IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 21, 2020, accepted January 18, 2021, date of publication February 18, 2021, date of current version March 3, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3057057

A Comparative Study of Four Merging Approaches for Regional Precipitation Estimation

ZEDONG FAN[1], WEIYUE LI[1,2,3], QIN JIANG[4,5], WEIWEI SUN[6,7,8], JIAHONG WEN[1], AND JUN GAO[1]

[1]School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai 200234, China
[2]Institute of Urban Studies, Shanghai Normal University, Shanghai 200234, China
[3]Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China
[4]Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China
[5]School of Geographic Sciences, East China Normal University, Shanghai 200241, China
[6]Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China
[7]Ningbo Universities Collaborative Innovation Center for Land and Marine Spatial Utilization and Governance Research at Ningbo University, Ningbo 315211, China
[8]Institute of East China Sea, Ningbo University, Ningbo 315211, China

Corresponding authors: Weiyue Li (lwy326@shnu.edu.cn) and Jiahong Wen (jhwen@shnu.edu.cn)

This work was supported in part by the Natural Science Foundation of Shanghai under Grant 19ZR1437500, in part by the National Natural Science Foundation of China under Grant 41730642, and in part by the Shanghai Gaofeng & Gaoyuan Project for University Academic Program Development and Shanghai Sailing Program under Grant 20YF1434900.

ABSTRACT To identify suitable merging methods to improve regional precipitation estimates using multiple sources of precipitation data, this study applied four different approaches (multiple linear regression (MLR), feedforward neural network (FNN), random forest (RF) and long short-term memory network (LSTM)) to merge four satellite precipitation products and one reanalysis data in the Jiangsu, Zhejiang and Shanghai of China. The pros and cons of the merging approaches are analyzed comprehensively, using correlation coefficient (CC), root mean square error (RMSE), relative bias (RB), probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) as evaluation indexes. Our results show that: (1) All merging approaches can improve the accuracy of precipitation estimations, but only RF and LSTM can improve the daily precipitation event detection capacity. These approaches can significantly reduce errors in moderate precipitation scenarios, but do not effectively improve accuracy in light and heavy precipitation scenarios. (2) MLR was the least expensive computing cost method in our study and performed better than the other three methods when gauge density was low. However, MLR had the worst daily precipitation event detection capacity (CSI = 0.67). (3) FNN performed moderately in most experiments (CC = 0.87, RMSE = 4.65 mm/day, RB = 1.19 %, POD = 0.94, FAR = 0.29, CSI = 0.70). (4) The merged data generated by RF was the most accurate and had the best daily precipitation event detection capacity (CC = 0.87, RMSE = 4.61 mm/day, RB = − 0.33 %, POD = 0.97, FAR = 0.20, CSI = 0.78). RF performed best in moderate precipitation scenarios. However, it performed worse than other methods when gauge density was low. (5) LSTM was the most robust methods and performed best in light precipitation scenarios. The FAR of the LSTM-generated data was the smallest (0.15) among four fusion methods. However, LSTM had the most expensive computing cost and the worst accuracy of the merged data (CC = 0.86, RMSE = 4.68 mm/day, RB = − 9.36 %).

INDEX TERMS Accuracy improvement, data merging, gridded precipitation data, precipitation estimation, robustness.

I. INTRODUCTION

Quantitative and accurate precipitation estimation is crucial for water resource management, natural disaster prevention, and risk management [1], [2]. However, current mainstream gridded precipitation products (e.g., interpolated precipitation products, satellite precipitation products and reanalysis datasets) have their own advantages and disadvantages. Generally, the quality of interpolated precipitation products is affected by the density of rain gauges and the consistency of the historical record periods [3], their spatial coverage is often limited to land, and they have thus larger errors in areas where the rain gauge networks are sparse [4]. Compared with interpolated data, satellite rainfall products

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

and reanalysis datasets have the advantages of quasi global spatial coverage. Satellite rainfall products can provide near real-time data [5], and reanalysis datasets have a long history of data records [6], which have been widely used in climate change research [7]. However, some studies have noted that there are some errors in these two types of data [8]. Varying sensor types, retrieval algorithms, and spatial sampling frequencies, can often cause satellite precipitation products to present regionally varying biases [9]. Furthermore, due to the influences of the assimilation model [10] and observation data error [8], there can be spatial differences in precipitation estimation accuracy [11]. Therefore, many scholars have attempted to obtain more accurate precipitation data by correcting these raw precipitation products or even by combining multiple products.

One of the most common examples of a combined approach is the multiple linear regression (MLR) algorithm. It has long been used in precipitation data fusion because it can quickly fit the weight of data for fusion and generate a relatively objective fusion scheme [12]. Moreover, ordinary kriging has also been applied to precipitation data fusion because it can improve the fused data's ability to capture extreme precipitation [13]. Machine learning algorithms have also been widely used in remote sensing data fusion because of their ability to solve nonlinear problems [14], [15]. For example, the merged result that was created utilizing artificial neural networks is far superior to all of the individual precipitation datasets used as inputs [16]. Baez-Villanueva *et al.* [17] used random forest (RF) to merge multi-source precipitation data, significantly improving the accuracy of the merged data.

