JOURNAL OF
HYDRO-METEOROLOGY

Check for updates

Research Article

Development of precipitation guidance for 36 regions in Vietnam up to 5 days ahead

Kiichi Sasaki[1]*, Vu Tuan Anh[2]

[1] Japan Meteorological Business Support Center, Tokyo101-0054 Japan; k-sasaki@jmbsc.or.jp
[2] National Center for Hydro-Meteorological Forecasting, Hanoi 10000, Vietnam; lamhoaanh@gmail.com

*Corresponding author: k–sasaki@jmbsc.or.jp; Tel.: +81–352810440

Received: 10 February 2023; Accepted: 21 March 2023; Published: 25 June 2023

Abstract: Development of forecast guidance is one of main activities of Output 3 of the JICA Project to improve forecasting services of VNMHA. Maximum and minimum temperature guidance was developed for 63 cities up to 10 days ahead in the first phase of the Project. Development of precipitation guidance was the primary activity of Output 3 in the second phase of the Project. Statistical analysis on 24-hour rainfall observations and predictions by JMA GSM and ECWMF IFS showed that the correlation between them was low for each station but relatively high for each region. Preliminary investigation of POP trial guidance calculated with logistic regression and multiple linear regression for 4 stations and 4 regions showed the verifications scores (BSS: Brier Skill Score) for regional POP were considerably higher than those for station POP. IFS-GSM-integrated precipitation guidance was developed to improve regional mean/max rainfall guidance and had slightly better verification results than IFS and GSM precipitation guidance. Based on these verification results, precipitation guidance on regional POP and regional mean and max for 12-hour rainfall (tonight and tomorrow daytime) and for 24-hour rainfall (after that) was developed for 36 regions in Vietnam up to 5-days ahead. Precipitation guidance was generally able to predict a reasonable amount of rainfall for heavy rain events caused by tropical cyclones in 2022, however there were several heavy rain events where IFS and GSM rainfall predictions were quite low, and rainfall guidance was also quite low.

Keywords: Precipitation guidance; Station POP; Regional POP; Regional mean and max rainfall.

1. Introduction

The JICA Project for Strengthening Capacity in Weather Forecasting and Flood Early Warning System started in April 2018 [1]. In the first phase of the project, maximum and minimum temperature guidance for 63 cities up to 10 days ahead was developed by the working group 3 (WG3) of the Project [2]. WG3 worked on precipitation guidance for POP and daily rainfall etc. in the second phase of the project. The National Center for Hydro-Meteorological Forecasting (NCHMF) issues city forecasts such as maximum and minimum temperatures for 63 cities in Vietnam but does not issue quantitative forecasts on precipitation such as probability of precipitation (POP). The main objective of WG3 activities in the second phase is to provide guidance materials to forecasters to assist them in issuing quantitative forecast information on precipitation.

MOS guidance is widely used as an objective forecast material to support the issuance of quantitative forecasts. MOS guidance has been used in the United States since around 1970 as an objective forecasting material using numerical predictions [3–4]. In Japan, the guidance operation started in the late 1970s. At JMA, all guidance was prepared using multiple linear regression until around 1996. Multiple linear regression is still widely used in the United States, Canada, and many other countries because it is easy to grasp the characteristics of the guidance and can be used effectively [5–9]. In this work, we used multiple linear regression and logistic regression based on training materials on guidance in the JICA group training course in meteorology implemented by JMA from the perspective of creating guidance that is easy for forecasters to understand.

A statistical analysis on 24-hour rainfall observations and Numerical Weather Prediction (NWP) rainfall predictions by JMA GSM Grid Point Values (GPVs) and ECMWF IFS GPVs was first performed to understand the accuracy of precipitation predictions by NWP models in Vietnam, and then preliminary investigation for POP guidance and for regional mean and max 24-hour rainfall was conducted. The statistical analysis and preliminary investigation showed that the correlation between rainfall observations and NWP predictions and verification results of rainfall guidance were not exceptionally good at each station but were relatively good in each reason. With respect to POP guidance, past research has shown that POP guidance by logistic regression has better verification scores than that by multiple linear regression [10]. In this study, the POP guidance by logistic regression scored better with respect to station POPs, but the regional POP guidance by multiple linear regression scored considerably better than the station POP guidance by logistic regression. Based on the results, experimental precipitation guidance for regional 12-hour/24-hour POP, regional mean and max 12-hour/24-hour rainfall was developed for 63 provinces up to 5 days ahead. This technical note reports on activities of WG3 related to the development of precipitation guidance during the second phase of the Project.

2. Materials and Methods

Guidance is one of the applied products of NWP. Guidance is produced by post-processing NWP output to improve the accuracy of the predictions of temperature and precipitation calculated by the NWP model by correcting them statistically. To create guidance, one must first use NWP data and observation data of the weather element to be forecasted to create a forecast equation using statistical methods. For this purpose, pairs of past NWP data and observation data are prepared. From these data, a forecast equation is created using the observation data of the weather element to be forecasted as the objective variable and the elements of the NWP data that have strong causal relationships with the observation data as explanatory variables. Statistical methods for guidance include linear multiple regression, logistic regression, Kalman filter, and neural network. Kalman filter was used for temperature guidance, while multiple linear regression and logistic regression were used to develop precipitation guidance. Multiple linear regression is a lump-sum statistical method, and the characteristics of the guidance can be easily understood and effectively applied. Logistic regression is one of the statistical methods used for the problem of classifying phenomena into two classes, such as the presence or absence of lightning, for example, and is often used for probability-type guidance.

2.1. NWP GPV and rainfall observation data set

2.1.1. Dataset with surface GPVs for the first statistical analysis

In developing precipitation guidance, it is necessary to know the meteorological characteristics of rainfall in Vietnam and the error characteristics of NWP rainfall predictions. We first created a two-year dataset of 24-hour rainfall observations and GSM surface predictions for 186 stations and conducted statistical analysis with the data set to obtain relations between rainfall observations and NWP predictions. The dataset was created using observation data files (186smMMyYY.xls) prepared by NCHMF, GSM GPVs obtained from JMA High-Resolution GSM Data Service (https://www.wis-jma.go.jp/cms/gsm/download.html), and IFS GPVs prepared by NCHMF's NWP division. In addition, we divided Vietnam into 21 regions based on its administrative division and created a regional dataset for each region. A sample of regional data set for Hanoi region Day1 for Jan 2018 to Dec 2019 is shown in Figure 1. For the IFS, datasets like those for the GSM were created, but the data period was different depending on data availability.

Date	Rain0_1	Robs_mean	Robs_max	Rgpv_mean	Rgpv_max	Hgpv_mean	Cgpv_mean	Ugpv_max	Vgpv_max	Fgpv_max
20180102	0	0.34	2.7	0.95	1.44	88.34	71.17	-4.12	-1.69	4.45
20180103	0	0.63	3.2	1.15	1.88	89.19	79.02	-3.93	-2.75	4.8
20180104	1	1.27	2.8	1.34	2.62	90.99	88.64	-3.34	-3.05	4.47
20180105	0	0.82	3.6	1.05	1.56	91.06	90.82	-3.48	1.67	3.67
20180106	0	0.24	0.9	1.14	1.56	90.72	84.32	-3.56	1.78	3.97
20180107	0	0.36	1.8	2.82	6.12	91.19	82.67	-3.31	-2.24	4
20180108	1	1.79	5.9	6.63	12.72	93.26	89.08	-7.77	-5.16	9.33
20180109	1	3.32	8.8	2.25	3.38	68.51	80.89	-2.84	-5.97	6.61
20180110	0	0.01	0.1	0	0.03	34.2	80.16	-0.57	-5.07	5.07

regional mean/max 24h-rain obs

regional mean/max GSM 24h-rain GPV

Figure 1. GSM GPV and regional mean and max observation data set for Hanoi region for the period of Jan 2018 - Dec 2019. Surface GPVs (Rgpv_mean: regional mean 24-hour rainfall, Rgpv_max: regional max 24-hour rainfall, Hgpv_mean: regional mean relative humidity, Cgpv_mean: regional mean total cloud amount, Ugpv_max: regional max wind u-component, Vgpv_max: regional max wind v-component, Fgpv_max: regional max wind speed).

2.1.2. Dataset with surface and upper GPVs for preliminary investigation and experimental precipitation guidance

Datasets for four representative stations and regions from northern to southern Vietnam were created for the period from Jun 2018 to Sep 2021 using GSM surface and upper GPVs to conduct a preliminary investigation of precipitation guidance. As for the upper level GPVs, each wind component and relative humidity at 950 hPa, 850 hPa and 700 hPa were used, and additional predictors such as KI (K index), SSI (Showalter stability index), PW (Possible Water) and PCWV were calculated and added to the dataset. PCWV is one of the predictors in the JMA precipitation guidance and expressed by the following equation 1:

$$\text{PCWV} = -\,\text{PW} \times (\text{wind–speed at 850hPa}) \times (\text{P–velocity at 850hPa}) \qquad (1)$$

If P-velocity at 850hPa > 0 then PCWV = 0.

For trial operation of precipitation guidance, regional datasets were created by subdividing Vietnam into 36 regions to make each region as close as possible to the 63 city forecast districts (5 municipalities and 58 provinces). In addition, 12-hour rainfall observations and GPV datasets corresponding to the 12-hour forecast for tonight (7:00 pm today to 7:00 am tomorrow) and tomorrow daytime (7:00 am to 7:00 pm tomorrow) were created in accordance with NCHMF's city forecasts. 24-hour rainfall observations and GPV datasets were created for Day 2, Day 3, Day 4 and Day 5. These datasets were created using GSM GPVs for GSM precipitation guidance, IFS GPVs for IFS precipitation guidance and both IFS and GSM GPVs for IFS-GSM integrated precipitation guidance.

2.1.3. Selection of predictors for POP guidance

Since there are so many predictors in the datasets, we subjectively selected about 10 potential predictors by making a correlation analysis between 24-hour POP observations and each predictor. Figure 2 shows correlation coefficients between station POP observations and predictors at the four stations and those between regional POP observations and predictors in the four regions. Correlations were not extremely high for either the stations or the regions, but in general correlation coefficients between regional POP observations and regional mean predictors were higher than those for station POP.

