# High-resolution typhoon precipitation integrations using satellite infrared observations and multi-source data

You Zhao[1,2], Chao Liu[1,2], Di Di[1,2], Ziqiang Ma[3], Shihao Tang[4]

[1]School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China.
[2]Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China.
[3]Institute of Remote Sensing and Geographical Information Systems, School of Earth and Space Sciences, Peking University, Beijing, 100871, China.
[4]Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center, China Meteorological Administration, Beijing 10008, China.

*Correspondence to*: Chao Liu (chao_liu@nuist.edu.cn)

**Abstract** Typhoon-related precipitation over land can result in severe disasters such as floods and landslides, and satellites are a valuable tool to estimate surface precipitation with high spatial-temporal resolutions. This study develops a high-resolution surface precipitation integration framework to combine observations from geostationary Fengyun-4A/ Himawari-8 (F4/H8) satellite radiometers, high-density rain-gauge observations or IMERG data and atmospheric reanalysis based on a random forest (RF) algorithm. The RF methods intend to integrate cloud and atmospheric features from radiometric observations and reanalysis information, and the intensity and spatial distribution of precipitation can be revealed by high-density rain-gauge or IMERG data. We take three typhoons made landfall in South China in 2018 as examples. Both F4-based and H8-based results using rain-gauge data as the predictand show excellent results, and the correlation coefficients (R) and probabilities of detection (POD) reach ~0.75 and ~0.95 respectively, both higher than satellite-only results. Meanwhile, when the IMERG data are used as the predictand, the corresponding R and POD drop to ~0.5 and 0.93, respectively, due to the uncertainties related to IMERG retrievals. By carefully choosing the predictors, our RF algorithms successfully summarize the information from satellite observations, surface observations and atmospheric reanalysis, resulting in precipitation results that are highly consistent with the actual ground observations. Our proposed integration framework can reconstruct hourly surface precipitation estimation datasets at high-spatial resolution for historical typhoon studies.

## 1 Introduction

Typhoon, also referred to as tropical cyclones, is a high-impact atmospheric phenomena, and cause some of the most significant socioeconomic damage due to their intense winds, immense storm surges and flood-inducing rainfall [Wang er al., 2019; Rappaport et al., 2014; Negri et al., 2005]. There are over eighty typhoons every year globally, and approximately one-third of which is generated in the Northwest Pacific (NWPAC) [Chan et al., 2005]. Furthermore, although the number of typhoons

in the NWPAC is exhibiting a decreasing trend with global warming, their average intensity is increasing, resulting in even more severe damage in coastal areas [Webster et al., 2005; Emanuel et al., 2013; Kang et al., 2016; Cai et al., 2004; Ho et al., 2021]. Unfortunately, because of the high spatial-temporal variability and complex physical processes, precipitation still accounts for one of the largest uncertainties in the forecasting of tropical cyclones [Su et al., 2012; Tu et al., 2013], and its high-quality observations are also limited. Thus, precipitation observations with a high spatial-temporal resolution during typhoon periods play an important role in not only researches on typhoon precipitation characteristics but also in disaster prevention.

Ground rain-gauges, weather radars and satellites are commonly used to acquire precipitation measurements [Villarini et al., 2008], and they all have their unique advantages and disadvantages. The ground-based rain-gauge is direct and accurate to measure local surface precipitation, while is limited by the station locations and coverages, especially in oceanic, mountainous and polar regions. [Susana et al., 2013; Gires et al., 2012; Rodriguez et al., 2019; Looper et al., 2012]. In contrast, microwave radars can survey larger areas and can better capture the spatial variability of rainfall fields; however, the accuracy of radar-based measurements is highly influenced by electromagnetic attenuation and the uncertainty in the relationship between the radar reflectivity factor and precipitation, particularly under extreme rainfall conditions [Marra et al., 2015; Bárdossy et al., 2017]. Satellite-based quantitative precipitation estimation (QPE) can be implemented on a large scale with a high spatial-temporal resolution, offering large scale capability with high spatial-temporal resolutions [Tang et al., 2016; Wang et al., 2018; Jozaghi et al., 2019], but quantitatively inferring the amount of surface precipitation from space is still a serious challenge, especially during typhoon periods. Nevertheless, with the continuous improvement of meteorological satellites, satellite-based QPE technologies have undergone considerable development [Boushaki et al., 2009; Nguyen et al., 2018]. Various models have been developed to generate satellite-based QPE products by relying on the relationships between passive infrared/microwave observations and precipitation; typical products include the Tropical Rainfall Measuring Mission Multisatellite Precipitation Analysis (TRMM TMA) [Huffman et al., 2007; Liu et al., 2016], Integrated Multisatellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG) [Huffman et al., 2015; Wang et al., 2017], Climate Prediction Center morphing technique (CMORPH) [Joyce et al., 2004] and Global Satellite Mapping of Precipitation (GsMaP) [Aonashi et al., 2009; Ushio et al., 2009]. At present, machine learning (ML) methods are widely used to establish the relationship between precipitation and satellite passive spectral observations as well [Albawi et al., 2017; Sehad et al., 2017; Min et al. 2018; Ahmed et al., 2020; Wang et al., 2021; Zhang et al., 2021]. Min et al. used a random forest (RF) algorithm to establish the relationship between spectral imager observations and numerical weather prediction results, and quite reasonable rainfall area and intensity can be obtained [Min et al., 2018]. Wang et al. used convolutional neural network (CNN) to establish a high-quality precipitation dataset based on Chinese FengYun geostationary satellite [Wang et al., 2017]. Most those studies intended to develop a general model for precipitations of all kinds globally or over large regions, while the different responses of precipitation caused by different kinds of cloudy and atmospheric conditions may limit their accuracy. Meanwhile, typhoons not only result in heavy and wide range precipitations, but also thick clouds that significantly influence/blocks satellites for surface variable observations. Considering the complex and unclear relationship between

2

precipitation and clouds, the QPE product particularly designed for typhoon precipitation remains limited, leaving a gap in accurate and high resolution surface precipitation. Meanwhile, FengYun-4A (F4) and Himawari-8 (H8) are new-generation geosynchronous equatorial orbit (GEO) satellites launched by the Chinese and Japanese meteorological agencies, respectively [Min et al., 2017; Bessho et al., 2016]. These two satellites cover a similar observational region of the NWPAC and a large

70    portion of East Asia, and their high temporal and spatial resolutions are favourable for continuously monitoring of NWPAC typhoons [Ma et al., 2021; Honda et al., 2018]. Thus, this study employs a ML technique to integrate F4/H8 radiometer observations and multi-source datasets to develop better surface precipitation integration algorithms particularly for typhoon precipitation, and, more importantly, investigates how multi-source (ground-based, satellite and reanalysis) data improve and contribute the surface precipitation integration differently.

75   **2 Data and Study area**

**2 Data**

To better estimate surface precipitation during the typhoon period, we would exert the advantages of multi-source data for different atmospheric variables, while also include only popular and public-available ones for the generality of the method. Thus, besides the F4/H8 radiometer measurements, we would consider ground-based observations, atmospheric reanalysis and

80    existing satellite-based surface precipitation estimations.

F4 and H8 can observe most of NWPAC and East Asia, while, over the region, South China is frequently influenced by the landfall of typhoons, resulting in serious disasters and casualties. The spectral radiometer onboard F4, namely the Advanced Geosynchronous Radiation Imager (AGRI), has six visible and near-infrared channels and eight thermal infrared channels. The Advanced Himawari Imager (AHI) onboard H8 has a similar channel design, and is equipped with two additional thermal

85    infrared channels. To develop an all-day available integration algorithm, we consider only the AGRI and AHI infrared channel radiances and their combinations. For a fair comparison between the F4/AGRI- and H8/AHI-based algorithms, only the channels that are similarly equipped on both instruments are used, including the two water vapor channels (WV1 and WV2) and four longwave infrared channels (LW1-LW4). The central wavelengths of each channel considered in our integration algorithm are shown in Table 1, and it becomes interesting whether such small differences on the channel characteristics as

90    well as those on channel spectral response functions and geolocation would influence the surface precipitation estimation. Passive spectral observations provide mostly the spatial distribution of clouds as well as cloud top properties, whereas surface precipitation is highly related to the atmospheric conditions. Thus, our algorithm also adopts atmospheric variables such as the column water vapor content, cloud water content, atmospheric profiles and so on. These variables are obtained from the state-of-the-art atmospheric reanalysis data, i.e., the European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis

95    v5 (ERA5), the latest atmospheric reanalysis product developed by ECMWF [Hersbach et al., 2020]. We consider nine variables from ERA5 related to typhoon precipitation to infer the general atmospheric conditions within the integration region [Min et al., 2018]. For example, the convective available potential energy (CAPE) is an indication of the instability (or stability)

3

of the atmosphere and can be used to assess the potential for the development of convection, which can lead to heavy rainfall, thunderstorms and other severe weather. The K-index, calculated from the temperature and dew point temperature in the lower part of the atmosphere, is a measure of the potential for a thunderstorm to develop. The total column rain water (TCRW) is the total amount of water in droplets of raindrop size (which can fall to the surface as precipitation) in a column extending from the surface of the Earth to the top of the atmosphere. Total precipitation (TP) is the accumulated liquid and frozen water (comprising both rain and snow) that falls to the Earth's surface. The total column liquid water (TCLW) is the total amount of supercooled water in a column extending from the surface of the Earth to the top of the atmosphere. Furthermore, four basic atmospheric variables are considered, namely, the relative humidity at 850 hPa (R850) and 950 hPa (R950) and the temperature at 850 hPa (T850) and 950 hPa (T950), which describe the humidity and temperature states in the lower atmosphere.