With the improvement of computer hardware and the increase in the amount and availability of data, deep learning has become widely used in many fields [18], [15]. Deep learning can also be used to fuse precipitation data. Wu *et al.* [19] proposed a model to merge multi-source precipitation data that combined a convolutional neural network (CNN) and a long short-term memory network (LSTM). The model improved the accuracy of satellite data and was more accurate than multi-layer perception.

Although these approaches usually improve the general accuracy of precipitation estimation, there are still problems that can arise with their merged data. For example, the accuracy of the merged data can have obvious spatial differences in regional precipitation estimations, owing to influences from sources such as data error or parameter setting [17], [19]. Moreover, the robustness of some merged methods is low [20]. Although some researchers have evaluated the performance of precipitation products in various scenarios [21]–[23], there are few studies that have focused on the applicability of merging approaches. Therefore, this study attempts to identify the strengths and weaknesses of data fusion in improving the quality of precipitation data. To this end, we comprehensively compare the pros and cons of different merging approaches that are used to fuse the same data, instead of only comparing the total accuracy of different merged data.

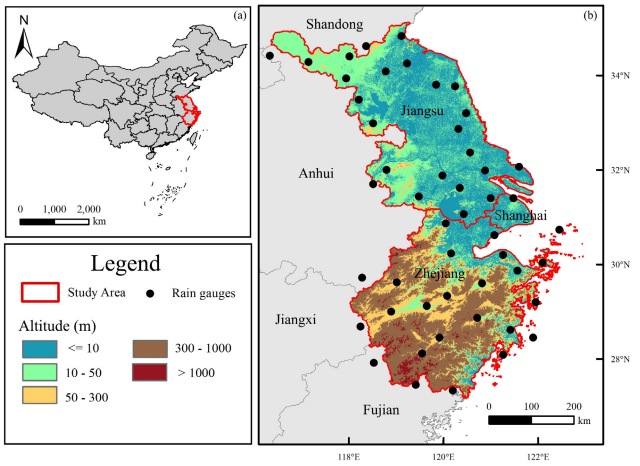


FIGURE 1. Study area and the distribution of rain gauges.

In this study, MLR, feedforward neural networks (FNN), RF, and LSTM are selected to merge five mainstream gridded precipitation products and subsequently generate merged data. MLR and FNN are selected for study because they have been used in the generation of two mainstream precipitation datasets to date: the CPC Merged Analysis of Precipitation (CMAP) and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) [12], [24]. RF and LSTM are chosen as they have become increasingly popular in recent years. In 2020, they were used to fuse precipitation datasets, creating merged results that were better than other mainstream precipitation datasets in some regions [17], [19].

In order to holistically evaluate these methods, this paper first outlines the study area and the datasets used in our study, including rain gauge observations, satellite precipitation products and a reanalysis dataset. Then, the merging approaches and accuracy verification methods are introduced in the METHODS section. The analysis of error and the daily precipitation event detection capacity for the four methods are presented in the RESULTS section, before being further analyzed in the following DISCUSSION. Finally, the study and its major results are summarized in the CONCLUSION.

II. STUDY AREA AND DATA

A. STUDY AREA

This study focused on the Jiangsu, Zhejiang and Shanghai, regions that cover an area of about 219000 km^2 and have dense rain gauge networks. This area has a subtropical monsoon climate. The terrain is high in the south and low in the north, with the altitude of most areas less than 1000 m, as shown in Fig. 1.

B. DATA

1) RAIN GAUGE DATA

The rain gauge data are from the daily dataset of surface climate data of China, which can be downloaded from the China

Meteorological Data Service Centre (http://data.cma.cn/). After quality control, 49 available rain gauges were selected and distributed in Jiangsu, Zhejiang, Shanghai, and their surrounding areas. In this study, the inverse distance weight interpolation method was used to interpolate the rain gauge data into gauge-based interpolated data (GID) with a spatial resolution of $0.25° \times 0.25°$, which was then used as the 'truth value' in the data fusion and accuracy tests.

TABLE 1. Details of the five gridded precipitation products.