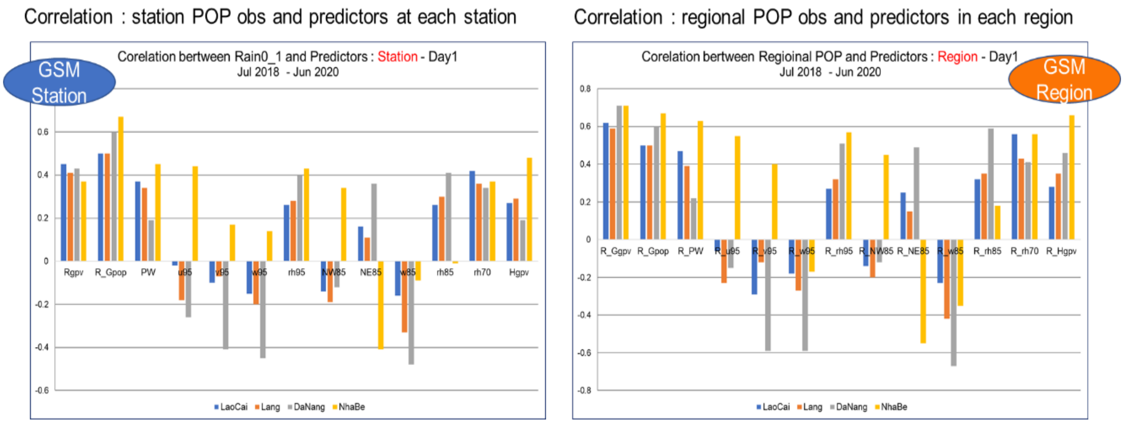

Figure 2. Correlation coefficients between station POP observations and GSM predictors for Day1 (left) and those between regional POP observations and GSM predictors for Day1 (right). Predictors (Rgpv: 24-hour rainfall, R_Gpop: regional mean POP, PW: possible water, u95: wind u-component at 950hPa, v95: wind v-component at 950hPa, w95: p-velocity at 950hPa, rh95: relative humidity at 950hPa, NW85: wind NW-component at 850hPa, NE85: wind NE-component at 850hPa, w85: p-velocity at 850hPa, rh85: relative humidity at 850hPa, rh70: relative humidity at 700hPa, Hgpv: surface relative humidity; prefix R_: regional mean).

For the calculation of station POPs, we subjectively selected potential predictors which had relatively high correlations with POP observations and calculated with logistic regression using all selected predictors at each station, and then objectively selected final predictors for each station using stepwise method. Selection of predictors for regional POPs was done in the same way as for station POPs, using multiple linear regression and stepwise methods. Selected potential predictors for station POP guidance and regional POP guidance are shown in table 1.

Table 1. Selected potential predictors for station POP, regional POP, regional mean rainfall, and regional max rainfall guidance. The prefix R_ stands for regional.

Station POP	Regional POP	regional mean rainfall	regional max rainfall
R_mean 24h rainfall	R_mean 24h rainfall	R_mean 24h rainfall	R_max 24h rainfall
R_mean POP	R_mean POP	R_mean POP	R_mean POP
Possible Water	R_mean Possible Water	R_mean Possible Water	R_mean Possible Water
NE wind at 850hPa	R_mean total cloud amount	R_mean PCWV	R_mean PCWV
NW wind at 850hPa	R_mean surface humidity	R_mean total cloud amount	R_mean total cloud amount
P-velocity at 850hPa	R_mean_u_wind at 850hPa	R_mean surface humidity	R_mean surface humidity
Humidity at 700hPa	R_mean_v_wind at 850hPa	R_mean_NE_wind at 850hPa	R_mean_NE_wind at 850hPa
	R_mean_P-velocity at 850hPa	R_mean_NW_wind at 850hPa	R_mean_NW_wind at 850hPa
	R_mean_Humidity at 850hPa	R_mean_P-velocity at 850hPa	R_mean_P-velocity at 850hPa
		R_mean_Humidity at 850hPa	R_mean_Humidity at 850hPa

2.1.4. Calculation procedure of station POP and regional POP

The predictand of station POP24 is 1 if more than 1mm of rainfall is observed at each station in 24 hours and 0 if less than 1 mm, and the predictand of regional POP24 is the mean of the station POPs in each region. In addition, we calculated regional POHP24 using the predictand, which is 1 if more than 30 mm of rainfall is observed at least 1 station in each region in 24 hours and 0 if less than 30mm. Since the predictand of station POP is 0 or 1 (binominal), station POP was calculated by logistic regression. On the other hand, the predictand of regional POP is 0 to 1, and the regional POP was calculated by multiple linear regression. The predictand of regional POHP is 0 or 1, and regional POHP was calculated by logistic regression. Each regression equation is as follows [11]:

Logistic regression:

forecast equation: $\ln(p/(1–p)) = a_0 + a_1x_1 + a_2x_2 + \ldots,$ (2)

$p = \exp(a_0 + a_1x_1 + a_2x_2 + \ldots) /(1+\exp(a_0 + a_1x_1 + a_2x_2 + \ldots)$ (3)

where p is the probability (supplied with 0 or 1 for response variable); x_i the predictor and a_i the coefficient of the predictor.

Multiple linear regression:

forecast equation: $p = a_0 + a_1x_1 + a_2x_2 + \ldots,$ (4)

where p is the probability (supplied with 0 to 1 for response variable), x_i the predictor and a_i the coefficient of the predictor.

The calculation procedure of station POP and regional POP is as follows: 1) select potential predictors subjectively for station POP and reginal POP based on correlation analysis, 2) calculate logistic regression for station POP and multiple liner regression for regional POP with the selected potential predictors, and then select final predictors for each station and region using stepwise method, 3) apply the obtained forecast equations to verification period datasets for verification. R's glm() and lm() functions were used to calculate logistic regression and multiple liner regression, and stepwise function was used for selecting predictors objectively [11].

2.1.5. Verification of POP and POHP

POP24 is the probability that more than 1 mm of precipitation will fall on any random point of the forecast area during 24-hours from 7:00 pm yesterday to 7:00 pm today. POHP24 is the probability that more than 30mm of precipitation will fall on at least one point of the forecast area during the 24-hours. Brier Score (BS) and Brier Skill Score (BSS) are used for the verification of probability forecasts [12–13]. The BS and BSS formulas are given below. As for BS, the smaller the better and perfect score is 0. As for BSS, the larger the better. If BSS ≤ 0 (BS_{clm} ≤ BS), probability forecasts are no improvement; If BSS > 0 (BS_{clm} > BS), probability forecasts are improvement; If BSS = 1 (BS = 0), probability forecasts are perfect (no error). BSS represents the rate of improvement from climatic probability.

$$BS = \Sigma (Fi – O_i)^2/n \quad (5)$$

where n is the number of data; F_i the probability forecast (0 to 1) and the O_i observation (0 or 1).

$$BSS = (BSclm – BS)/BS_{clm} \quad (6)$$

where BS_{clm} is the BS of climatic probability (use climate value (mean of observed POP) as Forecast).

2.1.6. Calculation procedure of regional mean/max rainfall guidance

Since the accuracy of precipitation predictions at each station was not exactly accurate, we decided to develop 12-hour and 24-hour regional mean and max rainfall guidance as the first experimental precipitation guidance. The rainfall guidance was calculated using multiple linear regression, which is easy to operate and to understand forecast equations and results. The calculation procedures of rainfall guidance are almost the same as the POP guidance: 1) select potential predictors subjectively for regional mean and max rainfall guidance based on correlation analysis; 2) Calculate multiple liner regression for regional mean and max rainfall with the selected potential predictors, and then select final predictors for each region using stepwise method; 3) Apply the obtained forecast equations to verification period data for verification.

2.1.7. Daily calculation procedure of precipitation guidance

The procedure for calculating daily guidance is illustrated in Figure 3: 1) download 6-hourly GSM surface and upper GPVs and collect IFS GPV data from NWP section of NCHMF, and calculate additional predictors; 2) create station-region datasets for 36 regions using the latest GSM and IFS GPVs for IFS, GSM, and IFS-GSM precipitation guidance; 3) apply the forecast equations obtained in each region with development period data to the latest daily dataset.

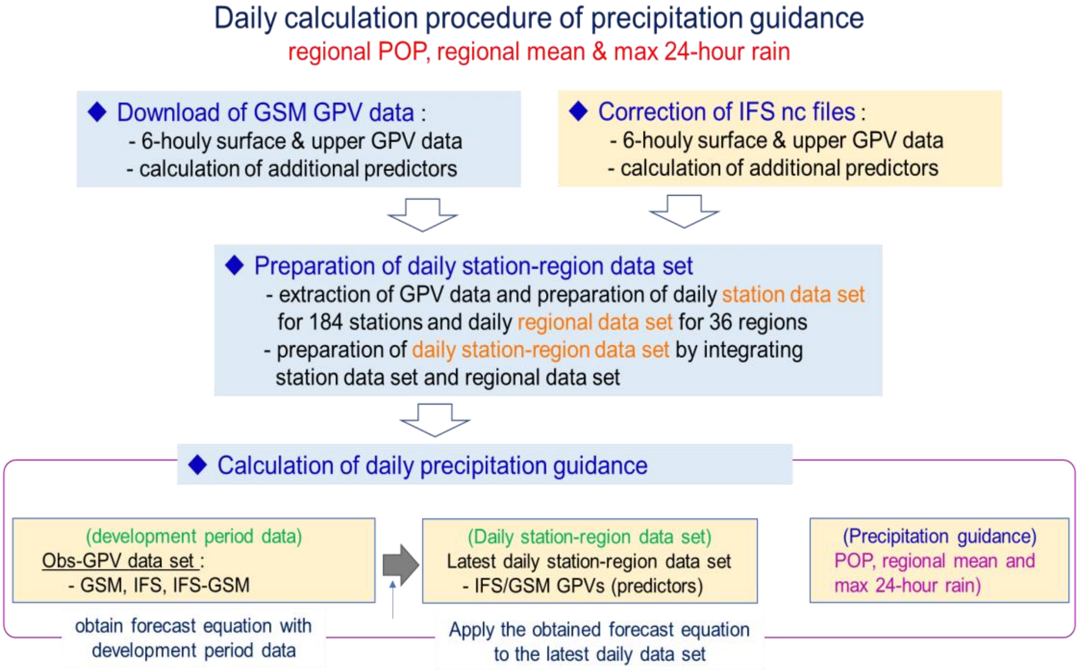

Figure 3. Illustration of daily calculation procedure of precipitation guidance.

3. Result and discussion

3.1. Statistical analysis using rainfall observation-GPV dataset

Rain is very localized and the accuracy of rainfall predictions by NWP model is not yet sufficient for quantitative forecasting. As a first step toward developing precipitation guidance, we conducted various statistical analyses using the rainfall observation-GPV dataset to understand the status of the accuracy of rainfall predictions by NWP models in Vietnam.