## 2.2 Study area

Accurate precipitation observations are needed as the references to train ML-based models, and thus are essential for high-quality integration. Two kinds of data are considered: high-density ground rain-gauge data and IMERG estimates. Offering some of the most reliable and fundamental precipitation observations, rain-gauge data are obtained from the National Meteorological Information Center of the China Metrological Administration (CMA). Due to the high frequency of rainfall and dense urban distribution in South China, automatic rain-gauge stations are densely distributed throughout the region. Within the land area of the studied region ($79 \times 10^4$ km$^2$), there are a total of 5024 rain-gauge stations, the distribution of which is shown in Figure 1. However, such a high density of ground-based observations may not always be feasible, so another type of data that is more commonly used in ML-based precipitation estimation is also considered, i.e., the IMERG final run-calibrated precipitation data. IMERG provide gridded precipitation estimation from PMW sensors on board various satellites in the GPM constellation and IR-based observations from geosynchronous satellites [Liu et al., 2016, Tang et al., 2016], and is used as one of the most reliable precipitation datasets. IMERG has a half-hourly temporal interval with a maximum rain rate of 50 mm/h and cover the area between the latitudes of 60°S and 60°N [Min et al., 2018].

This study investigates three typhoon events that made landfall in South China in 2018. Information on these three typhoon events is provided in Table 2, and the paths of the three typhoons are illustrated in Figure 1. We consider the area within the latitudes 15°-27°N and the longitudes 105°-125°E, including a large part of South China (15°-27°N; 105°-122.5°E) and NWPAC, which are completely covered by the observation areas of F4 and H8. We consider the whole processes from landing to dissipation of the three typhoons, model performance for a total of ~120 hours would be considered, resulting approximately 600,000 hourly precipitation observations from the ground rain-gauge stations.

4

**3 RF-based typhoon precipitation integration frameworks**

**3.1 Establishment of surface precipitation integration algorithm**

The RF algorithm is an ML method widely used in the inversion of meteorological elements and has been proven to perform well in applications such as the estimation of precipitation, detection of clouds, and inversion of PM2.5 concentrations [Oscar et al., 2020; Tan et al., 2021; Liu et al., 2021; Guo et al., 2021]. Thus, this study used this simple but promising RF algorithm to establish the nonlinear relationships among surface precipitation, satellite observations and atmospheric characteristics.

In the framework of RF algorithms, two types of data are utilized for the integration model: the predictor and the predictand. To study the influences of different satellite observations (F4/AGRI vs. H8/AHI) as the predictor and different precipitation data (rain-gauges vs. IMERG) as the predictand on the surface precipitation estimation algorithm, we established four independent surface precipitation integration models. The models will be referred to as F4-based and H8-based ones according to the satellite data used in the predictor, and, similarly, the models using rain-gauge observations and IMERG data as the predictand will be mentioned as RG-based and IM-based ones, respectively. Thus, a total of four models, i.e., F4-RG, F4-IM, H8-RG and H8-IM, are developed and studied. Table 3 differentiates these four models. It is worth noting that the only difference between the rain-gauge-based models and the IMERG-based models is the predictand used in the models, as will be described in more detail in the following.

The predictors for the surface precipitation integration algorithm include the geographic location, radiometer observations and atmospheric reanalysis data. By testing the performance of the RF-based models with different combinations of variables, our final models consider a total of 21 variables, as listed in Table 4. The rain-gauge observations and IMERG estimates are used to provide corresponding surface precipitation as model predictand. It should be noted that only F4/AGRI observations are used in F4-RG and F4-IM, and only H8/AHI observations are used in H8-RG and H8-IM. Considering the uncertainties related to IMERG estimated precipitation, we use ground rain-gauge data to validate all four models.

Figure 2 illustrates the general flowchart of our surface precipitation integration algorithm, including the model development and surface precipitation estimation. During the model development, both ERA5 data and satellite data are collocated with ground high-density rain-gauge data to obtain rain-gauge-based training datasets. During the collocation processes, the satellite data have a high spatial resolution (4 km for F4/AGRI and 5 km for H8/AHI), and the average radiances values of the nine satellite pixels (3×3 pixels) closest to each ground station are collocated to represent the satellite data corresponding to that station. In contrast, ERA5 has a much lower spatial resolution (0.25°), so the atmospheric variables from the ERA5 grid point closest to each ground station are used. In other words, all the aforementioned data are collocated with the ground rain-gauge stations, and the resulting datasets are used for further training and model development. For a fair comparison, the IMERG precipitation estimates are also collocated with the rain-gauge observations, resulting in a dataset with the same spatial-temporal sampling interval as that of the rain-gauge-based models for model training. Thus, the F4-IM/H8-IM and F4-RG/H8-RG models have completely the same predictor dataset, and differ only with regard to the predictand, i.e., replacing rain-gauge observations with the corresponding IMERG estimates for F4-IM/H8-IM. With the training datasets obtained above, the RF

algorithm is used for the training phase, and the four surface precipitation integration models (F4-IM, H8-IM, F4-RG and H8-RG) are developed.

The right part of Figure 2 depicts the procedure for the pixel-level component of our surface precipitation integration algorithm. At this stage, linear interpolation is employed to obtain the ERA5 atmospheric variables at each F4/H8 grid point. Then, the surface precipitation integration results from the direct RF-based integration models are classified into two sets: pixels with precipitation and those without precipitation. Considering that the minimum precipitation resolution of a rain-gauge is 0.1 mm/h, pixels with a fused rainfall rate below 0.1 mm/h are defined as those without precipitation, while the integration results are retained for pixels with an estimated rainfall rate greater than 0.1 mm/h. This threshold of 0.1 mm/h ensures a high probability of distinguishing precipitation from non-precipitation.

To fairly evaluate the performance of our surface precipitation integration products, we consider the rain-gauge observations to be the "truth" and perform 10-fold cross-validation (10-cv) to ensure the independence between the training and testing datasets. In other words, the original training dataset is evenly divided into ten parts, one of which is taken as the testing dataset each time (without repetition), while the remaining nine are taken as the training dataset; this process is repeated ten times. We adopt four popular parameters to quantify the model performance, including two categorical parameters (the probability of detection (POD) and false alarm ratio (FAR)) and two statistical parameters (the correlation coefficient (R) and root mean square error (RMSE)) between our results and the truth [Ebert et al., 2007; Mecikalski et al., 2008]. POD and FAR both range from 0 to 1, where larger POD and lower FAR values correspond to the better identification of precipitation events.

There are two important RF parameters that need to be considered: the number of trees to grow ($N_{tree}$) and the number of variables randomly sampled as candidates at each split ($M_{feature}$). Figure 3 shows the relationship of parameters $N_{tree}$ and max_features with the correlation coefficient (R), which is a measure of how well the RF models perform in testing datasets created by 10-cv. Based on the results, the value of R increases with the increase of $N_{tree}$ and $M_{feature}$, and tends to be stable when $N_{tree}$ is greater than 500 and $M_{feature}$ is greater than 10. To ensure that the four models are comparable, the RF parameters are fixed to be the same. Considering the accuracy, computational efficiency and comparability of the models, the number of trees to grow ($N_{tree} = 1000$) and the number of variables randomly sampled as candidates at each split ($M_{feature} = 10$) are both fixed for the four models.

## 3.2 Test and evaluation of the RF-based integration algorithm

Considering the distribution of precipitation throughout the year and the amount of damage caused by precipitation, the distribution of CMA ground rain-gauge stations is clearly uneven, with many more situated in eastern China than in western China. Because we employ ground rain-gauge observations for the integration, so their "resolution" is crucial for representing the spatial distribution of precipitation. To explore the impact of station density on the integration of typhoon precipitation, we select data from different numbers of stations to build training datasets of different sizes. The number of stations considered increases from 100 to 5000 with an interval of 100. Figure 4 shows the influence of the number of stations considered on the

model training results. The solid lines represent the rain-gauge-based model results, while the dashed lines represent the IMERG-based results; the blue and red colors denote the F4-based and H8-based results, respectively. The evaluation parameters of all four models exhibit similar trends with an increasing number of stations; i.e., the integration results become more accurate due to both additional input data and a better representation on precipitation spatial distribution. The POD values of the four models are all close to 1, and vary little with the number of stations, while the FAR of the four models decreases with an increase in the number of stations. The two statistical metrics fluctuate greatly when the number of stations is less than ~1000, i.e., $12.6 \times 10^{-4}/km^2$. When observations from more than 1000 stations are considered, the R/RMSE gradually increases/decreases with an increasing number of stations. In general, in the estimation of typhoon precipitation, when the number of stations covering the region exceed 1000, rain-gauge observations can generally reflect the spatial distribution of precipitation and thus can be used for surface precipitation integration. Of course, an increase in the number of stations does improve the integration performance, but only to a limited degree. Nevertheless, to ensure optimal model performance, this paper selects the data from all available stations.

The scatter plots in Figure 5 quantitatively compare the hourly precipitation estimated by our surface precipitation integration models with the 10-cv testing datasets, indicating only the generalization ability of the model. The two IMERG-based models (F4-IM and H8-IM) outperform the rain-gauge-based models (F4-RG and H8-RG) in the testing datasets according to the values of the POD (0.98 vs. 0.95, 0.97 vs. 0.95), FAR (0.25 vs. 0.46, 0.25 vs. 0.46), R (0.89 vs. 0.79, 0.89 vs. 0.79) and RMSE (1.25 mm/h vs. 2.19 mm/h, 1.25 mm/h vs. 2.17 mm/h). This may be because the satellite-based precipitation data of IMERG as the predictand may demonstrate better spatial consistency with the satellite observations than the rain-gauge observations. In contrast, due to the similar satellite channels used in the prediction variables, the results based on different satellites are very similar in both the rain-gauge-based and the IMERG-based models. Noted that the "better performance" here only illustrates the better correlations between the predictors and predictands, but is not a high accuracy of the final model estimations, which also depend on the accuracy of the predictands compared to the truth (see Section 4).