Datasets	Spatial Resolution	Time Resolution	Study Period	Spatial Coverage
CMORPH-BLD	0.25°	1d	2003-2015	90°N-90°S
3B42V7	0.25°	3h	2003-2015	50°N-50°S
PERSIANN-CDR	0.25°	1d	2003-2015	60°N-60°S
CHIRPS	0.25°	1d	2003-2015	50°N-50°S
ERA5	0.25°	1h	2003-2015	Global

2) GRIDDED PRECIPITATION PRODUCTS

In total, five mainstream gridded precipitation products (four satellite gridded precipitation products and one reanalysis product) were selected for the merging test, including Tropical Rainfall Measuring Mission 3B42 version 7 (3B42V7) [25], Climate Prediction Center Morphing technique satellite-gauge blended product (CMORPH-BLD) [26], PERSIANN-CDR [24], Climate Hazards Group InfraRed Precipitation with Station data version 2.0 (CHIRPS) [27] and European Centre for Medium-range Weather Forecasts Reanalysis 5 (ERA5) [28]. Previous studies have shown that these products are able to reflect spatial distributions and temporal changes of precipitation [23], [29], [30]. A summary of these products is presented in Table 1.

a: 3B42V7

Tropical Rainfall Measuring Mission (TRMM) data is produced using the TRMM multi-satellite precipitation analysis (TMPA) method, which first combines microwave and infrared data to estimate precipitation, and then corrects these estimates using rain gauge data [25]. TMPA provides two kinds of standard precipitation products: near-real-time data 3B42RT and non-real-time analysis data 3B42. Version 7 of 3B42, provided by National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAEA), can be downloaded from https://pmm.nasa.gov/data-access/downloads/trmm, was used in this study.

b: CMORPH-BLD

Climate Prediction Center Morphing (CMORPH) data uses motion vectors derived from half-hourly infrared imagery to propagate precipitation estimates derived from passive microwave data [26]. CMORPH-BLD is a product that blends raw CMORPH data with rain gauge data. It is provided by the National Centers for Environmental Prediction (NCEP), and can be downloaded from ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/BLD/.

c: PERSIANN-CDR

PERSIANN-CDR is generated from infrared data using artificial neural networks, and is adjusted using the Global Precipitation Climatology Project (GPCP) monthly product [24]. PERSIANN-CDR is developed by the Center for Hydrometeorology and Remote Sensing (CHRS) at the University of California, Irvine (UCI), and can be downloaded from https://www.ncei.noaa.gov/data/precipitation-persiann/access/.

d: CHIRPS

CHIRPS blends satellite information, cold cloud duration (CCD) observations and rain station data to generate precipitation estimates [27]. CHIRPS is provided by the Climate Hazards Group (CHG) at the University of California, Santa Barbara (UCSB) and can be downloaded from http://chg.geog.ucsb.edu/data/chirps/.

e: ERA5

ERA5 is the latest reanalysis product of the European Centre for Medium-Range Weather Forecast (ECMWF), and is produced using the Integrated Forecast System (IFS) cycle 41r2 with four-dimensional variational analysis [28]. ERA5 is provided by ECMWF, and can be downloaded from https://cds.climate.copernicus.eu/cdsapp#!/home.

III. METHODS

A. DATA PREPROCESS

First, the rain gauge data were interpolated into gridded data. Then, the time benchmarks (UTC) of satellite precipitation and reanalysis products were matched to the observation time (GMT + 8) of the rain gauges. All data were accumulated on a daily scale for normalization. In this study, data from 2003 to 2013 (4018 samples) were selected as the training set, data from 2014 (365 samples) were selected for parameter validation, and data from 2015 (365 samples) were selected as the test set. The simulations were executed on a computer running Windows 10, with an Intel i7-67000 (3.40 GHz) CPU and 20GB of RAM. The code was programmed in Python 3.6.

B. MERGING APPROACHES

The precipitation data were fused, pixel by pixel, using four different methods. The fusion approaches used in this study include: MLR, FNN, RF, and LSTM.

1) MULTIVARIATE LINEAR REGRESSION

Linear algorithms have been applied in many precipitation data fusion studies [4], [31]. The linear algorithm used in this study is as follows:

$$P_{MLR} = \alpha_0 \cdot P_0 + \ldots + \alpha_4 \cdot P_4 + \beta \quad (1)$$

where $P_0 \ldots P_4$ are the precipitation values of CMORPH-BLD, 3B42V7, PERSIANN-CDR, CHIRPS and ERA5, respectively; $\alpha_0 \ldots \alpha_4$ are the corresponding regression coefficients of $P_0 \ldots P_4$, and β is the intercept. All values are defined by the least squares method. P_{MLR} is the merged result based on MLR.

2) FEEDFORWARD NEURAL NETWORK

FNN is a popular neural network. The information in FNN only travels forward and the connections between the nodes do not form cycles [32]. FNN has good self-learning and self-organizing ability and can simulate complex nonlinear mapping [33], [34]. The FNN model used in this study consists of three layers: an input layer, a hidden layer, and an output layer. The input variables are the precipitation values of CMORPH-BLD, 3B42V7, PERSIANN-CDR, CHIRPS and ERA5. The output is the merged result based on the FNN.