3.1.1. Scatter diagram of 24-hour rainfall at stations

First, we created scatter diagrams of observations and GSM predictions for 24-hour rainfall at 186 SYNOP stations using 2-years of rainfall observation-GPV datasets from Jan 2018 to Dec 2019. Figure 4 shows scatter diagrams, regression equations and correlation coefficients for 24-hour rainfall observations and predictions by GSM at Dien Bien, Ha Noi (Lang), Hue and Ho Chi Minh (Nha Be). The scatter diagrams all showed large variations, and heavy rainfalls were often observed even if the predicted rainfalls were low, especially in Hue. The correlation coefficients ranged from 0.45 to 0.62. Looking at the 186 stations, the correlation was generally low in the northern and southern regions, and relatively high in the central region.

3.1.2. Scatter diagram of 24-hour rainfall in regions

Since the correlations were not very high at each station, we next divided Vietnam into 21 regions so that each region would contain more than several SYNOP stations and examined the correlations in each region. The regions were divided based on administrative districts, and if there were only a few SYNOP stations in each province, two or three provinces were merged into one region. Figure 5 shows scatter diagrams for 24-hour regional mean rainfall observations and predictions by GSM in region 1 (Dien Bien), region 8 (Ha Noi), region 12 (Hue) and region 19 (Ho Chi Minh). The variation of scatter diagrams in each region was smaller than that at each station, and the correlation coefficients ranged from 0.65 to 0.74, considerably higher than the correlations at the stations.

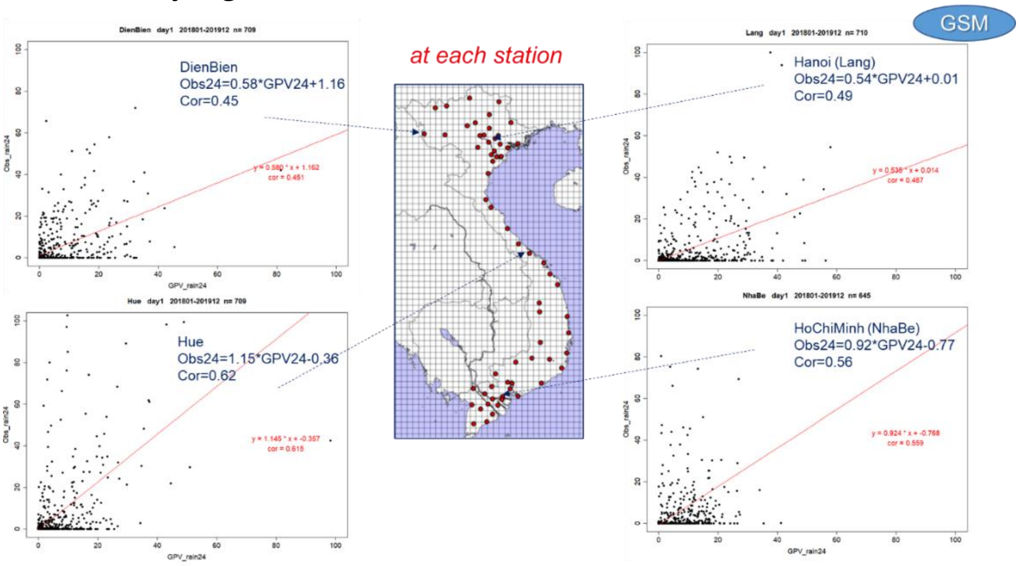

Figure 4. Scatter diagram of 24-hour rainfall observations and predictions for Day1 by GSM at 4 stations.

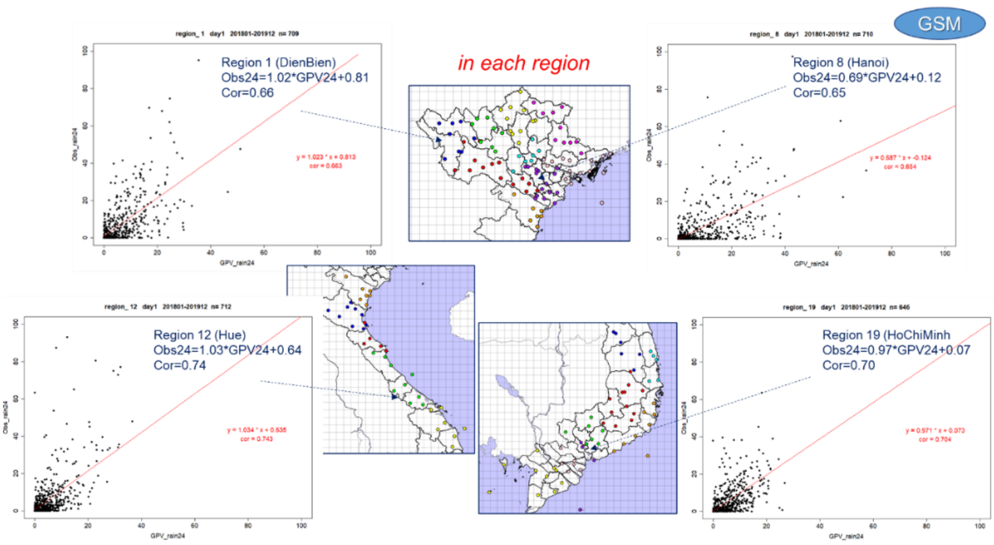

Figure 5. Scatter diagram of 24-hour rainfall observations and predictions for Day1 by GSM in 4 regions. Data period is from Jan 2018 to Dec 2019.

Then, we created similar scatter diagrams with IFS predictions using rainfall observation-GPV datasets from Jun 2019 to Dec 2020. The variation of the IFS regional scatter diagrams was much smaller compared to the GSM station scatter diagrams, and the correlation coefficients were also quite high, ranging from 0.62 to 0.84 (Figure not shown).

3.1.3. Correlation between 24-hour rainfall observations and IFS/GSM predictions

Figure 6 shows the distribution of correlation coefficients between 24-hour rainfall observations and IFS predictions (left side), GSM predictions (right side) for 186 stations and 21 regions. Correlation coefficients for each station were higher in the central region and lower in the northern mountainous region and the south region in both IFS and GSM. In the central region, many stations had correlation coefficients above 0.7, while in the northern mountainous region and the south region, many stations had correlation coefficients below 0.5, and several stations had correlation coefficients below 0.4.

The correlation coefficients of regional mean 24-hour rainfall were considerably higher than those of the stations in each region concerned in both IFS and GSM, especially in the northern mountainous regions and the southern region, where the station correlations were relatively low. IFS regional correlation coefficients were higher than 0.7 in many regions, and they were slightly higher than GSM correlation coefficients for both station and regional correlations in many stations and regions.

The correlation coefficients for regional max 24-hour rainfall were lower than those for regional mean 24-hour rainfall but were greater than 0.6 in many regions. There were a few regions with a correlation coefficient of less than 0.5 in the south.

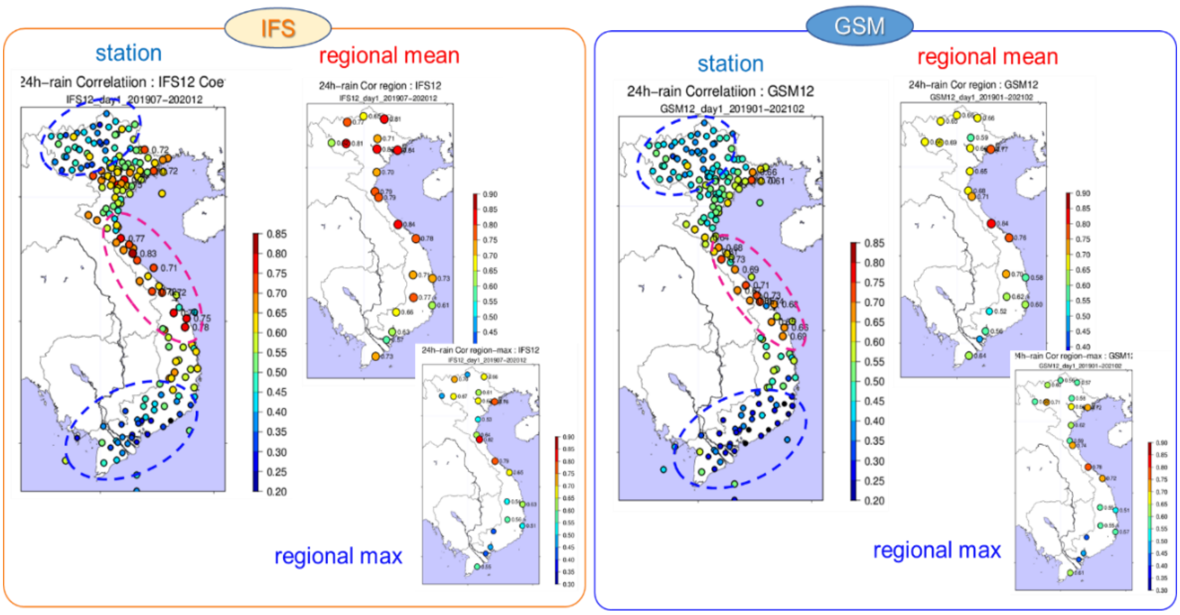

Figure 6. Distribution of correlation coefficients between 24-hour rainfall observations and IFS/GSM predictions at 186 SYNOP stations and in 21 regions.

3.2. Preliminary investigation for POP guidance

3.2.1. Monthly days with 24-hour rainfall of 1mm or more and 30 mm or more

Toward the development of POP guidance, we created two-year datasets of 24-hour rainfall observations and GSM predictions from Mar 2020 to Mar 2022 for 186 stations, and calculated the number of monthly rainy days with daily rainfall of 1 mm or more and the number of monthly heavy rainy days with daily rainfall of 30 mm or more. Figure 7 shows monthly days with 24-hour rainfall of 1mm or more and 30 mm or more in Lao Cai, Hanoi, Da Nang and Ho Chi Minh. Rainy days were more frequent from August to October in Hanoi and Ho Chi Minh, in August in Lao Cai, and from October to December in Da Nang at least once every two days. Heavy rainfall was most common from August to October in Lao Cai, Hanoi and Ho Chi Minh, about once every several days. In Da Nang, heavy rainfall was especially frequent in October, about once every three days.

