 | 74.44 | 9 | | BIN | 20.6 | 85.45 | 7 |