According to the tutorial of scikit-learning [35], the limited memory BFGS (L-BFGS) method can converge faster and perform better than other methods for small datasets (detail about L-BFGS can be found in [36]). Therefore, L-BFGS was selected as the optimizer of the FNN model in this study. The number of neurons in the hidden layer, n_{FNN}, was determined from the candidate set {6, 7, 8, 9, 10, 11, 12}. For this study, n_{FNN} was set to 11, as this number allowed the FNN model to perform best for the validation dataset. Similarly, the learning rate and the number of iterations of the FNN were set as 0.05 and 2000, respectively.

3) RANDOM FOREST

RF is a type of algorithm proposed by [37], which performs well when facing classification and regression problems [38]–[40]. RF has been used in precipitation data fusion [17] and precipitation downscaling [41]. The RF model used in this study is as follows:

$$P_{RF} = \frac{1}{\gamma} \cdot \sum_{e=1}^{\gamma} t_e^*(X) \tag{2}$$

where e represents the individual bootstrap sample, X represents the input vector for the test, t_e^* represents the individual decision tree, γ is the number of trees, and P_{RF} is the merged result based on RF.

In this study, we constructed regression trees based on the classification and regression tree algorithm, using the mean square error as the basis of node splitting. γ was determined from {25, 50, 100, 200}. The RF model generally performed better as the number of trees increased. However, when $\gamma = 200$, it took more time to train the model, and the RMSE of the merged result did not decrease. Therefore, the number of regression trees was set to 100. The minimum number of samples required for a leaf node, n_{min}, was determined from {1, 2, 3, 4, 5, 6, 7, 8, 9}. The RF model performed best when $n_{min} = 6$. Therefore, any split point in this study left at least 6 training samples in each of the left and right branches. Additionally, the number of features to consider when searching for the optimal split, $n_{features}$, was set as 3, since the RF model performed better with this number than when $n_{features} = 1, 2, 4$ or 5.

4) LONG SHORT-TERM MEMORY NETWORK

LSTM is an improved version of a recurrent neural network, which is typically used to process sequential data [19]. A basic LSTM module usually consists of a memory cell and 3 gates (input, forget and output gates) [42]. The gates are used to determine the amount of information that should be remembered and the cell is used to store the information [43]. The related formulations are as follows:

$$I_t = \sigma(W_I S_t + U_I H_{t-1} + \delta_I) \tag{3}$$
$$F_t = \sigma(W_F S_t + U_F H_{t-1} + \delta_F) \tag{4}$$
$$O_t = \sigma(W_O S_t + U_O H_{t-1} + \delta_O) \tag{5}$$
$$\tilde{C}_t = tanh(W_C S_t + U_C H_{t-1} + \delta_C) \tag{6}$$
$$C_t = F_t \otimes C_{t-1} + I_t \otimes \tilde{C}_t \tag{7}$$
$$H_t = O_t \otimes \tanh(C_t) \tag{8}$$

where the variables I, F, O, and C represent the input gate, forget gate, output gate, and memory cell, respectively. S and H are the input vector and the hidden vector, respectively. Subscript t denotes that a given variable refers to time, t. Similarly, $t-1$ refers to the previous time step. W is the weight from each gate or memory cell to the input and U represents the weight from each gate or memory cell to the hidden state. δ represents the bias of each gate or memory cell.

The subscripts I, F, O, and C indicate given variables corresponding to the input gate, forget gate, output gate, or memory cell, respectively. σ represents the activation function (sigmoid).

In this study, a dense (fully connected) layer is added after the LSTM layer to generate the value of the merged result using H_t. The Adam method was chosen as the optimized for this model [44], as it performed better than both the RMSprop and SGD methods. We use early stopping in our study, meaning that the training of the model was ended when the loss stopped decreasing (the models of all pixels stopped within 1000 epochs). The number of neurons in the hidden layer, n_{hidden}, was selected from {3, 4, 5, 6, 7, 8, 9, 10, 50, 100}. n_{hidden} was set to 4, as the LSTM model performed best with this number. Similarly, the learning rate and batch size were set to 0.001 and 30, respectively.

C. ACCURACY VERIFICATION METHODS

The fusion results were evaluated from two aspects: error statistics and detection capability analysis. Error statistics reflect the error between merged results and interpolated data and include the correlation coefficient (CC), root mean square error (RMSE), and relative bias (RB). CC reflects the synchronicity of precipitation variation between the evaluated data and gauge data; RMSE indicates the averaged error magnitude of the evaluated data and gauge data, and RB reflects the probability of overestimation (RB $<$ 0) or