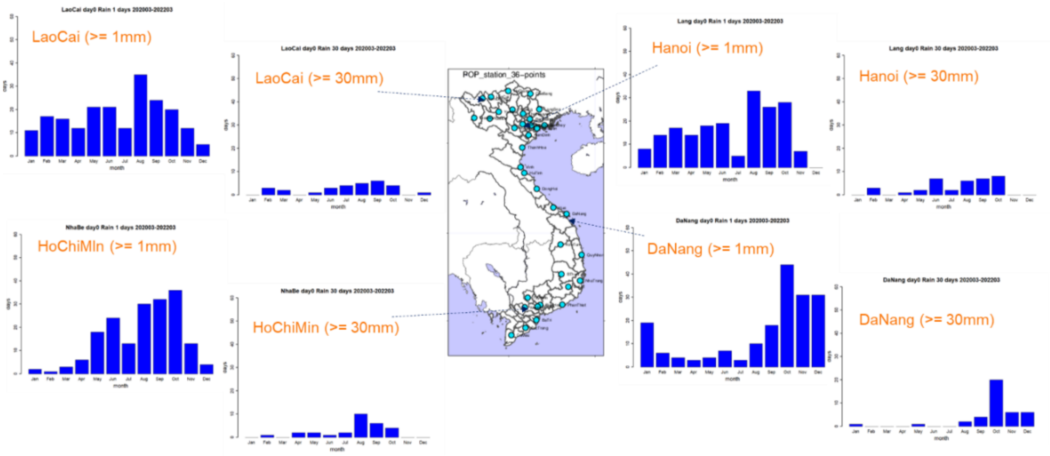

Figure 7. Monthly days with 24-hour rainfall of 1mm or more and 30mm or more in the two–year period from March 2020 to March 2022.

3.2.2. Verification of POP trial guidance for 4 stations and 4 regions

We created 24-hour rainfall observations and GSM predictions datasets for 4 stations and 4 regions (Lao Cai, Ha Noi, Da Nang and Ho Chi Minh) for POP trial guidance and its verification. Development period is from Mar 2020 to Sep 2021 and verification period is Jul 2018 to Jun 2020. One major station in each relevant region was used to calculate the station POP and several stations in each relevant region were used to calculate the regional POP.

Figure 8 shows reliability diagrams of POPs at 4 stations of Lao Cai, Lang (Ha Noi), Da Nang and Nha Be (Ho Chi Minh) for Day 1 (tomorrow). Daily calculated POPs (0 to 1) were subtotaled every 0.1, and the average of them and the average of the corresponding POP observations were plotted on the diagram. It's ideal if the plot rides on a yellow straight line. Brier Skill Score (BSS) and correlation coefficient between POP observations and calculated POPs are also shown on the top of each diagram. The plots for Lao Cai and Lang (Ha Noi) were generally on the yellow line, while those for Da Nang and Nha Be (Ho Chi Minh) were a bit more scattered. BSSs ranged from 0.26 to 0.29. Figure 9 shows reliability diagrams of POPs in 4 regions of Lao Cai, Ha Noi, Da Nang and Ho Chi Minh for Day 1. The plots for the 4 regions were generally on the yellow line and BSSs were considerably higher than those of the POPs at 4stations, ranging from 0.39 to 0.58.

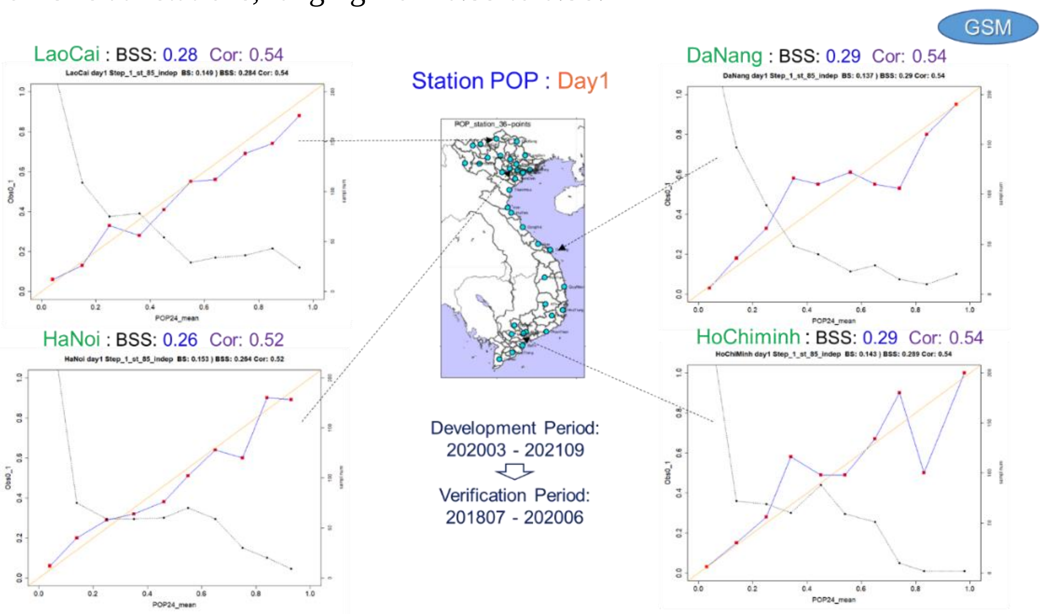

Figure 8. Reliability diagrams of station POPs at 4 stations for Day1.

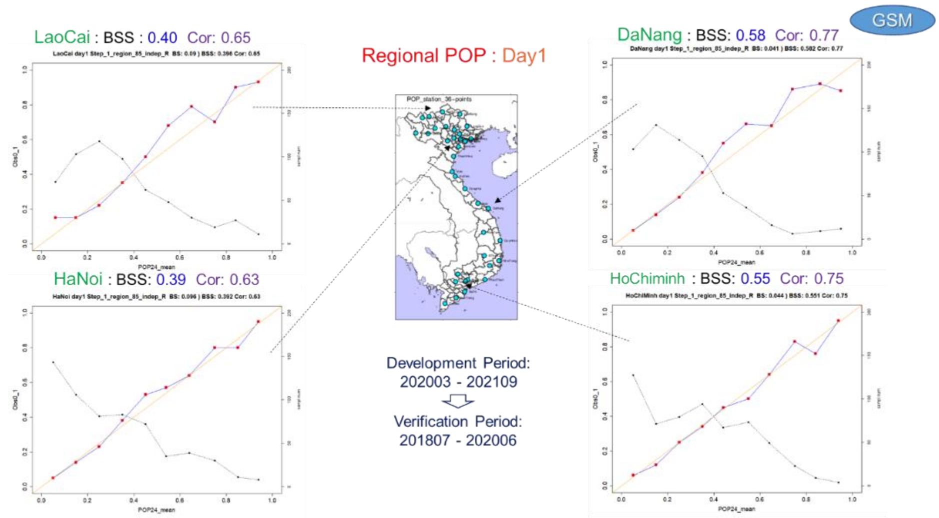

Figure 9. Reliability diagrams of regional POPs in 4 regions for Day1.

Looking at reliability diagrams at the 4 stations and in the 4 regions for Day 5 (5-days ahead), the variation in the diagrams increased than those for Day 1, and BSSs of station POPs were much lower ranging from 0.12 to 0.24, while BSSs of regional POPs were ranging from 0.19 to 0.42. BSSs of regional POPs were considerably lower than those for Day 1, but considerably higher than BSSs of station POPs for Day 5 (Figure not shown).

3.3. Preliminary investigation for regional mean and max 24-hour rainfall guidance

Following the verification of trial POP guidance in 4 regions, we calculated regional mean and max 24-hour rainfall guidance and conducted verification with the same datasets used for POP trial guidance. Figure 10 shows scatter diagrams of regional mean 24-hour rainfall observations and regional mean 24-hour rainfall guidance for Lao Cai, Ha Noi, Da Nang and Ho Chi Minh. Scatter diagrams for Lao Cai and Ho Chi Minh City showed relatively little variation, while those for Hanoi and Da Nang showed a large variation. Correlation coefficients between regional mean 24-hour rainfall observations and guidance for Hanoi and Danang were relatively low (0.5 and 0.56 respectively), while those for Lao Cai and Ho Chi Minh were relatively high (0.7 and 0.71 respectively). Scatter diagrams of regional max 24-hour rainfall observations and regional max 24-hour rainfall guidance for the 4 regions showed a large variation in all regions, with correlation coefficients ranging from 0.52 to 0.68 (Figure not shown).

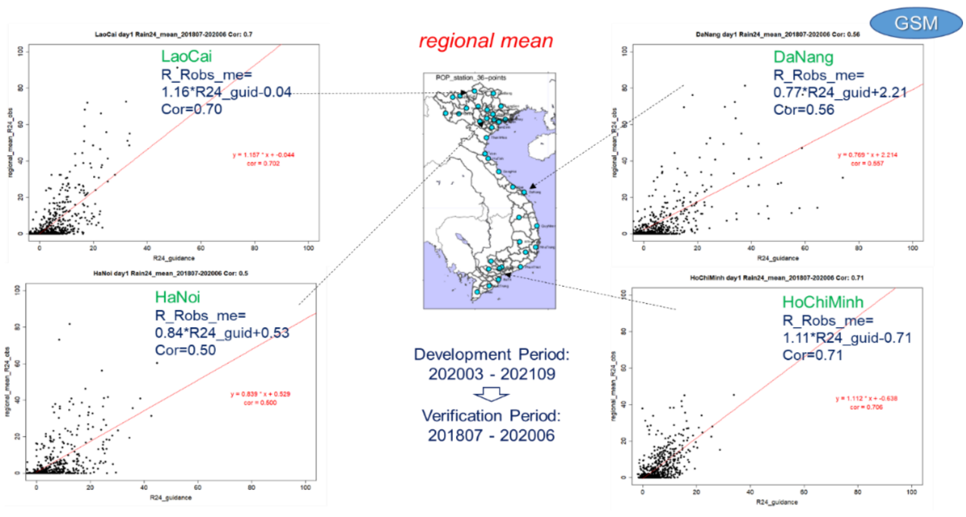

Figure 10. Scatter diagrams of regional mean 24-hour rainfall observations and regional mean 24-hour rainfall guidance in 4 regions for Day1.

3.4. Calculation of regional POP, regional mean and max 24-hour rainfall for 36 regions

In the preliminary investigation, Vietnam was divided into 21 regions to calculate regional mean rainfall etc. We however decided to subdivide Vietnam into 36 regions to make the regional allocation as close as possible to the forecast areas of NCHMF's city forecasts (5 municipalities and 58 provinces). It was desirable for a region to contain more than several stations, so if a province had only a few stations, it was combined with one or two neighboring provinces to form a region. New rainfall observation-GPV datasets for the 36 regions were created using 186 SYNOP station data for the period from Mar 2021 to Mar 2022 for IFS and from Mar 2020 to Mar 2022 for GSM, and regional POP, regional mean, and max rainfall guidance for the 36 regions were calculated and verified with the datasets.

3.4.1. Brier Skill Score of POP12 and POP24 for 36 regions

In line with the forecast period of NCHMF's city forecasts, we decided to prepare 12-hour precipitation guidance for tonight and tomorrow daytime, and 24-hour precipitation guidance after the day after tomorrow. Figure 11 shows BSSs of regional POP12 of IFS guidance and GSM guidance for tonight and tomorrow daytime in the 36 regions. BSSs of both IFS and GSM regional POP12 for tonight and tomorrow daytime were higher than 0.4 in most regions in the north and central. There were several regions with BSSs of less than 0.3 in the south for tonight in both IFS and GSM guidance. Comparing IFS and GSM, many regions had a slightly higher BSSs of POP12 in IFS, although the statistical periods were different.