To better understand the surface precipitation integration models, Figure 6 shows the importance of all the predictors considered in the four models. The importance of each variable is given in the form of the mean accuracy decrease (%IncMSE), which can represent the relative contribution of each variable, and the sum of the importance of all variables is 100%. For F4-RG and H8-RG, the most important input variables are TCRW and TP, which represent the potential precipitable water in the atmosphere and the sum of large-scale precipitation and convective precipitation given by the atmospheric reanalysis. Geographical location data are also highly important, perhaps because the obvious spatial patterns of the typhoon precipitation could be captured by the RF-based training. The most important radiometer observations are a few BTs, such as BT12.0-BT10.7 for F4-RG and BT6.25-BT10.7 for H8-RG, both of which are related to the state of water vapor in the atmosphere. It is interesting that the important variables of F4-IM/H8-IM differs significantly from those of F4-RG/H8-RG. Since the IMERG data is a satellite-based precipitation estimation product (Min et al., 2018), there is a high consistency between IMERG data and satellite observation, which is also reflected in the importance of variables. For F4-IM and H8-IM, the importance ranking of variables such as BT8.5, BT6.2-BT10.7, BT6.2-BT11.2 and BT7.0-BT12.3 increased significantly, while the importance

225  <span style="color:red">ranking of atmospheric reanalysis data and geographic data decreased, resulting in the precipitation predicted by the two models being more dependent on satellite observation. Meanwhile, the higher rankings of reanalysis-based variables in the RG-based models indicate their reasonable performance for distributing accurate atmospheric states (i.e., surface precipitation from rain gauges).</span>

## <span style="color:red">4 Performance for typhoon precipitation integrations</span>

230  To evaluate the surface precipitation integration results, Figure 7 shows the ground rain-gauge observations and the results from the four integration models at three time steps (from top to bottom) during Typhoon Mangkhut, which made landfall at 0900 UTC on 16 Sep 2018. Three typical cases before the typhoon made landfall (0100 UTC on 16 Sep 2018), after the typhoon made landfall (1200 UTC on 16 Sep 2018) and as the typhoon dissipate (2300 UTC on 17 Sep 2018) are illustrated. Before the typhoon made landfall, the number of stations with precipitation on land was small, and heavy precipitations (e.g.,

235  stations with rain rates greater than 20 mm/h) was concentrated mainly over the coastline. After the typhoon made landfall, the precipitation area increased significantly, and both the extent of heavy precipitation area and the rain rates increased, with over 29 stations with recording rain rates exceeding 30 mm/h. Before the typhoon started to dissipate, the center of heavy rainfall disappeared, although there were still large regions of weak precipitation. At all three time steps, the four models consistently yield reasonable typhoon precipitation intensities and spatial distributions, especially the F4-RG and H8-RG,

240  whereas F4-IM and H8-IM slightly overestimate the precipitation extent and can hardly represent the heavy precipitation centers. <span style="color:red">First, the differences are introduced by the different predictands considered, which are clearly shown between the upper and lower panels. As expected, the precipitation predicted by the RG-based models is closer to the rain-gauge precipitation and the IM-based model results are closer to the IMERG precipitation. Meanwhile, the differences between the two RG-based models (the upper panels) are less significant, similarly for the two IM-based model results (the lower panels),</span>

245  <span style="color:red">because the F4 and H8 provide very similar satellite observations for the estimations.</span>

To quantitatively understand the abovementioned differences, Figure 8 illustrates the differences in hourly precipitation between our integration results and ground rain-gauge observations at the same three time steps in Figure 7. Warm colors indicate stations with surface precipitation integration results larger than the rain-gauge observations, while cold colors indicate those with underestimated surface precipitation integration results. The differences between the F4-RG/H8-RG results and

250  rain-gauge observations are mostly within 2 mm/h, significantly smaller than the differences from the F4-IM/H8-IM results. However, all four models tend to underestimate the precipitation intensity in the area of heavy precipitation, while overestimating the precipitation at the rain-gauge stations with relatively weak precipitation (rain rates below 5 mm/h). Ultimately, more than 70% of the surface precipitation was overestimated because there were far more stations experiencing weak precipitation than there were recording heavy precipitation within the studied region.

255  Table 5 summarizes the percentages of the precipitation that was overestimated by the four integration models during the three typhoon precipitation periods. The percentages in this table denote ratios of the numbers of our integration results larger than the true values (i.e., rain-gauge observations) to the total amount of data at the particular rainfall rate. As our precipitation

8

differences are strongly consistent with the rainfall rates, we divided the results into two groups: one group for stations with rainfall rates smaller than 5 mm/h and the other group for those with rainfall rates larger than 5 mm/h. For rain-gauge precipitation less than 5 mm/h, more than ~85% of these data are overestimated. However, for rain rates greater than 5 mm/h, only approximately 20% of ours results are larger than the true values, which means that the rate of underestimation of our models is approximately about 80%.

Figure 9 compares the daily precipitation of our four models during Typhoon Mangkhut on 16 Sep 2018. The daily precipitation results of based on the rain-gauge predictand (F4-RG and H8-RG, top panels) agree closely, and similar consistency is noticed between the F4-IM and H8-IM results (bottom panels). However, the rain-gauge-based results differ substantially from the IMERG-based results. The daily precipitation distributions of F4-RG and H8-RG indicate that during Typhoon Mangkhut, the precipitation over land was concentrated mainly in Guangdong Province, and the daily precipitation in some areas surpassed 200 mm/day. In contrast, F4-IM and H8-IM overestimate most of the daily precipitation on land but significantly underestimate the daily precipitation over 200 mm/day. It is worth noting that our models can give precipitation distribution on ocean as well, while its performance could hardly be fairly evaluated due to the lack of the ground-based observations.

The spatial distributions of the above biases in our surface precipitation integration results in daily scale are shown in Figure 10. In general, both the rain-gauge-based models and the IMERG-based models share considerable similarities in their spatial distribution of the bias. For the two rain-gauge-based models, the average errors at most stations are between -20 mm/day and 20 mm/day, while the errors in the two IMERG-based models are significantly larger than those in the two rain-gauge-based models. Moreover, ~75% of the stations in all four models overestimate the rain rate, which is consistent with the analysis of the aforementioned hourly scale results.

For more sample validation, Figure 11 shows the surface precipitation integration results from our integration models against ground rain-gauge observations at both hourly (top panels) and daily (bottom panels) scales for all three landfalling typhoon events. The surface precipitation integration results and rain-gauge observations during all three typhoon events are compared simultaneously, with the color bars indicating the occurrence frequency on a log scale at intervals of 0.5 mm/h on the hourly scale and 5 mm/day on the daily scale. In general, the surface precipitation integration results from the rain-gauge-based models show better consistency with the rain-gauge observations than those from the IMERG-based models, not only at the hourly scale but also at the daily scale.

To explore the temporal characteristics of the four surface precipitation integration models, Figure 12 illustrates time series plots of the four evaluation metrics at the hourly scale during the three typhoon events. The solid and dotted lines represent the rain-gauge-based and IMERG-based models, respectively, and the blue and red lines represent the F4-based and H8-based models. Generally, the rain-gauge-based models (F4-RG and H8-RG) perform significantly better than the IMERG-based models (F4-IM and H8-IM) with relatively better classification metrics without apparent fluctuations and with better statistical metrics. Because the rain-gauge-based models use ground rain-gauge observations as the predictor while the IMERG-based models use IMERG estimates based on satellite observations as the predictor, the surface precipitation integration results of

these two kinds of models differ greatly. In addition, because of the similar satellite observation channels adopted for the model development, the F4- based and H8-based models yield very similar surface precipitation integration results.

295     Note that all surface precipitation integration models perform better during Ewiniar and Mangkhut than during Bebinca, with the former two having higher POD and R values and lower FAR values; this was caused by the uneven distribution of surface precipitation. As shown in Table 1, there was either light precipitation or no precipitation over most of the land area during Bebinca, accounting for 78% and 18%, respectively. Therefore, the surface precipitation integration results of all models for the precipitation process during Bebinca are relatively poor.

**5 Summary**

300     This paper proposes an RF-based surface precipitation integration framework for typhoons making landfall with geostationary spectral radiometer observations, atmospheric reanalysis data, high-density rain-gauge observations and IMERG estimates. To develop the model, we consider either F4 or H8 observations as the predictor and either rain-gauge observations or IMERG estimates as the predictand. Thus, the performances of four models, i.e., F4-RG, H8-RG, F4-IM, and H8-IM, are systematically evaluated. All four models are capable of capturing precipitation events caused by typhoons making landfall. Regardless of

305 hourly precipitation or daily precipitation, POD is greater than 0.9, and FAR is approximately 0.5, and the rain-gauge-based models (F4-RG and H8-RG) can estimate surface precipitation well. For hourly precipitation, the R values between F4-RG and H8-RG and the ground rain-gauge observations are greater than 0.7, and the RMSE is approximately 2.5 mm/h. For daily precipitation, the R values between F4-RG and H8-RG and the ground rain-gauge observations are approximately 0.9, and the RMSE is approximately 25 mm/d. In contrast, while the two IMERG-based models achieve good success in model

310 development, when the surface precipitation integration results of the two models are compared with the ground rain-gauge observations, the comparison results are obviously worse than those of the rain-gauge-based models, which indicates that the ground rain-gauge data are more suitable as the model development ground truth for the typhoon surface precipitation integration algorithm. The great performance of our typhoon-only model indicate that the surface precipitation estimation could be further improved by developing and using different models for different precipitation types.

315     Note that the input variables of the surface precipitation integration models include satellite observations, geographic locations and channel combinations. The key point of establishing surface precipitation integration models is how to accurately discern the nonlinear relationship between the model input variables and precipitation. According to the importance of the variables shown by the RF algorithm used in the model development, TCRW, TP and geographic location rank much higher in importance, which is useful for confirming the accuracy of the surface precipitation integration results for typhoons.