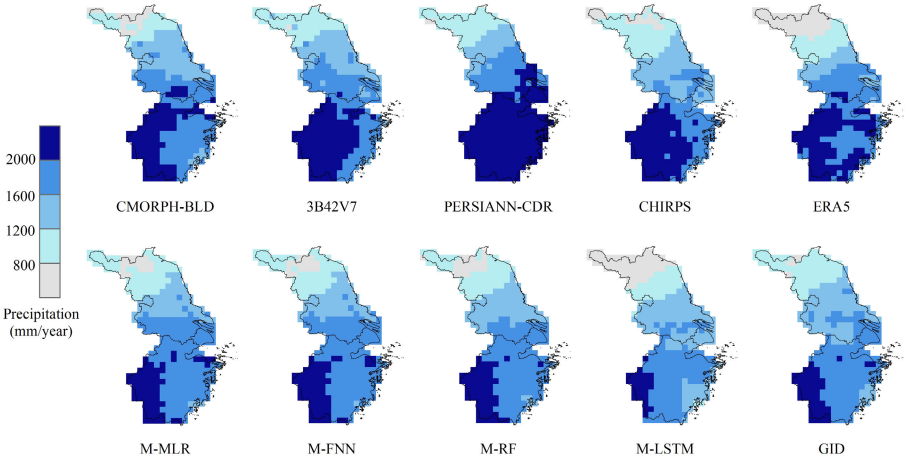


FIGURE 2. The spatial distribution of five precipitation products and four merged data and GID at a yearly scale.

TABLE 2. Specifications of the four merging methods.

Approach	Specifications
MLR	least squares method
FNN	Optimizer: L-BFGS
	Number of neurons in the hidden layer: 11
	Iterations: 2000
	Learning rate: 0.05
RF	Number of trees: 100
	Minimum number of samples required for a leaf node: 6
	Number of features to consider when searching the optimal split: 3
LSTM	Number of neurons in the hidden layer: 4
	Optimizer: Adam
	Epochs: 1000
	Learning rate: 0.001
	Batch size: 30

TABLE 3. Accuracy verification formulas.

Indexes	Calculation Formula	Unit
CC	$\mathrm{CC}=\frac{\sum_{i=1}^{n}(S_i-\bar{S})(G_i-\bar{G})}{\sqrt{\sum_{i=1}^{n}(S_i-\bar{S})^2}\sqrt{\sum_{i=1}^{n}(G_i-\bar{G})^2}}$	/
RMSE	$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i-G_i)^2}$	mm
RB	$\mathrm{RB}=\sum_{i=1}^{n}(S_i-G_i)/\sum_{i=1}^{n}G_i$	%
POD	$\mathrm{POD}=\frac{H}{H+M}$	/
FAR	$\mathrm{FAR}=\frac{F}{H+F}$	/
CSI	$\mathrm{CSI}=\frac{H}{H+M+F}$	/

underestimation (RB $>$ 0) from the evaluated data. Detection capability analysis, which is used to evaluate the monitoring capability of merged results for precipitation events, includes the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). POD describes the probability that precipitation events are detected correctly by the evaluated data. FAR describes the probability that the precipitation events are detected by the evaluated data but not detected by gauge data. CSI describes the total capability of the evaluated data relative to gauge data [45]. The equations for these indicators are described in Table 3.

In Table 3, n indicates the number of samples and G_i indicates the values of GID, S_i indicates the values of the evaluated data; $\bar{G}$ and $\bar{S}$ indicate the average values of GID and the evaluated data, respectively. H indicates the number of rainfall events detected both by gauges and the evaluated data, with M indicating the number just detected by gauges, and F indicating the number detected by the evaluated data.

IV. RESULTS

A. SPATIAL DISTRIBUTION OF ANNUAL PRECIPITATION

The four sets of merged data generated in this study are based on MLR (M-MLR), FNN (M-FNN), RF (M-RF), and LSTM (M-LSTM), respectively. The spatial distribution of annual precipitation from GID, the precipitation products, and the merged data are shown in Fig. 2. The annual precipitation of GID in the southwest was the highest. In general, annual precipitation increased gradually from north to south. The spatial distribution characteristics generated from the precipitation products and merged data are all similar to those from the GID. However, all five precipitation products clearly overestimated the precipitation in the southern regions of the study area. M-LSTM underestimated precipitation in the north, east, southeast and southwest regions of the study area. GID showed only 5 pixels where annual precipitation was less than 800 mm/year, where there were 57 pixels placed in this category by M-LSTM. According to GID, there were 46 pixels that exhibited annual precipitations greater than 2000 mm/year, whereas M-LSTM showed only 26 pixels.

B. ERROR ANALYSIS OF DAILY PRECIPITATION

Fig. 3 illustrates the relationships between the GID and the studied precipitation products and merged datasets. In general, among all gridded precipitation products before fusion, CMORPH-BLD has the highest accuracy, followed by ERA5.