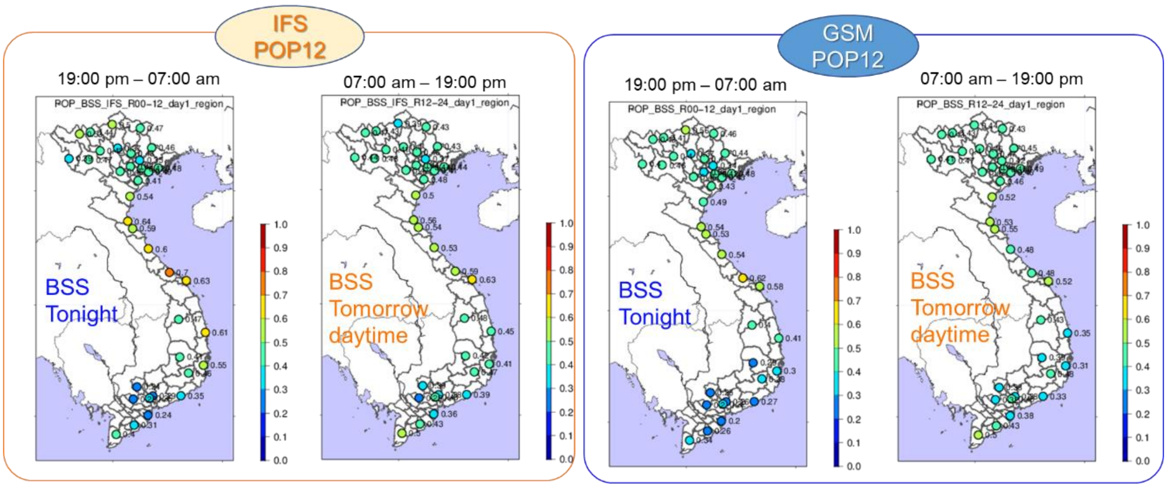

Figure 11. Distribution of Brier Skill Scores (BSSs) of regional IFS POP12 guidance and GSM POP12 guidance for tonight and tomorrow daytime for 36 regions. BSSs were calculated with dependent data (IFS: Mar 2021 to Mar 2022; GSM: Mar 2020 to Mar 2022).

Figure 12 shows BSSs of regional IFS POP24 and GSM POP24 for Day 1 and Day 5 in the 36 regions. BSSs of regional POP24 for Day 1 were higher than 0.5 in most regions in both IFS and GSM. Those for Day 5 were lower than POP24 for Day 1, but higher than 0.3 in most regions in both IFS and GSM. BSSs of IFS POP24 were a little bit higher than GSM POP24 in most regions, and BSSs of POP24 in the central were higher than those in the north and south.

In addition, POHP was calculated and verified similarly to the POP. However, BSSs of POHP were considerably lower than those of POP due to the small sample of heavy rainfall. BSSs were less than 0.2 in many regions even in the dependent data especially for long forecast periods (Figure not shown). It is considered necessary to accumulate heavier rain sample data for the development of POHP guidance.

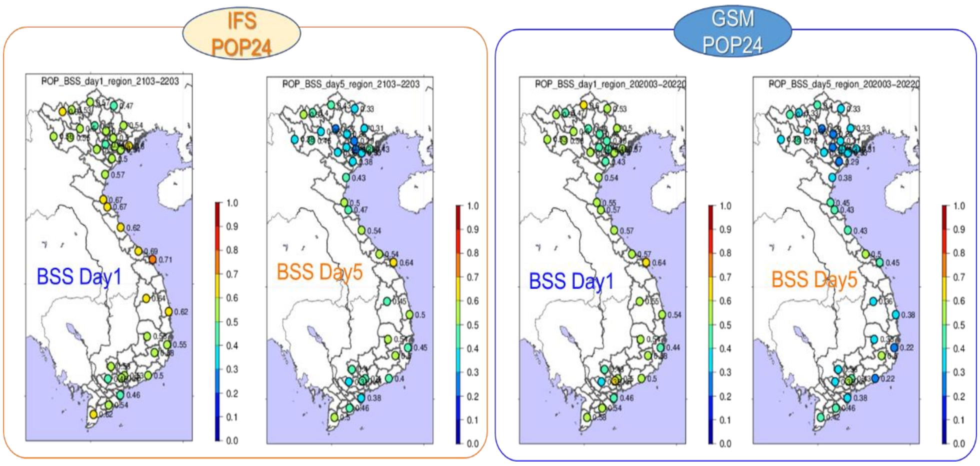

Figure 12. Distribution of Brier Skill Scores (BSSs) of regional IFS POP24 guidance and GSM POP24 guidance for Day1 and Day 5 for 36 regions. BSSs were calculated with dependent data (IFS: Mar 2021 to Mar 2022; GSM: Mar 2020 to Mar 2022).

3.4.2. IFS-GSM integrated precipitation guidance

We then developed IFS-GSM integrated precipitation guidance by using IFS and GSM predictors together. Figure 13 shows BSSs of POP and POHP with dependent data averaged over the 36 regions in IFS, GSM and IFS-GSM guidance for tonight, tomorrow daytime, Day 1 and Day 3. BSSs of POP12 and POP24 in IFS-GSM guidance were slightly higher than those in IFS and GSM guidance for all forecast periods. As for POHP12 and POHP24, BSSs in IFS-GSM guidance were higher than those in IFS and GSM guidance but were generally much lower than BSSs of POP guidance.

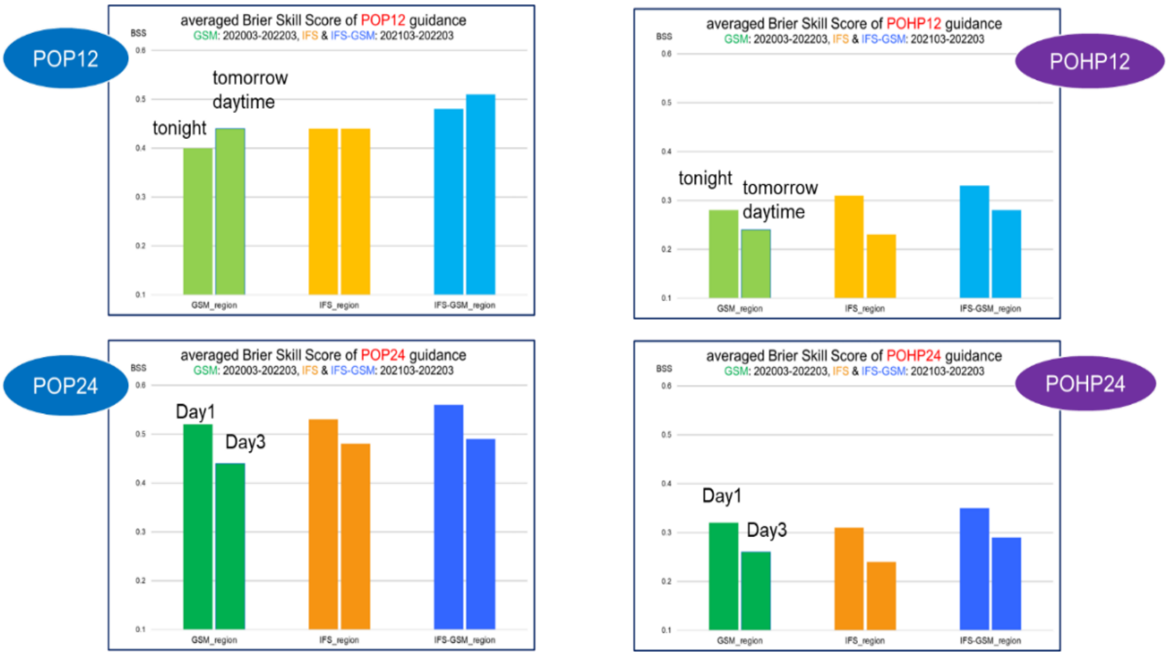

Figure 13. Averaged BSSs of POP12, POHP12, POP24 and POHP24 guidance by GSM, IFS and IFS–GSM for tonight, tomorrow daytime, Day1 and Day3.

Figure 14 shows scatter diagrams of regional mean and max 24-hour rainfall of IFS-GSM guidance in Hue region for Day 1. Correlation coefficients with dependent data were quite high at 0.88 and 0.83 for regional mean and max 24-hour rainfall respectively, but the variation was quite large, especially in maximum rainfall.

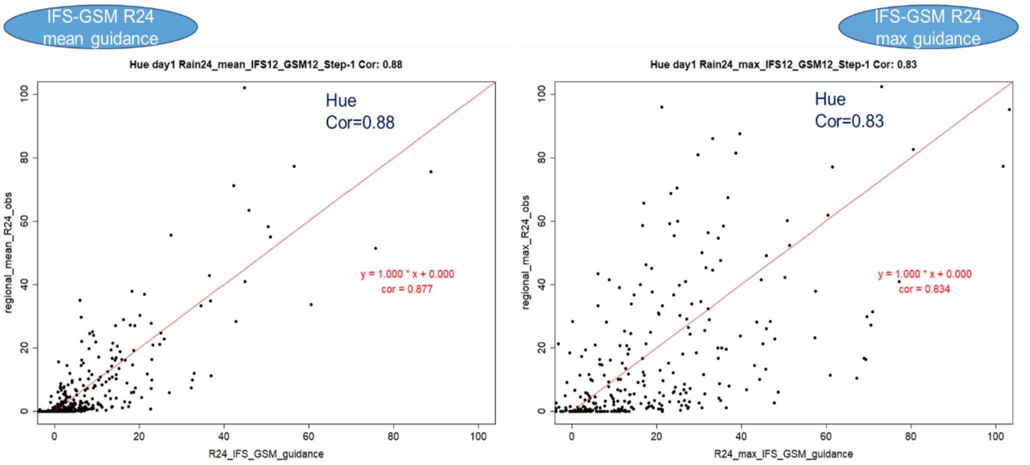

Figure 14. Scatter diagrams of regional mean/max 24-hour rainfall observations and IFS–GSM regional mean and max 24-hour rainfall guidance for Hue, Day1.

3.5. Daily calculation of precipitation guidance for 63 provinces

Based on the verification results, we decided to develop precipitation guidance on regional POP12 (tonight, tomorrow daytime) and POP24 (Day 2 to Day 5), regional mean and max 12-hour rainfall (tonight, tomorrow daytime) and 24-hour rainfall (Day 2 to Day 5) for 5 municipalities and 58 provinces up to 5-day ahead. Precipitation guidance is calculated for 36 regions and output for 5 municipalities and 58 provinces. Figure 15 is an example of precipitation guidance for August 19, 2022. It includes POP12, regional mean and max 12-hour rainfall for tonight and tomorrow daytime, POP24, regional mean and max 24-hour rainfall from the day after tomorrow to 5 days ahead. In addition, IFS and GSM regional mean rainfall predictions are included for forecaster's reference.