320 Furthermore, considering the complexity and variability of typhoon precipitating cloud systems, the vertical factors of clouds derived by passive microwave sensors (e.g., cloud water vapor profiles, cloud thickness) with environmental conditions (e.g., wind shear, relative humidity) from a global forecast system should be introduced into the RF model to improve the QPE accuracy in the future.

**Code availability**

325 The model in this paper is based on the random Forest data package in the Python language, and our implementation and analysis code are available upon request to the corresponding author (chao_liu@nuist.edu.cn).

**Data Availability**

Fengyun-4A/AGRI data (http://satellite.nsmc.org.cn/PortalSite/Data/Satellite.aspx) and Himawari-8/AHI data (http://www.eorc.jaxa.jp/ptree/) was used to establish surface precipitation integration models. IMERG data

330 (https://gpm.nasa.gov/data/directory) and surface precipitation observations collected from the China Meteorological Data Service Center (http://data.cma.cn/en) were used as predictands of surface precipitation integration algorithm.

**Supplement**

The supplement related to this article is available online at: https://amt.copernicus.org/

**Author contributions**

335 Conceptualization, C.L.; designed the study layout; methodology, Y.Z. and C.L. ; Data Curation: D.D. and Y.Z.; software, Y.Z.; validation, Y.Z., C.L. and Z.M.; formal analysis, Y.Z., D.D., Z.M., and S.T.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z, C.L., D.D., Z.M. and S.T.; visualization, D.D. and Z.M.; supervision, C.L.; funding acquisition, C.L.. All authors have read and agreed to the published version of the manuscript.

340 **Competing interests**

The authors declare that they have no conflict of interest.

11

**Acknowledgements**

350

## References

Ahmed, K., Sachindra, D. A., Shahid, S., Iqbal, Z., Nawaz, N., and Khan, N.: Multi-model ensemble predictions of precipitation and temperature using machine learning algorithms, Atmospheric Research, 236, 104806, https://doi.org/10.1016/j.atmosres.2019.104806, 2020.

Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi.: Understanding of a convolutional neural network, 2017 International Conference on Engineering and Technology (ICET). Ieee, 2017.

Aonashi, K., Awaka, J., Hirose, M., Kozu, T., Kubota, T., Liu, G., and Takayabu, Y. N.: GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation, Journal of the Meteorological Society of Japan. Ser. II, 87, 119-136, https://doi.org/10.2151/jmsj.87A.119, 2009.

Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., McNamara, I., Ribbe, L., Nauditt, A., and Thinh, N. X.: RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements, Remote Sensing of Environment, 239, 111606, https://doi.org/10.1016/j.rse.2019.111606, 2020.

Bárdossy, A., and Pegram, G.: Combination of radar and daily precipitation data to estimate meaningful sub-daily point precipitation extremes, Journal of Hydrology, 544, 397-406, https://doi.org/10.1016/j.jhydrol.2016.11.039, 2017.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., and Yoshida, R.: An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites, Journal of the Meteorological Society of Japan. Ser. II, 94(2), 151-183, https://doi.org/10.2151/jmsj.2016-009, 2016.

Boushaki, F. I., Hsu, K. L., Sorooshian, S., Park, G. H., Mahani, S., and Shi, W.: Bias adjustment of satellite precipitation estimation using ground-based measurement: A case study evaluation over the southwestern United States. Journal of Hydrometeorology, 10(5), 1231-1242, https://doi.org/10.1175/2009JHM1099.1, 2009.

Chan, J. C.: Interannual and interdecadal variations of tropical cyclone activity over the western North Pacific, Meteorology and Atmospheric Physics, 89(1), 143-152, https://doi.org/10.1007/s00703-005-0126-y, 2005.

Emanuel, K. A.: Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century, Proceedings of the National Academy of Sciences, 110(30), 12219-12224, https://doi.org/10.1073/pnas.1301293110, 2013.

Ebert, E. E., Janowiak, J. E., and Kidd, C.: Comparison of near-real-time precipitation estimates from satellite observations and numerical models, Bulletin of the American Meteorological Society, 88(1), 47-64, https://doi.org/10.1175/BAMS-88-1-47, 2007.

Ehsan B., Md A., Efthymios I. N., and Emmanouil N. A.: Machine learning–based blending of satellite and reanalysis precipitation datasets: A multiregional tropical complex terrain evaluation, Journal of Hydrometeorology, 20(11), 2147-2161, 2019

Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., and Simoes, N.: Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. Journal of Hydrology, 442,

385  117-128, https://doi.org/10.1016/j.jhydrol.2012.04.005, 2012.

Guo, B., Zhang, D., Pei, L., Su, Y., Wang, X., Bian, Y., and Guo, L.: Estimating PM2. 5 concentrations via random forest method using satellite, auxiliary, and ground-level station dataset at multiple temporal scales across China in 2017, Science of The Total Environment, 778, 146288, https://doi.org/10.1016/j.scitotenv.2021.146288, 2021.

390  Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146(730), 1999-2049, https://doi.org/10.1002/qj.3803, 2020.

Ho, C. H., Baik, J. J., Kim, J. H., Gong, D. Y., and Sui, C. H.: Interdecadal changes in summertime typhoon tracks, Journal of Climate, 17(9), 1767-1776, https://doi.org/10.1175/1520-0442(2004)017<1767:ICISTT>2.0.CO;2, 2004.

Honda, T., Miyoshi, T., Lien, G. Y., Nishizawa, S., Yoshida, R., Adachi, S. A., and Bessho, K.: Assimilating all-sky Himawari-

395  8 satellite infrared radiances: A case of Typhoon Soudelor (2015), Monthly Weather Review, 146(1), 213-229, https://doi.org/10.1175/MWR-D-16-0357.1, 2018.

Hong, Y., Hsu, K. L., Sorooshian, S., and Gao, X.: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system, Journal of Applied Meteorology, 43(12), 1834-1853, https://doi.org/10.1175/JAM2173.1, 2004.

400  Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of hydrometeorology, 8(1), 38-55, https://doi.org/10.1175/JHM560.1, 2007.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., and Yoo, S. H.: NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD)

405  Version, 4, 26, 2015.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of hydrometeorology, 5(3), 487-503, https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.

Jozaghi, A., Nabatian, M., Noh, S., Seo, D. J., Tang, L., and Zhang, J.: Improving multisensor precipitation estimation via

410  adaptive conditional bias–penalized merging of rain gauge data and remotely sensed quantitative precipitation estimates. Journal of Hydrometeorology, 20(12), 2347-2365, https://doi.org/10.1175/JHM-D-19-0129.1, 2019.

Kühnlein, M., Appelhans, T., Thies, B., and Nauss, T.: Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—A random forests-based approach applied to MSG SEVIRI, Remote Sensing of Environment, 141, 129-143, https://doi.org/10.1016/j.rse.2013.10.026, 2014.

14

415     Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T.: Precipitation estimates from MSG SEVIRI daytime, nighttime, and twilight data with random forests, Journal of Applied Meteorology and Climatology, 53(11), 2457-2480, https://doi.org/10.1175/JAMC-D-14-0082.1, 2014.

Liu, C., Yang, S., Di, D., Yang, Y., Zhou, C., Hu, X., and Sohn, B. J.: A Machine Learning-based Cloud Detection Algorithm for the Himawari-8 Spectral Image, Advances in Atmospheric Sciences, 1-14, https://doi.org/10.1007/s00376-021-0366-x,

420     2021.

Liu, Z.: Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: initial results. Journal of Hydrometeorology, 17(3), 777-790, https://doi.org/10.1175/JHM-D-15-0068.1, 2016.

Looper, J. P., and Vieux, B. E.: An assessment of distributed flash flood forecasting accuracy using radar and rain gauge input

425     for a physics-based distributed hydrologic model, Journal of Hydrology, 412, 114-132, https://doi.org/10.1016/j.jhydrol.2011.05.046, 2012.

Ma, Z., Li, J., Han, W., Li, Z., Zeng, Q., Menzel, W. P., and Liu, C. Y.: Four-Dimensional Wind Fields From Geostationary Hyperspectral Infrared Sounder Radiance Measurements With High Temporal Resolution, Geophysical Research Letters, 48(14), e2021GL093794, https://doi.org/10.1029/2021GL093794, 2021.

430     Marra, F., and Morin, E.: Use of radar QPE for the derivation of Intensity–Duration–Frequency curves in a range of climatic regimes, Journal of hydrology, 531, 427-440, https://doi.org/10.1016/j.jhydrol.2015.08.064, 2015.

Mecikalski, J. R., Bedka, K. M., Paech, S. J., & Litten, L. A.: A statistical evaluation of GOES cloud-top properties for nowcasting convective initiation, Monthly Weather Review, 136(12), 4899-4914, https://doi.org/10.1175/2008MWR2352.1, 2008.

435     Min, M., Wu, C., Li, C., Liu, H., Xu, N., Wu, X., and Dong, L.: Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: Fengyun-4 series, Journal of Meteorological Research, 31(4), 708-719, https://doi.org/10.1007/s13351-017-6161-z, 2017.

Min, M., Bai, C., Guo, J., Sun, F., Liu, C., Wang, F., and Li, J.: Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning, IEEE transactions on geoscience and remote sensing, 57(5), 2557-2570,

440     2018.

Mountrakis, G., Im, J., and Ogole, C.: Support vector machines in remote sensing: A review, ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247-259, https://doi.org/10.1016/j.isprsjprs.2010.11.001, 2011.

Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., and Thorstensen, A. R.: The PERSIANN family of global satellite precipitation data: a review and evaluation of products, Hydrol. Earth Syst. Sci., 22,

445     5801–5816, https://doi.org/10.5194/hess-22-5801-2018, 2018.

Ochoa-Rodriguez, S., Wang, L. P., Willems, P., and Onof, C.: A review of radar-rain gauge data merging methods and their potential for urban hydrological applications, Water Resources Research, 55(8), 6356-6391, https://doi.org/10.1029/2018WR023332, 2019.