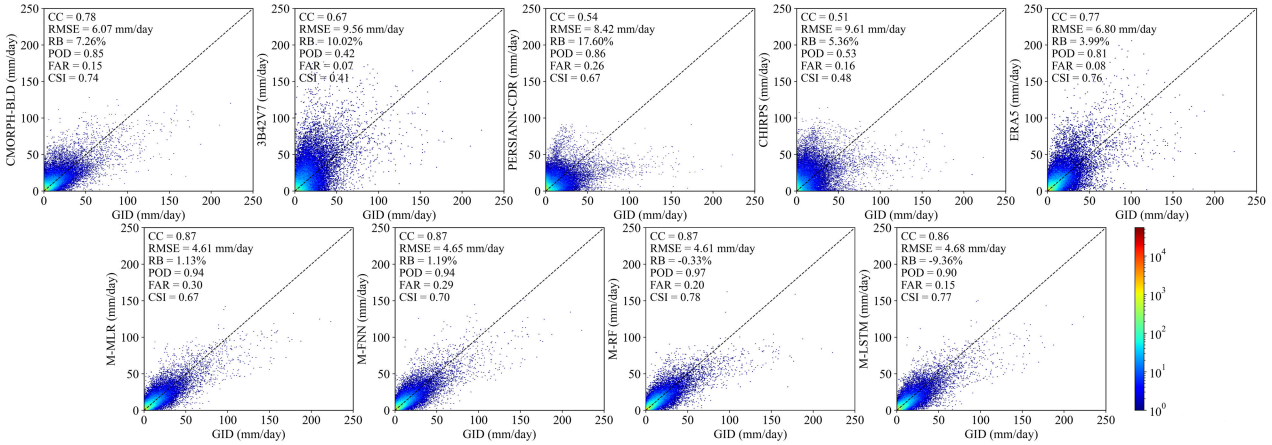


FIGURE 3. The scatter plots of five precipitation products and four merged data.

However, each of the merged datasets was more accurate than any of the gridded precipitation products. The merged datasets all had greater CCs and smaller RMSEs than the five precipitation products, showing that these merging approaches can reduce errors between precipitation estimations and GID. Among the four merged datasets, M-RF performed best (CC = 0.87, RMSE = 4.61 mm/day, RB = − 0.33%), with a CC 11% higher than the best gridded precipitation product (CMORPH-BLD) and an RMSE 24% lower. The other merged datasets had similar CCs and RMSEs. However, M-LSTM not only performed worse in terms of RB when compared to other merged datasets, but also compared to the CHIRPS and ERA5 products. It is worth mentioning that the gridded precipitation products before fusion all overestimated the actual precipitation, whereas M-RF and M-LSTM underestimated the actual precipitation.

According to [46], the rainy season over this study area is from March to September, while the dry seasons are from January to February and October to December. The average rainfall in the rainy season (5.85 mm/day) is higher than in the dry season (2.88 mm/day). Fig. 4 shows the accuracy of the five precipitation products and four merged datasets at different time periods. Regardless of the period, the CCs and RMSEs of the merged data were always better than those of the precipitation products. The first quartiles, median values, and third quartiles of merged data's CCs were all larger, whereas the first quartiles, median values, and third quartiles of the merged data's RMSEs were smaller. The CCs of the four merged datasets were better in the dry season. During this time, their median values (0.89, 0.89, 0.90, 0.90) were higher and their ranges (0.36, 0.37, 0.34, 0.40) were smaller than those in the rainy season. The RMSEs of the four merged datasets were also better in the dry season, with median values (2.72 mm/day, 2.73 mm/day, 2.65 mm/day, 2.76 mm/day) and ranges (6.42 mm/day, 6.26 mm/day, 7.63 mm/day, 6.44 mm/day) that were smaller than those in the rainy season.

However, the RBs were actually better in the rainy season, exhibiting median values (2.82%, 2.71%, − 0.30%, − 4.56%)

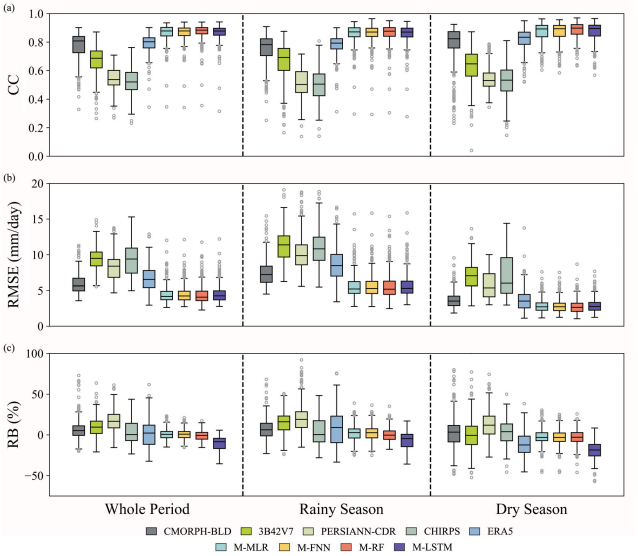


FIGURE 4. Box charts of three error statistical indexes of the precipitation products and merged data (Circles represent outliers and the length of each boxes represents the interquartile range).

that were closer to 0% and ranges (63.78%, 61.61%, 52.85%, 52.91%) that were smaller than those in the dry season.