IFS-GSM precipitation guidance 19th Aug 2022 : 63 provinces for next 5 days

63 provinces

Tonight: 19th 19:00pm-20th 07:00am

POP12 for tonight

regional mean 12h-rain guidance

regional max 12h-rain guidance

IFS regional mean

GSM regional mean

Tomorrow daytime 20th 07:00am-19:00pm

Day2 20st 19pm- 21st 19:00pm

POP24 for Day2

regional mean 24h-rain guidance

regional max 24h-rain guidance

Station	Tonight	POP12_0	R12me_0	R12mx_0	IFSme_0	GSMme_0	Tomorrow	POP12_1	R12me_1	R12mx_1	IFSme_1	GSMme_1	Day2	POP24_2	R24me_2	R24mx_2	IFSme_2	GSMme_2	Day3
TamDuong	19-Aug	70	14.9	29.8	16	9.7	20-Aug	30	0.9	4.01	1.4	4.2	21-Aug	70	18.8	42.5	17.9	14.8	22-Aug
DienBien	19-Aug	70	11.5	21.5	17.7	13.6	20-Aug	30	1.4	5.96	1.1	2.6	21-Aug	70	8.2	26.4	25.1	10.8	22-Aug
SonLa	19-Aug	60	8.2	20.9	8.2	10.6	20-Aug	40	4.4	17.9	6.2	6	21-Aug	60	10.3	23.6	16	7.5	22-Aug
HoaBinh	19-Aug	60	8.3	18.2	2.4	16.3	20-Aug	70	16.1	26.4	16.9	15.8	21-Aug	100	25.1	31.5	29.2	14.8	22-Aug
LaoCai	19-Aug	60	10.1	23	18.1	15.6	20-Aug	40	3.8	9.69	8	2.4	21-Aug	80	16	36.1	25.5	14.2	22-Aug
YenBai	19-Aug	60	7.9	13.3	17.1	7.1	20-Aug	50	4.3	15.5	14.8	4.7	21-Aug	90	16	30.2	28.3	16.2	22-Aug
HaGiang	19-Aug	70	14.9	34.9	12.7	9.4	20-Aug	40	3.6	13.3	8.4	5.5	21-Aug	80	20.2	53.6	17	15.2	22-Aug
TQuang	19-Aug	60	13.6	33.1	25.3	8.4	20-Aug	60	8.2	22.4	11.5	12.2	21-Aug	80	17.3	42.9	27.8	22	22-Aug
BacCan	19-Aug	60	13.6	33.1	25.3	8.4	20-Aug	60	8.3	22	11.5	12.2	21-Aug	80	17.3	42.9	27.8	22	22-Aug
ThaiNguyen	19-Aug	60	12.5	28.1	15.2	13.2	20-Aug	60	11.1	22.3	11.7	16.9	21-Aug	90	22.1	42.2	27.5	24.1	22-Aug
VietTri	19-Aug	60	12.5	28.1	15.2	13.2	20-Aug	60	10.9	22.3	11.7	16.9	21-Aug	90	22.1	42.2	27.5	24.1	22-Aug
VinhYen	19-Aug	60	12.5	28.1	15.2	13.2	20-Aug	60	10.9	22.3	11.7	16.9	21-Aug	90	22.1	42.2	27.5	24.1	22-Aug
CaoBang	19-Aug	40	4.3	4.27	8.7	1.9	20-Aug	50	7.6	10.9	7	7.7	21-Aug	70	8.9	17.2	12.6	8	22-Aug
LangSon	19-Aug	50	8.2	21.6	5.4	6	20-Aug	70	7.6	23.6	10.6	9	21-Aug	60	11.7	30.4	23.2	5.9	22-Aug
BaiChay	19-Aug	50	9.4	26.8	6.9	4.4	20-Aug	80	22.6	29.1	23.5	6	21-Aug	100	33.8	77.1	44.5	14.4	22-Aug
BacGiang	19-Aug	50	7.3	18.3	9.9	5.7	20-Aug	70	8.4	14.3	18.6	16.7	21-Aug	70	10.7	21	21.4	13.3	22-Aug
BacNinh	19-Aug	50	7.3	18.3	9.9	5.7	20-Aug	70	8.4	14.4	18.6	16.7	21-Aug	70	10.7	21	21.4	13.3	22-Aug
PhuLien	19-Aug	60	12	16.7	13.3	6.7	20-Aug	80	17.7	31.4	20.7	6.2	21-Aug	100	25.1	46.2	30.8	15.9	22-Aug
Lang	19-Aug	60	6.4	14.8	10.5	10.1	20-Aug	80	12.5	21.9	15	17.8	21-Aug	90	27.7	40.1	27.2	23.5	22-Aug
HaiDuong	19-Aug	50	9	17.4	8	7.8	20-Aug	70	17.7	24.3	19.4	13.8	21-Aug	80	19.7	32.4	25.7	19.5	22-Aug

Figure 15. Example of IFS–GSM precipitation guidance for 63 provinces up to 5 days ahead for 19th Aug 2022.

3.6. Examples of precipitation guidance for heavy rain events in 2022

3.6.1. Heavy rain event caused by Typhoon NORU

Intense Typhoon NORU hit central Vietnam on the night of 27th September 2022 and brought heavy rainfall exceeding 300 mm in 12 hours in central Vietnam. Track forecasts of NORU by most NWP models fairly accurately predicted that it would make landfall in central Vietnam and bring heavy rainfall from the night of the 27th September to the next morning. Figure 16 shows the track of NORU analyzed by RSMC Tokyo, predictions of surface pressure and 6-hour rainfall around NORU for 28th 00UTC Sep by GSM 26th 12UTC initial.

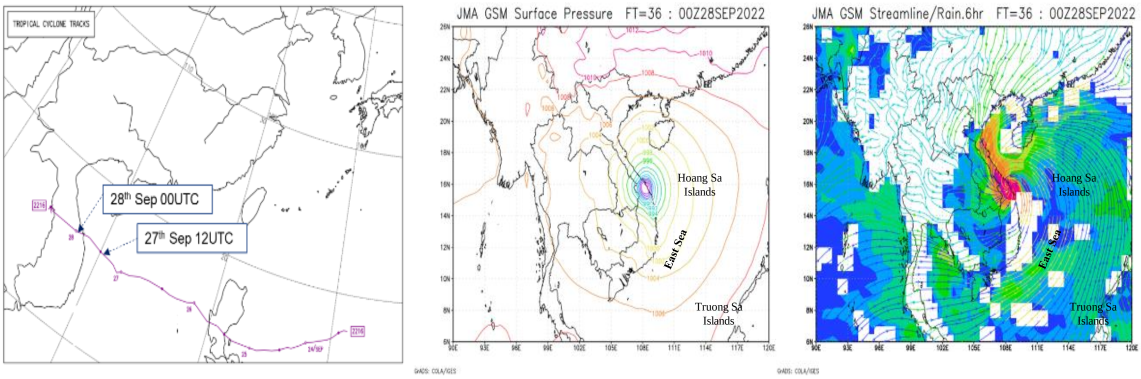

Figure 16. Track of TY NORU analyzed by RSMC Tokyo, 36–hour predictions of Surface pressure, streamline and 6–hour rainfall around NORU by GSM 26th 12UTC initial.

Three days prior to landfall, the 24-hour maximum rainfall predictions by IFS, NCEP GFS and GSM from the night of the 27th to the next day were 200-250 mm, 250-300 mm and 250-300 mm respectively, in central Vietnam (Figure 17). Figure 18 shows regional mean and max 24-hour rainfall guidance for 25th September 2022 based on IFS, GSM, and IFS-GSM predictions of 24th 12UTC initial. IFS guidance and GSM guidance had the highest rainfall of about 250 mm (IFS) and 420 mm (GSM) around Hue, while IFS-GSM guidance had the highest rainfall of about 350 mm around Da Nang.

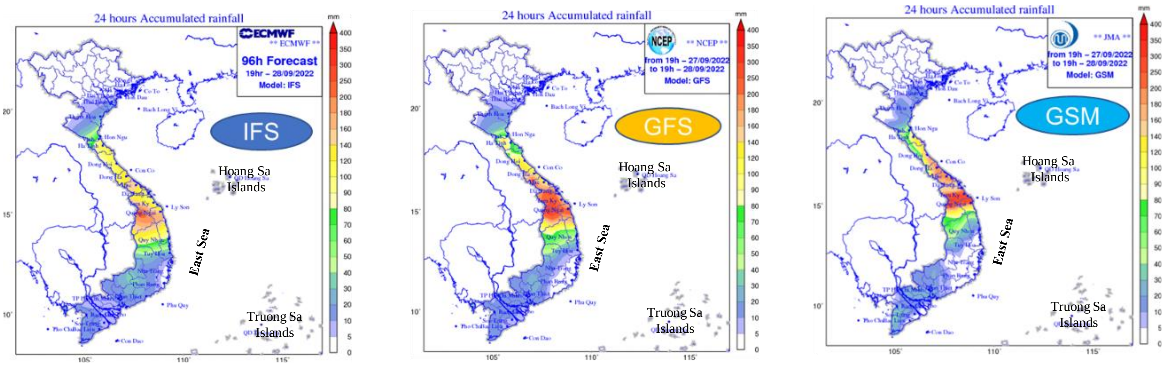

Figure 17. 24-hour rainfall predictions for 28th Sep by IFS, GFS and GSM of 24th 12UTC initial.