15

Rappaport, E. N.: Fatalities in the United States from Atlantic tropical cyclones: New data and interpretation, Bulletin of the
450    American Meteorological Society, 95(3), 341-346, https://doi.org/10.1175/BAMS-D-12-00074.1,  2014.

Sehad, M., Lazri, M., and Ameur, S.: Novel SVM-based technique to improve rainfall estimation over the Mediterranean
region (north of Algeria) using the multispectral MSG SEVIRI imagery, Advances in Space Research, 59(5), 1381-1394,
https://doi.org/10.1016/j.asr.2016.11.042, 2017.

Tan, Z., Liu, C., Ma, S., Wang, X., Shang, J., Wang, J., and Yan, W.: Detecting Multilayer Clouds From the Geostationary
455    Advanced Himawari Imager Using Machine Learning Techniques. IEEE Transactions on Geoscience and Remote Sensing,
doi: 10.1109/TGRS.2021.3087714, 2021.

Tang, G., Ma, Y., Long, D., Zhong, L., and Hong, Y.: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products
over    Mainland    China    at    multiple    spatiotemporal    scales,    Journal    of    hydrology,    533,    152-167,
https://doi.org/10.1016/j.jhydrol.2015.12.008, 2016.

460    Tang, G., Ma, Y., Long, D., Zhong, L., and Hong, Y.: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products
over    Mainland    China    at    multiple    spatiotemporal    scales,    Journal    of    hydrology,    533,    152-167,
https://doi.org/10.1016/j.jhydrol.2015.12.008, 2016.

Wang, H., Xu, M., Onyejuruwa, A., Wang, Y., Wen, S., Gao, A. E., and Li, Y.: Tropical cyclone damages in Mainland China
over 2005–2016: Losses analysis and implications, Environment, Development and Sustainability, 21(6), 3077-3092,
465    https://doi.org/10.1007/s10668-019-00481-7, 2019.

Negri, A. J., Burkardt, N., Golden, J. H., Halverson, J. B., Huffman, G. J., Larsen, M. C., and Wieczorek, G. F.: The hurricane–
flood–landslide    continuum,    Bulletin    of    the    American    Meteorological    Society,    86(9),    1241-1247,
https://www.jstor.org/stable/26221353, 2005.

Su, S. H., Kuo, H. C., Hsu, L. H., & Yang, Y. T.: Temporal and spatial characteristics of typhoon extreme rainfall in Taiwan,
470    Journal of the Meteorological Society of Japan. Ser. II, 90(5), 721-736, https://doi.org/10.2151/jmsj.2012-510, 2012.

Tu, J. Y., and Chou, C.: Changes in precipitation frequency and intensity in the vicinity of Taiwan: typhoon versus non-
typhoon events, Environmental Research Letters, 8(1), 014023, 2013.

Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J.: Rainfall and sampling uncertainties: A rain gauge
perspective, Journal of Geophysical Research: Atmospheres, 113(D11), 2008.

475    Ochoa-Rodriguez, S., Wang, L. P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., and ten Veldhuis, M. C.: Impact
of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment
investigation, Journal of Hydrology, 531, 389-407, https://doi.org/10.1016/j.jhydrol.2015.05.035, 2015.

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K. I., Aonashi, K., and Kawasaki, Z. I.: A Kalman filter approach to
the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data,
480    Journal of the Meteorological Society of Japan. Ser. II, 87, 137-151, https://doi.org/10.2151/jmsj.87A.137, 2009.

Wang, C. Tang, G. Han, Z Guo, X., and Hong, Y.: Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements, Journal of hydrology, 564, 342-356, https://doi.org/10.1016/j.jhydrol.2018.06.064, 2018.

Wang, C., Tang, G., Xiong, W., Ma, Z., and Zhu, S.: Infrared Precipitation Estimation using Convolutional neural network for

485   FengYun satellites, Journal of Hydrology, 127113, https://doi.org/10.1016/j.jhydrol.2021.127113, 2021.

Wang, R., Chen, J., and Wang, X.: Comparison of IMERG Level-3 and TMPA 3B42V7 in estimating typhoon-related heavy rain, Water, 9(4), 276, https://doi.org/10.3390/w9040276, 2017.

Yao, C., Xiao, Z., Yang, S., and Luo, X.: Increased severe landfall typhoons in China since 2004, International Journal of Climatology, 41, E1018-E1027, https://doi.org/10.1002/joc.6746, 2021.

490   Zhang, L., et al.: Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach, Journal of Hydrology, 594, 125969, 2021.

**Table 1: Channel information for F4/AGRI and H8/AHI considered in our integration algorithms.**

| Band name | Central wavelength (μm) | |
|-----------|:-----------------------:|:-------:|
|           | F4/AGRI                 | H8/AHI  |
| WV1       | 6.2                     | 6.2     |
| WV2       | 7.1                     | 7.0     |
| LW1       | 8.5                     | 8.6     |
| LW2       | 10.7                    | 11.2    |
| LW3       | 12.0                    | 12.3    |
| LW4       | 13.5                    | 13.3    |

**Table 2: Information on the three examples of typhoon events that made landfall in South China in 2018.**

| Typhoon name | Period | No rain | Rain<=5 mm/h | Rain>5 mm/h |
|--------------|--------|---------|--------------|-------------|
| Ewiniar  | 6-7 Jun. 2018 (48 hrs)   | 55.7% | 35.9% | 8.4%  |
| Bebinca  | 14-15 Aug. 2018 (48 hrs) | 78.9% | 18.3% | 2.8%  |
| Mangkhut | 16 Sep. 2018 (24 hrs)    | 50.9% | 36.0% | 13.1% |

515

**Table 3: Differences among the four precipitation integration models used in this paper.**

| Satellite data / Precipitation data | FengYun-4A/AGRI (F4-based models) | Himawari-8/AHI (H8-based models) |
|---|---|---|
| Rain-gauge observations (RG-based models) | F4-RG | H8-RG |
| IMERG estimates (IMERG-based models) | F4-IM | H8-IM |

520

**Table 4: Data and variables considered in this study for precipitation integration.**

| | Parameters | Type | Resolution |
|---|---|---|---|
| | Longitude, Latitude | Geographic Location | - |
| Predictor | WV1, WV2, LW1, LW2, LW3, LW4, WV1-LW2, LW1-LW2, WV2-LW3, LW2-LW3 | AGRI observation | 4 km/1 h |
| | | AHI observation | 5 km/1 h |
| | CAPE, K-Index, TCRW, TP, TCLW, R850, R950, T850, T950 | Atmospheric Reanalysis | 0.25°/1 h |
| Predictand | Rain-gauge data | Rain-gauge-based | - |
| | IMERG data | IMERG-based | 0.1°/1 h |
| Validation | Rain-gauge data | Ground-based | - |

525 **Table 5: Surface precipitation overestimation percentages for the four models during the three typhoons.**

| Typhoon | Rain Rate | F4-RG | H8-RG | F4-IM | H8-IM |
|---------|-----------|-------|-------|-------|-------|
| Ewiniar | < 5 mm/h | 87% | 87% | 86% | 86% |
| | > 5 mm/h | 22% | 23% | 19% | 19% |
| Bebinca | < 5 mm/h | 88% | 89% | 85% | 84% |
| | > 5 mm/h | 11% | 9.1% | 7.2% | 6.7% |
| Mangkhut | < 5 mm/h | 89% | 88% | 90% | 90% |
| | > 5 mm/h | 25% | 24% | 16% | 15% |
| Total | < 5 mm/h | 88% | 88% | 87% | 87% |
| | > 5 mm/h | 19% | 19% | 14% | 14% |
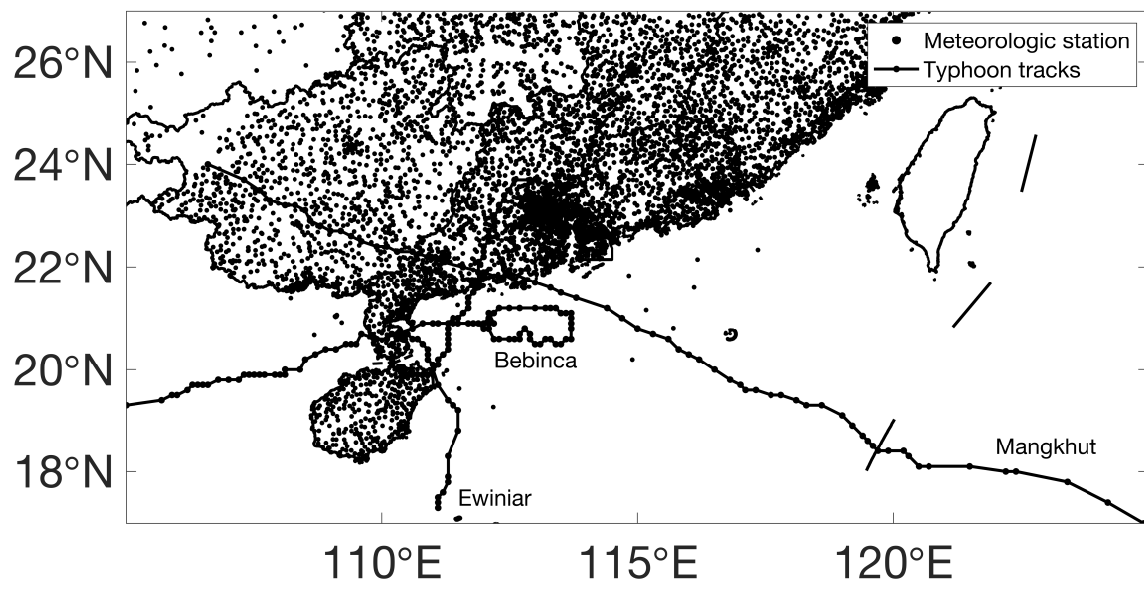
**Figure 1: Distribution of high-density automatic stations over the studied area.**