The spatial distribution of errors for the studied approaches (Fig. 5, Fig. 6 and Fig. 7) showed that, in general, M-RF performed best. Additionally, the merged datasets clearly performed better than the precipitation products in terms of CC and RMSE. CMORPH-BLD, 3B42V7, and ERA5 only had 193, 21, and 197 pixels with CC larger than 0.80, respectively. PERSIANN-CDR and CHIRPS did not have any. Conversely, 90% of the pixels in the four merged datasets had a CC larger than 0.80. In terms of RMSE, CMORPH-BLD, PERSIANN-CDR, CHIRPS, and ERA5 only had 101, 16, 1, and 69 pixels smaller than 5.00 mm/day, respectively. 3B42V7 did not have any. The merged datasets, however, had more than 75% of pixels with an RMSE smaller than 5.00 mm/day. There were no obvious trends in the spatial distribution of CCs or RMSEs in any of the tested methods. The results regarding the relative

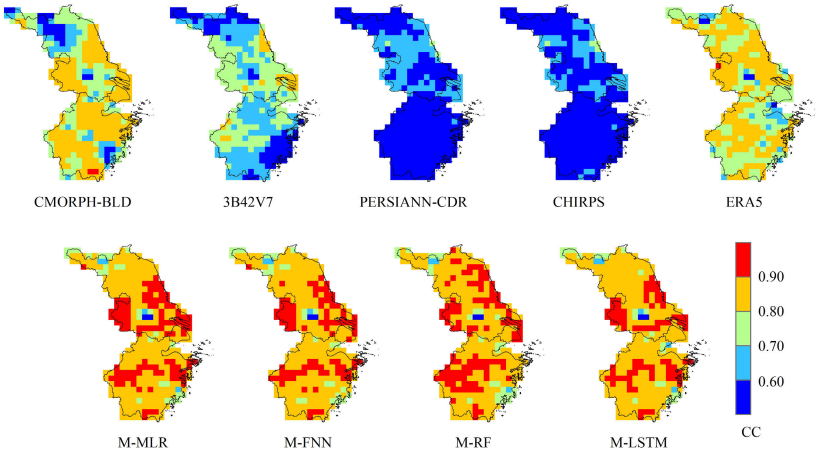


FIGURE 5. Spatial distribution of CCs for the five precipitation products and four merged data at a daily scale.

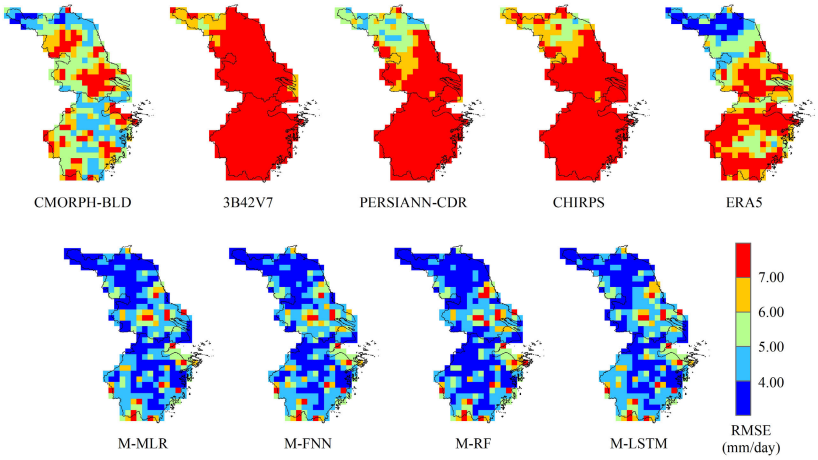


FIGURE 6. Spatial distribution of RMSEs for the five precipitation products and four merged data at a daily scale.

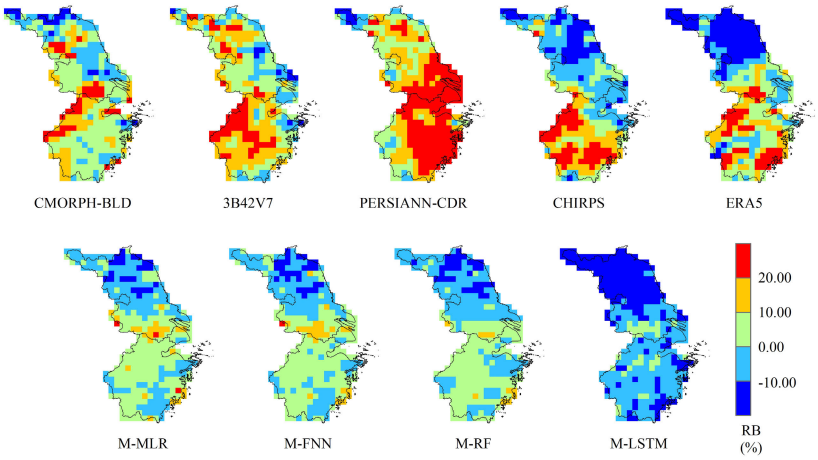


FIGURE 7. Spatial distribution of RBs for the five precipitation products and four merged data at a daily scale.