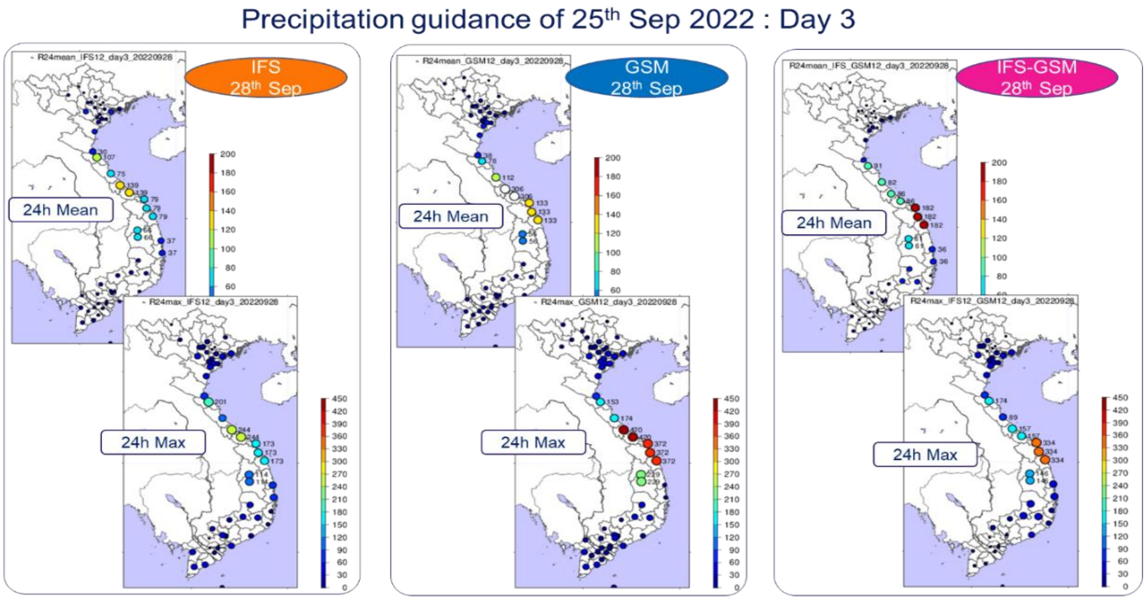

Figure 18. IFS, GSM and IFS-GSM precipitation guidance (regional mean/max 24-hour rainfall) of 25th Sep for Day 3 (28th Sep) based on 24th Sep 12 UTC initial.

Looking at 24-hour rainfall observations for the corresponding time period, the regional max rainfall in Hue and Da Nang was almost the same at 330 mm and 340 mm, respectively, and the regional mean rainfall was almost the same at 160 mm (Figure 19). Comparing the guidance to observations, IFS guidance had considerably too little rainfall in the Da Nang region and GSM guidance had considerably too much mean rainfall in the Hue region; IFS-GSM guidance had a fairly good representation of rainfall in the Da Nang region but less in the Hue region. All guidance was too little for rainfall in the Vinh region.

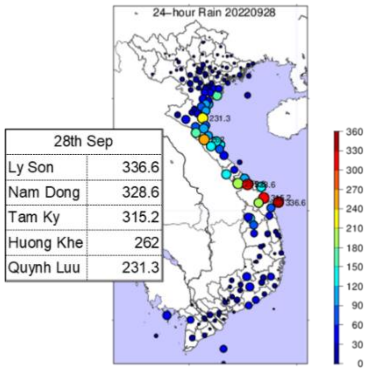

Figure 19. 24-hour rainfall observation for 19:00 27th – 19:00 28th Sep 2022.

One day prior to landfall, the 24-hour maximum rainfall predictions by IFS, GFS, GSM and VNMHA WRF for the night of the 27th and next day were 200-250 mm, 300-350 mm, 250-300 mm, and more than 400 mm respectively, in central Vietnam (Figure 20). Figure 21 shows regional mean and max 12-hour rainfall guidance of 27th for 27th tonight based on IFS, GSM, and IFS-GSM predictions of 26th 12UTC initial. Regional mean and max 12-hour rainfalls of IFS guidance were mean: 90 mm, max: 180 mm around Hue, mean: 60 mm, max: 180 mm around Da Nang, those of GSM guidance were mean:170 mm, max: 270 mm around Hue, mean: 120 mm, max: 280 mm around Da Nang and those of IFS-GSM guidance were mean: 90 mm, max: 250 mm around Hue, mean: 170 mm, max:430 mm around Da Nang.

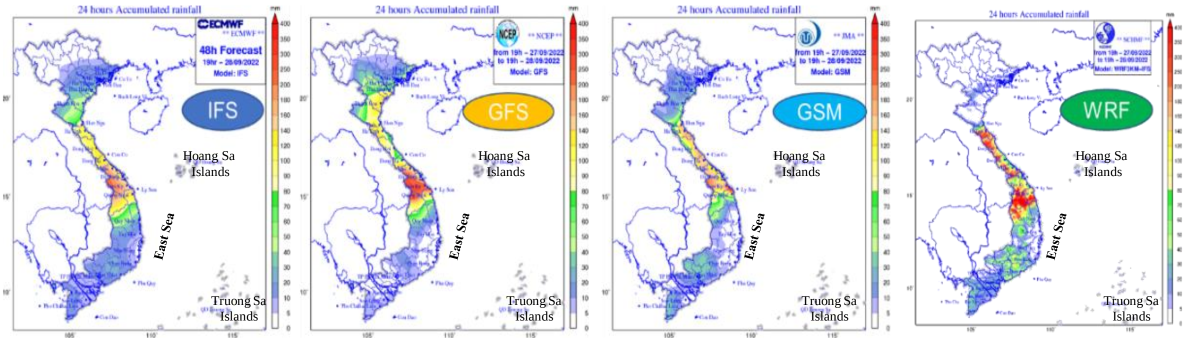

Figure 20. 24-hour rainfall predictions for 28th Sep by IFS, GFS and GSM of 26th 12UTC initial.

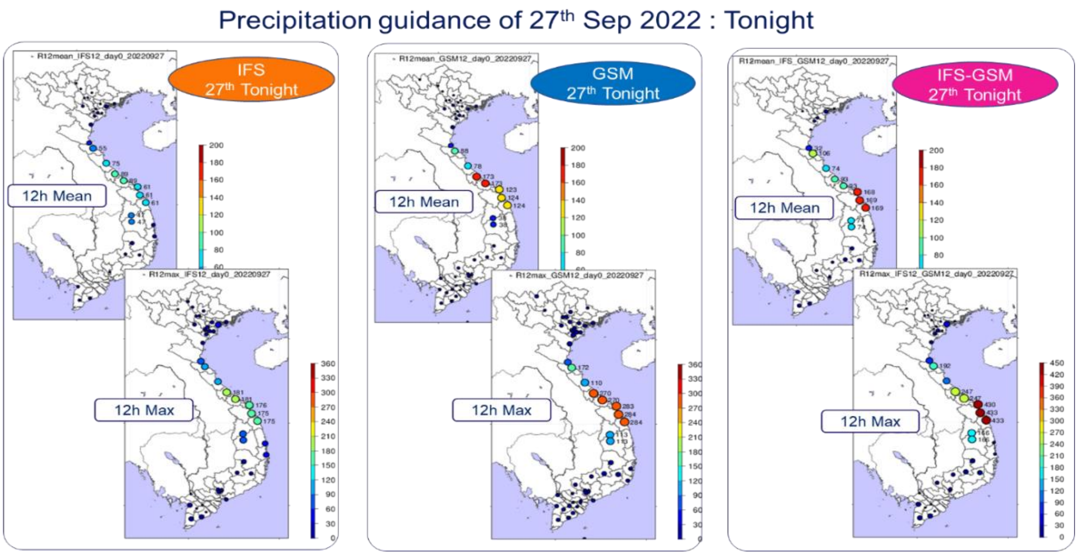

Figure 21. IFS, GSM and IFS-GSM precipitation guidance (regional mean/max12-hour rainfall) of 27th Sep for 27th tonight based on 26th Sep 12 UTC initial.

Looking at 12-hour rainfall observations, the regional max rainfall in Hue and Da Nang was almost the same at 330 mm, and the regional mean rainfall was also almost the same at 150 mm (Figure 22a). Comparing the guidance to observations, IFS guidance had considerably too little rainfall in the Da Nang region and GSM guidance was generally adequate rainfall in both Hue and Da Nang regions IFS-GSM guidance was slightly less for Hue region and more for Da Nang region. All guidance was generally adequate around Vinh region on the night of 27th September, but too little for the next daytime.

3.6.1. Heavy rain event with strong northeasterly winds blowing

Strong northeasterly winds blew into central Vietnam and heavy rainfall of 400-600 mm was observed in central Vietnam from the night of 9th to daytime of 10th October 2022. On that day, well-developed cloud systems and very strong radar QPE developed by Output 2 radar expert team were continuously observed over central Vietnam (Figure 22b) [13]. QPE provides detailed information on heavy rainfall that cannot be obtained from satellite imagery. Figure 23 shows predictions of surface winds and 6-hour rainfall and winds at 850hPa for 10th 00UTC Oct by GSM 7th 12UTC Oct initial.

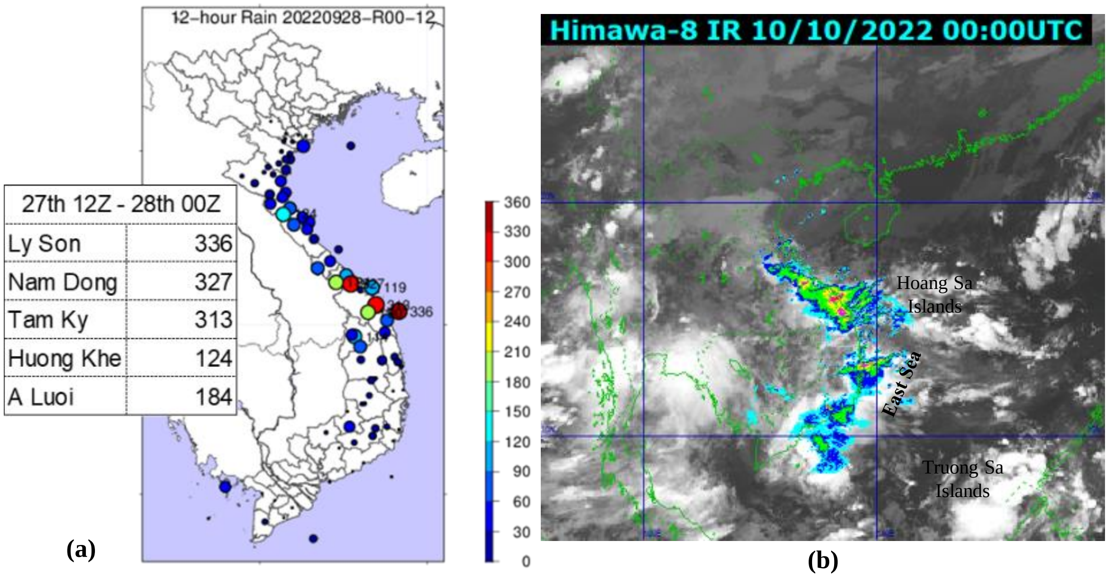

Figure 22. (a) 12-hour rainfall observation for 19:00 27th – 00:00 28th Sep 2022; (b) Radar QPE overlayed on IR image for 10th 00UTC Oct 2022.