530

## Model development

| IMERG estimates | | Satellite data (FY-4A/H8) | ERA5 data |

↓ Matching     ↓ Matching ↓

High-density rain-guage station data     High-density rain-guage station data

Replace the rain-gauge data with IMERG estimates ←

IMERG-based datasets     Station-based datasets

Random forest algorithm (10-fold cross-validation)     Random forest algorithm (10-fold cross-validation)

F4-IM & H8-IM    SPF models    F4-RG & H8-RG

## Surface precipitation estimation

Satellite data (FY-4A/H8)     ERA5 data

Linear interpolation

Satellite position (FY-4A/H8)

Estimation datasets (FY-4A/H8)

↓ input

Surface precipitation estimation

↓ output

Surface precipitation results (mm/h)

Precipitation classification

Surface precipitation products

Application →



**Figure 2: Flowchart of the surface precipitation integration algorithm.**

535
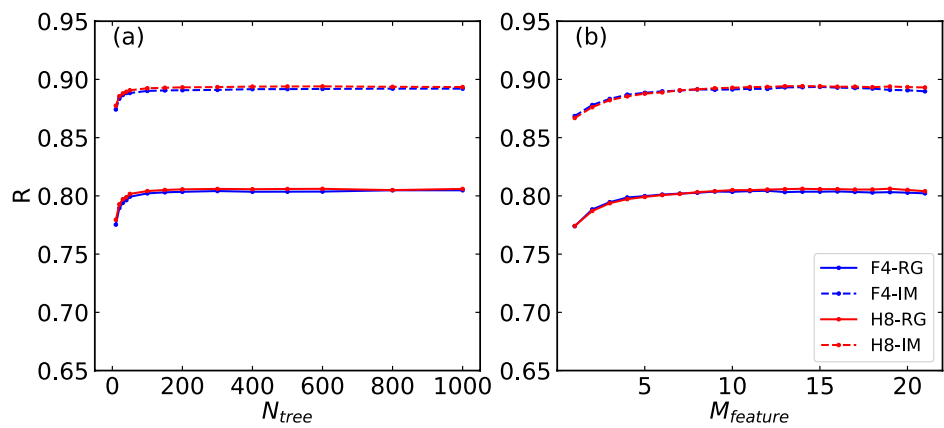
Figure 3: Dependence of the correlation coefficient (R) on the parameters $N_{tree}$ (a) and $M_{feature}$ (b).
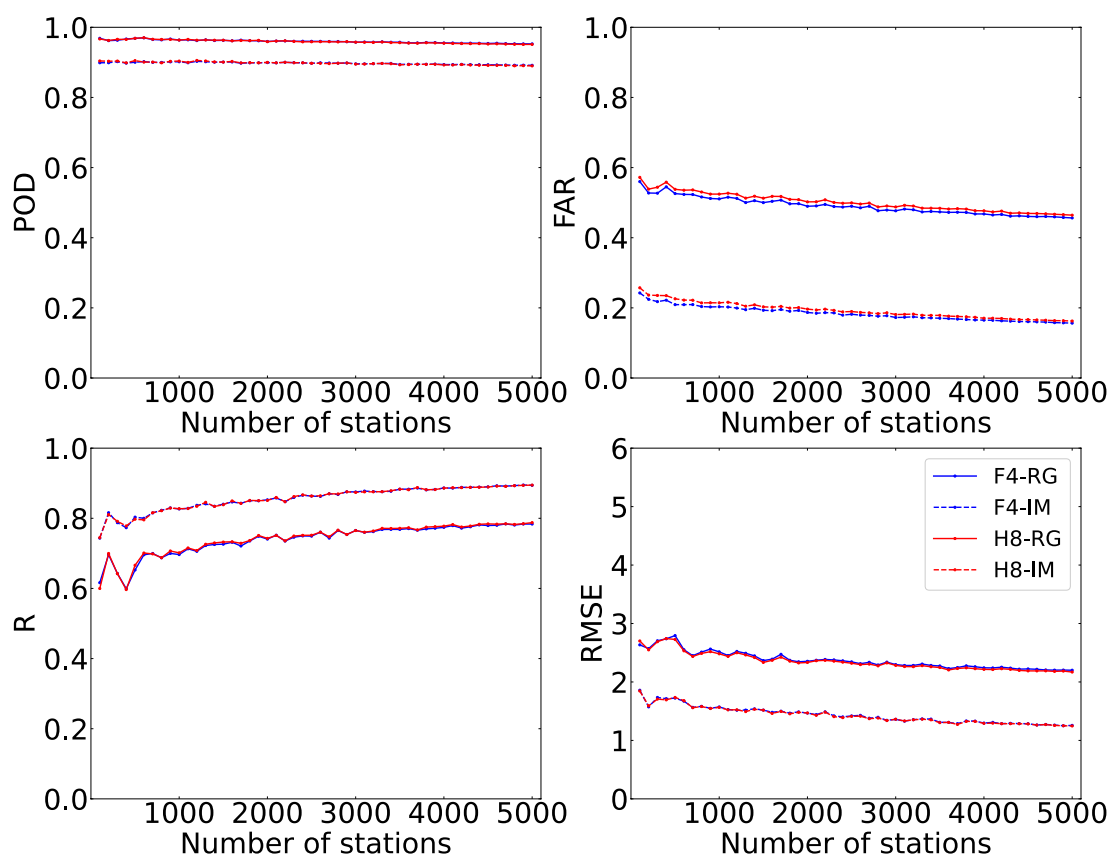
**Figure 4: Influence of the number of ground stations on the model training results.**

540
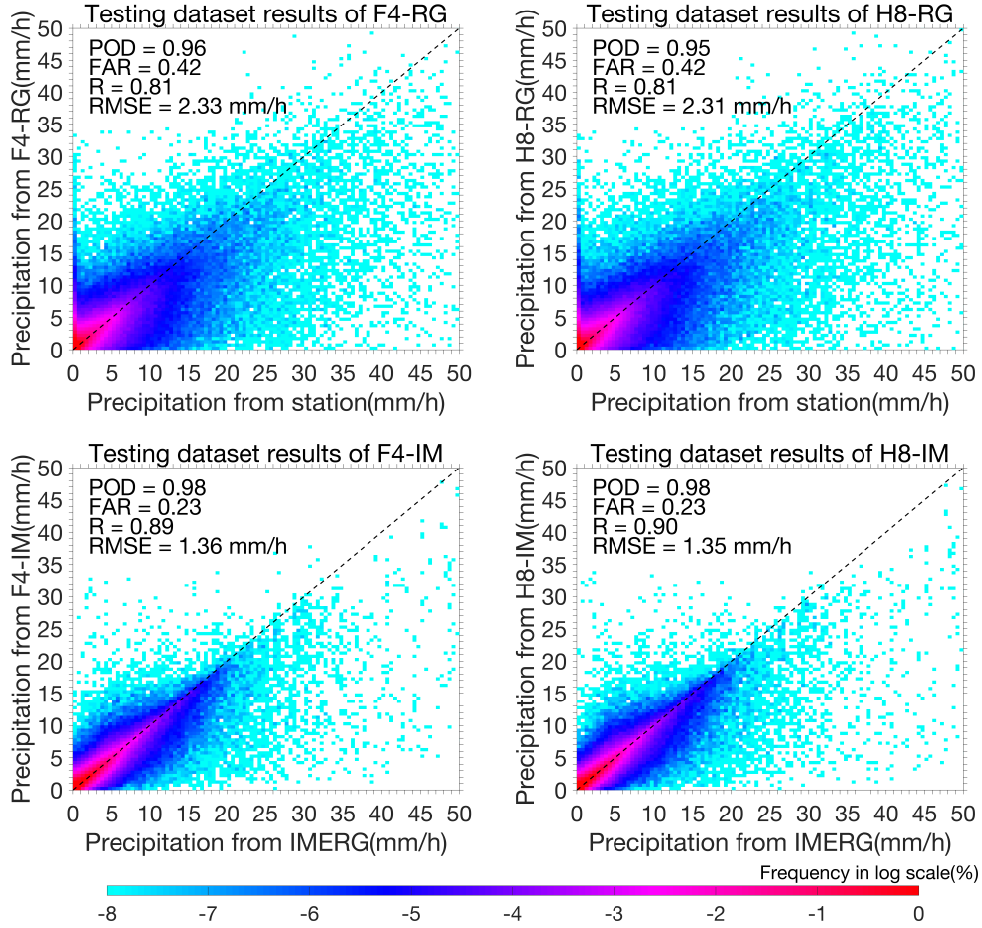
**Figure 5: Scatter plots of the hourly precipitation between different surface precipitation integration models (F4-RG, H8-RG, F4-IM and H8-IM) and rain-gauge observations over the testing datasets created by the 10-cv method during the RF algorithm. The black dotted line in all panels represents the 1:1 line.**

545
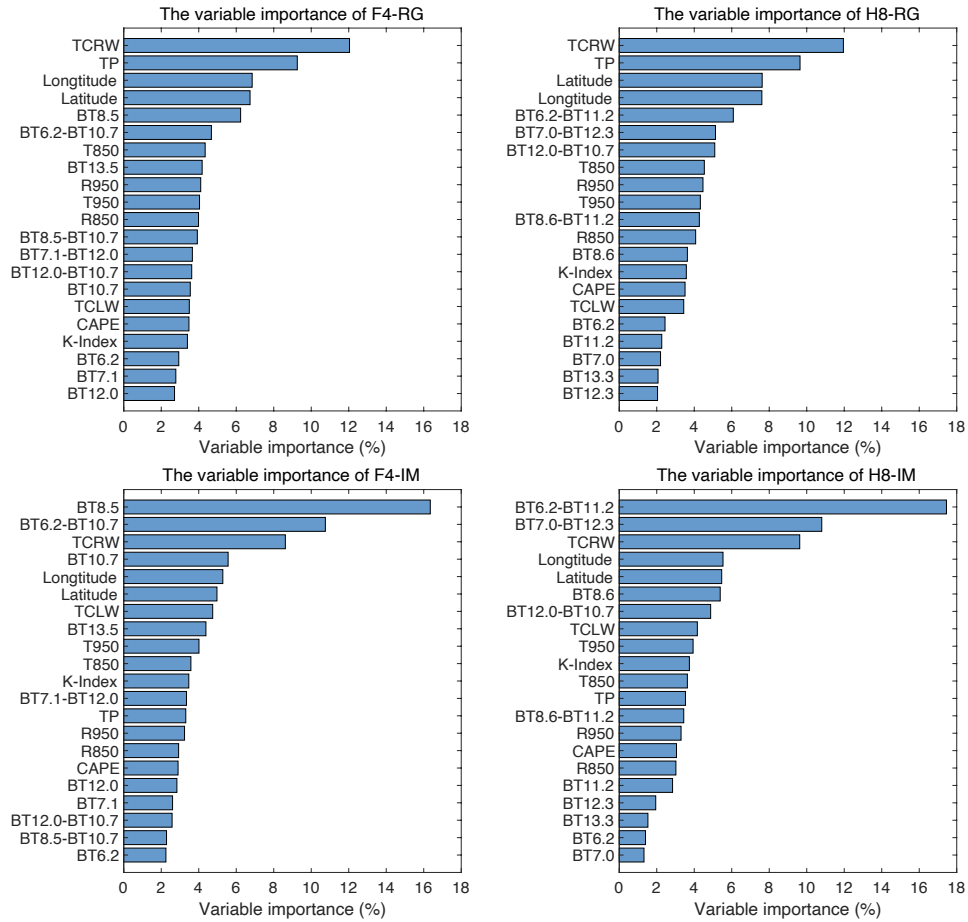
**Figure 6: Importance of each variable in the RF algorithm when estimating the surface precipitation.**
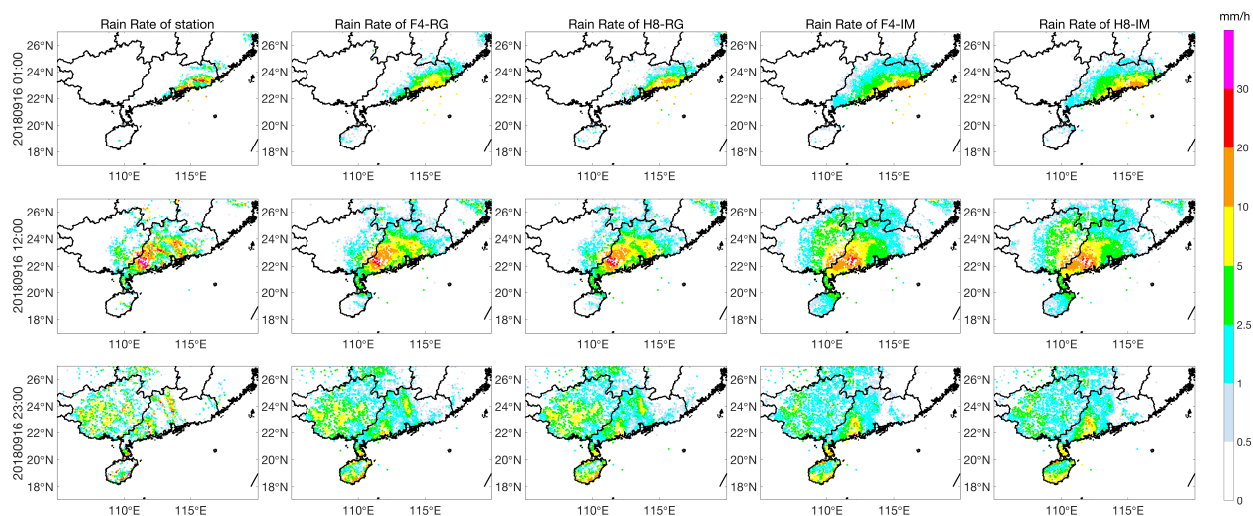
550 **Figure 7: Comparison of the hourly precipitation from the ground rain-gauge observations with the F4-RG, H8-RG, F4-IM and H8-IM estimates at three time steps (0100 UTC on 16 Sep 2018, 1200 UTC on 16 Sep 2018, and 2300 UTC on 17 Sep 2018).**
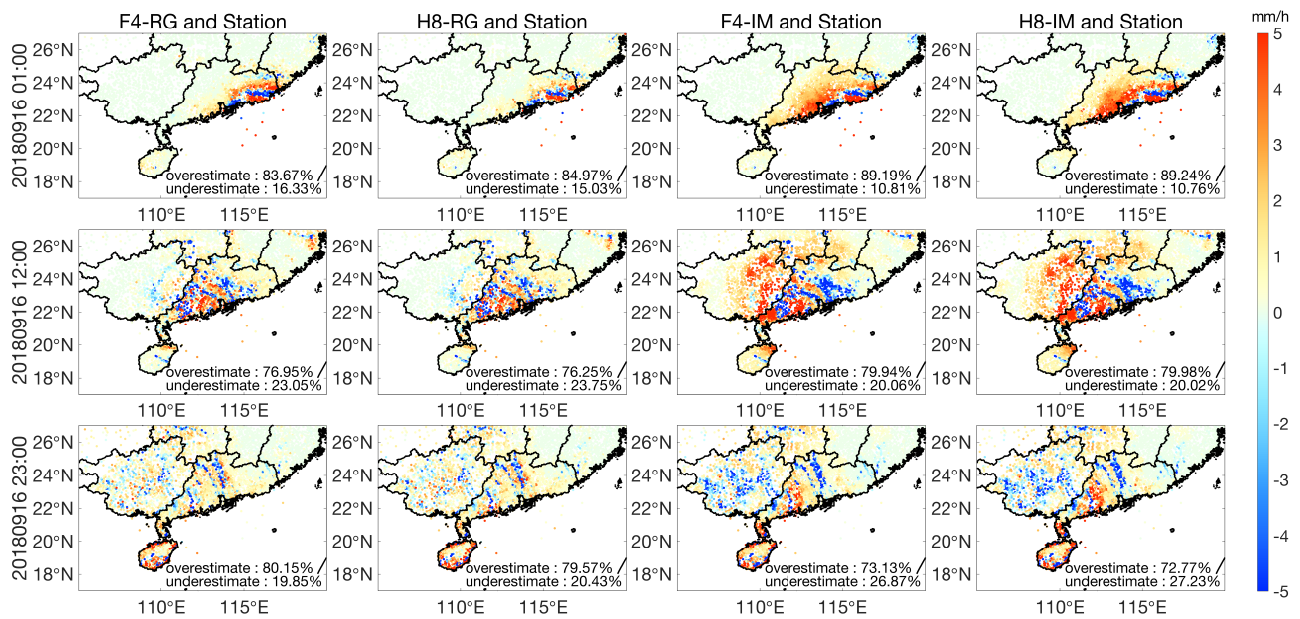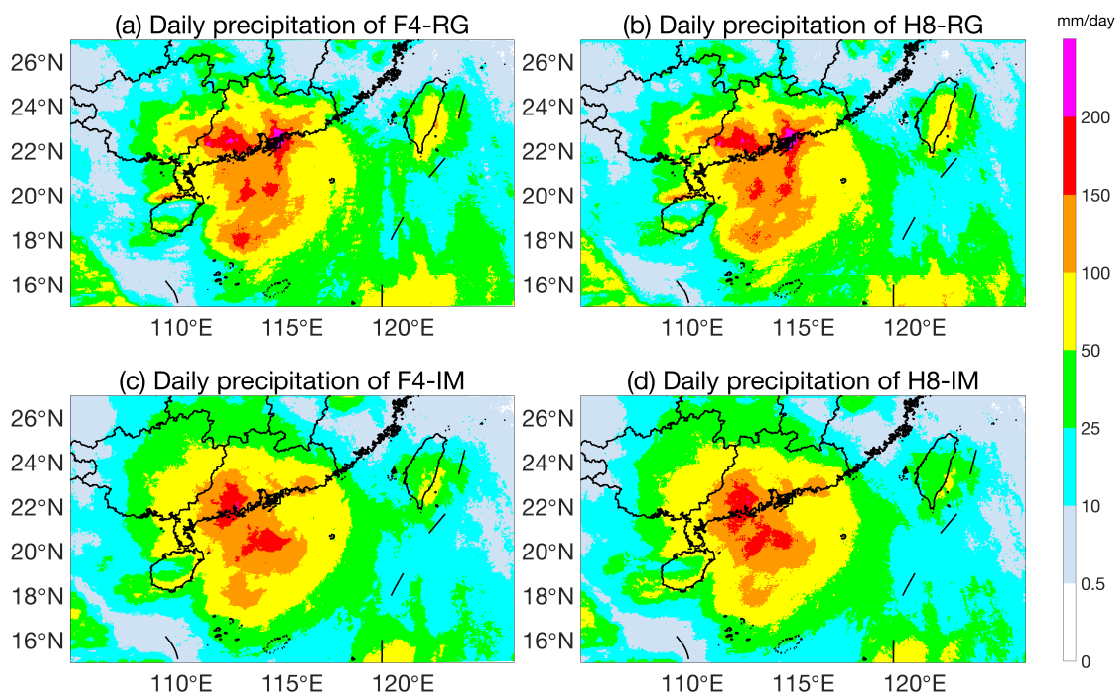
**Figure 8: Differences in hourly precipitation between F4-RG, H8-RG, F4-IM and H8-IM and the ground rain-gauge observations at three time steps (0100 UTC on 16 Sep 2018, 1200 UTC on 16 Sep 2018, and 2300 UTC on 17 Sep 2018).**