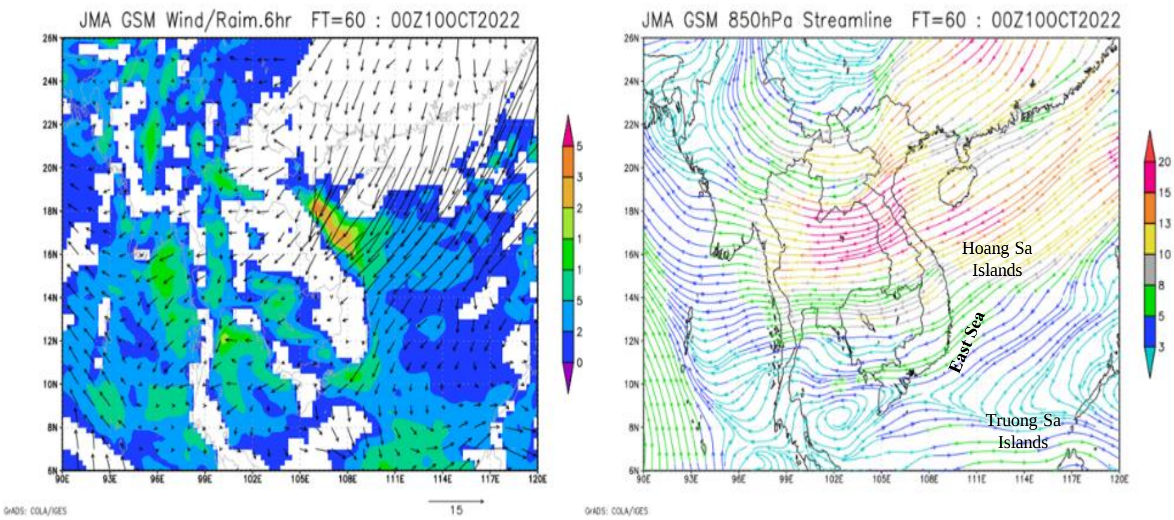

Figure 23. Surface winds and 6-hour rainfall, 850hPa stream lines for 10th 00UTC Oct 2022.

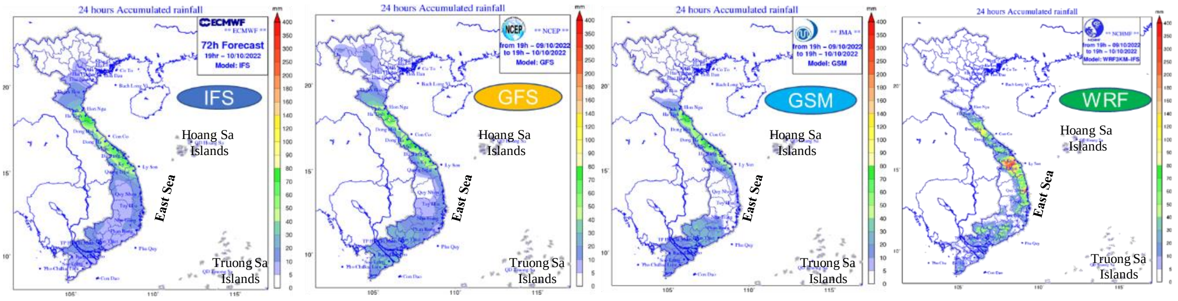

Figure 24. 24-hour rainfall predictions for 10th Oct 2022 by IFS, GFS and GSM of 7th Oct 12UTC.

IFS, GFS and GSM all predicted strong northeasterly winds blowing over central Vietnam, bringing 70-90 mm of rainfall during the time period, but they did not predict any heavy rains reaching 600 mm. WRF predicted heavy rainfall of 200-250 mm for some regions in central Vietnam (Figure 24). Regional max 24-hour rainfall of precipitation guidance for 8th Oct 2022 based on IFS-GSM predictions of 7th 12UTC initial was about 150 mm around Hue, much less than actual observations (Figure 25).

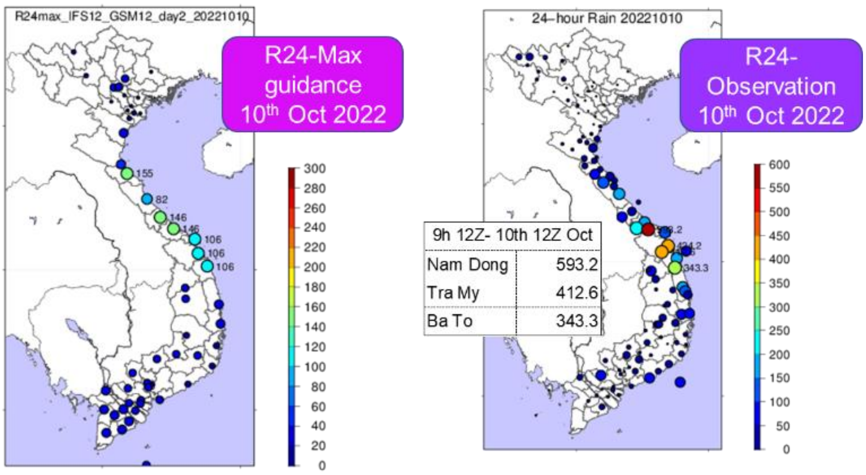

Figure 25. IFS-GSM precipitation guidance (regional max24-hour rainfall) of 8th Oct for 10th Oct 2022 based on 7th Oct 12 UTC initial.

In 2022, there were many heavy rain events caused by tropical cyclones. As for tropical cyclones, IFS and GSM predicted a little less rainfall but a reasonable amount, so precipitation guidance was generally able to predict a reasonable amount of rainfall for such heavy rain events caused by tropical cyclones. However as in the case of a heavy rain event with strong northeasterly winds blowing, there were several events where IFS and GSM rainfall predictions were quite low, and rainfall guidance was also quite low. Post-analysis of various heavy rainfall cases should be accumulated to improve rainfall precipitation guidance.

4. Conclusions

Precipitation guidance was developed with logistic regression and multiple linear regression for 36 regions up to 5 days ahead using JMA GSM GPV data and ECMWF IFS GPV data as an activity of the second phase of the JICA Project for Strengthening Capacity in Weather Forecasting and Flood Early Warning System. 2-year datasets of 24-hour rainfall observations and GSM/IFS predictions were first created at each station and in each region. Statistical analysis with the datasets showed correlation coefficients between 24-hour rainfall observations and IFS/GSM predictions were higher in the central region and lower in the northern mountainous region and south region. The correlation coefficients of regional mean 24-hour rainfall were considerably higher than those of the stations in each region. Verification of POP trial guidance for 4 stations and 4 regions showed reliability diagrams were generally good for Day 1 for both at station and in region. BSSs of regional POP, which ranged from 0.39 to 0.58, were considerably higher than those of station POP ranged from 0.26 to 0.29. Scatter diagrams of regional mean 24-hour rainfall observations and guidance for 4 regions showed a large variation and correlation coefficients of them were relatively low in Hanoi and Da Nang (0.5 and 0.56) and relatively high in Lao Cai and Ho Chi Minh City (0.7 and 0.71).

Based on the verification results, we decided to develop precipitation guidance on regional POP12 (tonight, tomorrow daytime) and POP24 (Day 2 to Day 5), regional mean and max 12-hour rainfall and 24-hour rainfall for 36 regions (5 municipalities and 58 provinces) up to 5-day ahead. In addition to IFS and GSM guidance, we developed IFS-GSM-integrated precipitation guidance by using IFS and GSM predictors together and confirmed that its verification results were slightly better than those of IFS and GSM guidance. Test operation of the precipitation guidance started in September 2022.

Case studies of heavy rain events in September and October 2022 were conducted using observation data, NWP products and precipitation guidance. Precipitation guidance was generally able to predict a reasonable amount of rainfall for heavy rain events caused by tropical cyclones, however there were several events where IFS and GSM rainfall predictions were quite low, and rainfall guidance was also quite low. Post-analysis of various heavy rainfall cases should be accumulated to improve rainfall precipitation guidance.

Author Contributions: Conceptualization: K.S.; Methodology: K.S.; Software: K.S.; Observation and forecast data curation: V.T.A.; Verification: K.S.; Writing–original draft preparation: K.S.; Writing–review and editing: V.T.A.

Acknowledgments: Method of this work is based on training materials on guidance in the JICA group training course in meteorology being implemented by the Japan Meteorological Agency (JMA). We would like to thank JMA for their support. Ms. Do Thuy Trang of the NCHMF NWP division helped us in providing the IFS data. We would like to thank the Working Group 3 members, NCHMF staff members and JICA experts of this Project for their cooperation to our activities.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tonouchi, M.; Kasuya, Y.; Tanaka, Y.; Akatsu, K.; Akaeda, K.; Nguyen, V.T. Activities of JICA on disaster prevention and achievement of JICA project in Period 1. *VN. J. Hydrometeorol.* **2020**, *5*, 1–12.
2. Sasaki, K.; Anh, V.T.; Hang, N.T.; Trang, T. Development of maximum and minimum temperature guidance with Kalman filter for 63 cities in Vietnam up to 10 days ahead. *VN. J. Hydrometeorol.* **2020**, *5*, 51–64.
3. Glahn, H.R.; Lowry, D.A. The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meter.* **1972**, *11*, 1203–1211.
4. Klein, W.H.; Glahn, H.R. Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.* **1974**, *55*, 1217–1227.
5. Takada, S. Overview of Guidance. Special volume of numerical prediction division report. *Japan Meteorological Agency* **2018**, *64*, 1–18. (In Japanese)
6. Kudo, J. Guidance Development Techniques. Special volume of numerical prediction division report. *Japan Meteorological Agency* **2018**, *64*, 19–74. (In Japanese)

7. Shiroyama, Y. Precipitation Guidance. Special volume of numerical prediction division report. *Japan Meteorological Agency* **2018**, *64*, 95–118. (In Japanese)
8. Hoshina, M.; Yasutomi, Y.; Konoda, S. Guidance development techniques. research bulletin. *Japan Meteorological Agency* **1982**, *34*, 239–276. (In Japanese)
9. Ebihara, S. Improvement and verification of precipitation guidance. textbook for nwp training numerical prediction division. *Japan Meteorological Agency* **1999**, 23–33. (In Japanese)
10. Applequist, S.; Gahrs, G.E.; Pfeffer, R.L.; Niu, X.F. Comparison of methodologies for probabilistic quantitative precipitation forecasting. *Weather Forecasting* **2001**, *17*, 783–799.
11. Wilks, D.S. Statistical methods in the atmospheric sciences. 3rd Ed., Academic Press, Chapter 7, 2011, pp. 214–300.
12. Brier, G.W. Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.* **1950**, *78*, 1–3.
13. Kimpara, C.; Tonouchi, M.; Bui, T.K.H.; Nguyen, V.H.; Nguyen, M.C.; Akaeda, K. Quantitative precipitation estimation by combining rain gauge and meteorological radar networks in Viet Nam. *VN. J. Hydrometeor.* **2020**, *5*, 36–50.