**Figure 9: Comparison of daily precipitation among F4-RG, H8-RG, F4-IM and H8-IM on 16 Sep 2018.**
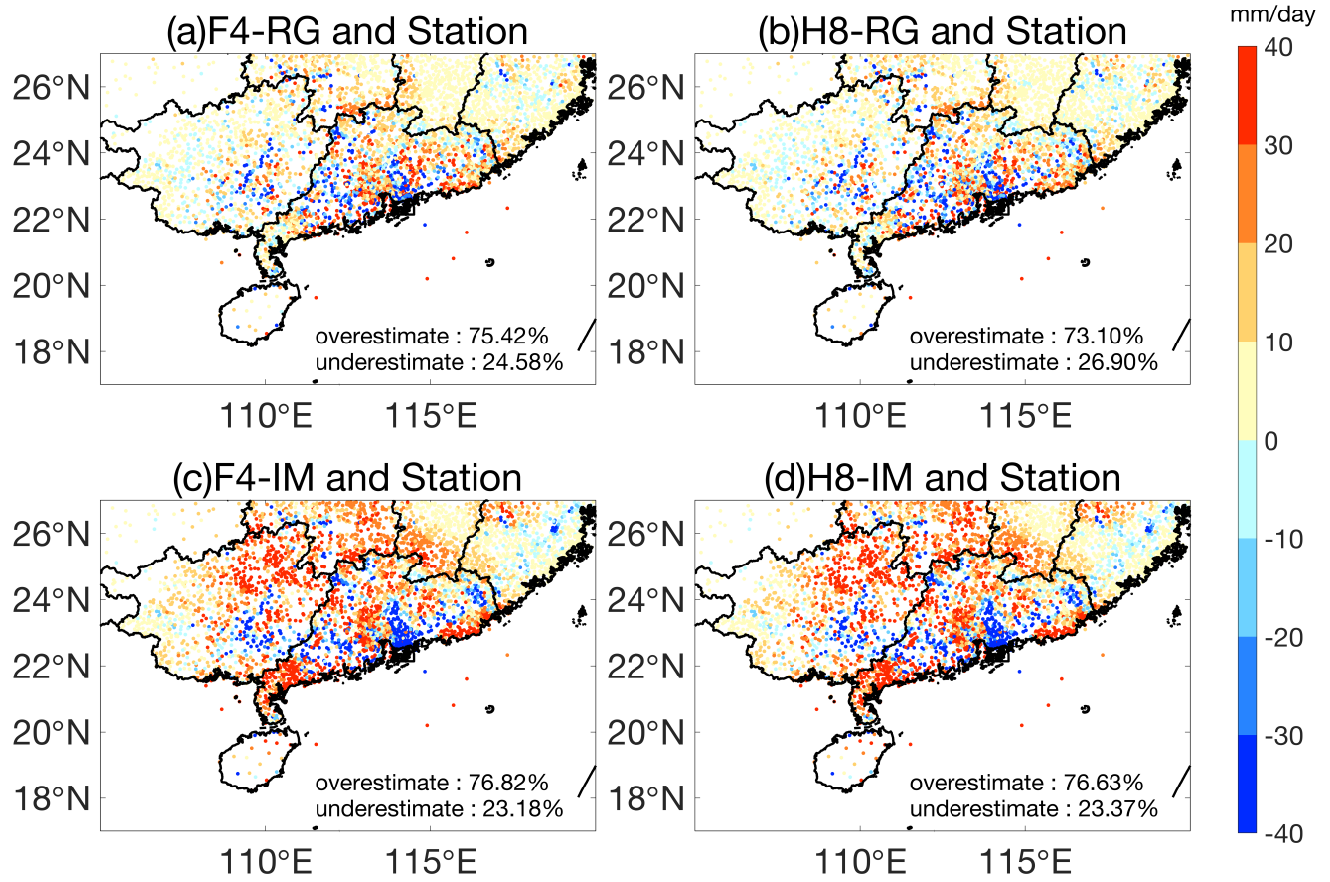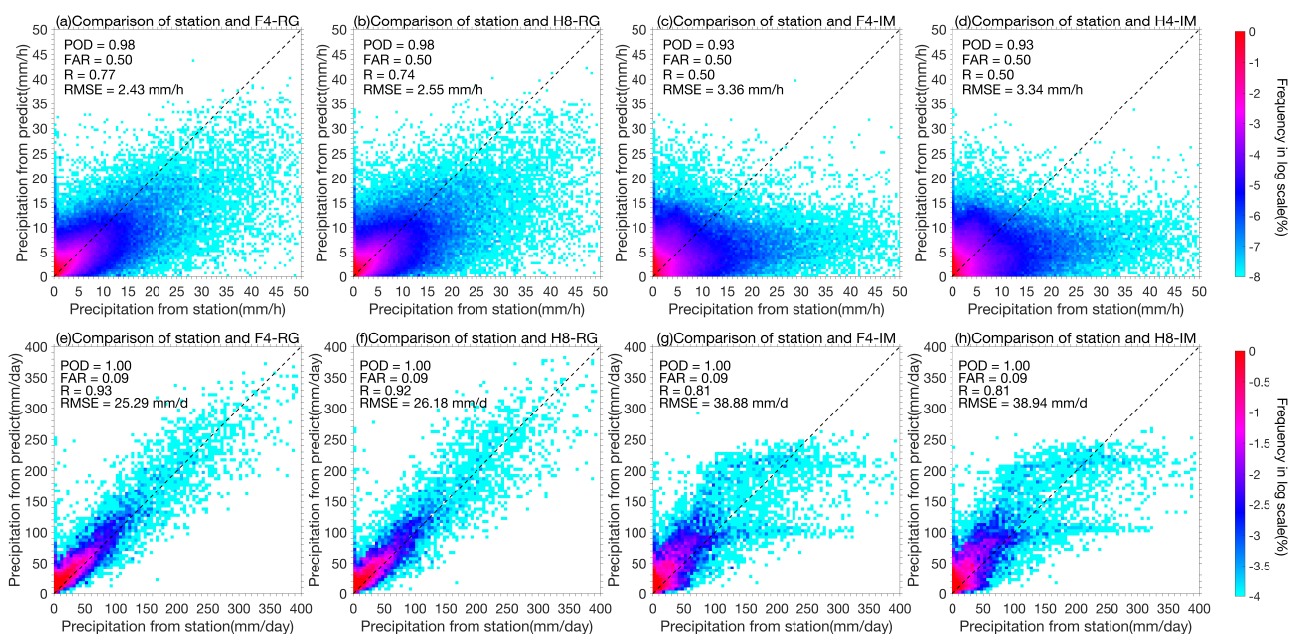
560

**Figure 10: Differences in daily precipitation between F4-RG, H8-RG, F4-IM and H8-IM and the ground rain-gauge observations on 16 Sep 2018.**

565 **Figure 11: Probability density distribution of the hourly ((a)~(d)) and daily ((e)~(h)) precipitation of F4-RG, H8-RG, F4-IM and H8-IM during three typhoon events. The black dotted line in all panels represents the 1:1 line.**
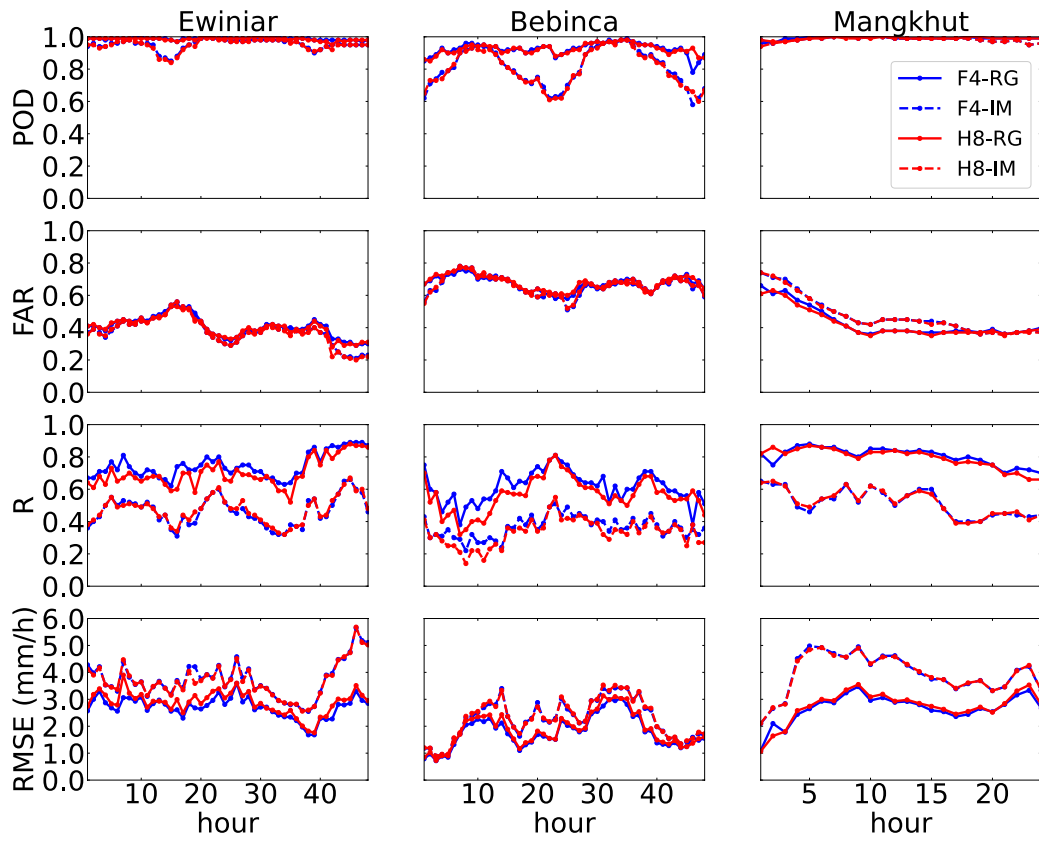
**Figure 12: Time series plots of the POD, FAR, R, and RMSE (mm/h) from F4-RG, H8-RG, F4-IM and H8-IM throughout the evolution of the three typhoon events.**