bias of the estimations, however, were more varied. Values for CHIRPS, ERA5, and the four merged datasets in the south were generally larger than those in the north. M-MLR, M-FNN and M-RF all performed significantly better than the

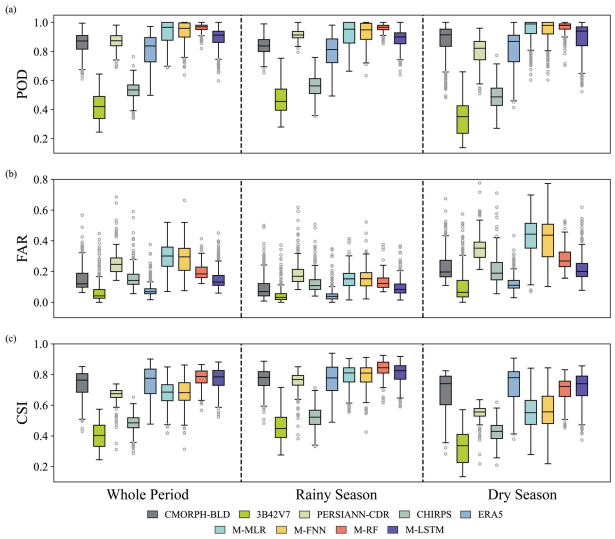


FIGURE 8. The box chart of the daily precipitation event detection capacity of the precipitation products and the merged data.

precipitation products, with more than 85% of pixels having RBs between − 10% and 10%. Only 66% of pixels from the five precipitation products fell into this range and only 56.1% from M-LSTM.

C. ANALYSIS OF DAILY PRECIPITATION EVENT DETECTION CAPACITY

Compared with the five gridded precipitation products, M-RF and M-LSTM had much greater daily precipitation event detection capacity. However, the CSIs of M-MLR and M-FNN were lower than those of CMORPH-BLD and ERA5. Furthermore, the FARs of the four merged datasets were all greater than 3B42V7 and ERA5, signifying that there was a greater probability that the merged data would misjudge precipitation events. Even the smallest FAR (0.15) among the merged data was 0.08 greater than the FAR of 3B42V7. However, each of the merged data's POD was greater than any of the precipitation products, signaling that merged datasets could correctly detect more daily precipitation events. Even the smallest total POD (0.90) among the four merged datasets was 0.12 greater than that of CMORPH-BLD, which had the largest POD among the precipitation products.

As shown in Fig. 8, the daily precipitation event detection capacity of each pixel was analyzed across different time periods. In general, M-RF performed best, too. In the rainy season, the total daily precipitation event detection capacity of the merged data was better than that of the precipitation products. However, ERA5 performed better than the merged data in the dry season. The PODs of the four merged datasets were better in the dry season, with median values (0.99, 0.98, 0.98, 0.94) that were greater than those in the rainy season. However, the POD ranges (0.40, 0.40, 0.30, 0.48) in dry season were larger than those in the rainy season. Conversely, the FARs of the merged datasets were better in the rainy season, where they had median values (0.15, 0.15, 0.12, 0.08) and ranges (0.40, 0.50, 0.31, 0.35) that were smaller than those in dry season. The CSIs were also better in the rainy season. The median CSI values (0.81, 0.81, 0.84, 0.83) in the rainy season were greater than those in the dry season, whereas the ranges (0.35, 0.49, 0.31, 0.33) were all smaller.

Fig. 9, Fig. 10 and Fig. 11 showed the spatial differences of the daily precipitation event detection capacity of the merged data. Using these metrics, M-RF again showed the best performance of all tested approaches. M-RF and M-LSTM clearly had improved CSI values compared to the precipitation products. CMORPH-BLD, PERSIANN-CDR, and ERA5 only had 71.00%, 19.24% and 69.65% pixels with CSIs greater than 0.80. 3B42V7 and CHIRPS did not have a single pixel with a CSI over 0.80. However, more than 80% of pixels from M-RF and M-LSTM had a CSI greater than 0.70. CSI values from CMORPH-BLD, 3B42V7, CHIRPS, ERA5, and the four merged datasets all increased from north to south, signifying that the pixels in the south had better daily precipitation event detection capacity than those in the north.

The four merged datasets also had significantly better PODs than the precipitation products. CMORPH-BLD, PERSIANN-CDR and ERA5 only had 31.98%, 32.79% and 21.41% pixels, respectively, that had a POD greater than 0.90. 3B42V7 and CHIRPS had none. The four merged datasets, however, had more than 57% of pixels with a POD greater than 0.90. The values of CMORPH-BLD, PERSIANN-CDR, ERA5 M-MLR, M-FNN and M-LSTM all increased from north to south. M-RF had the smallest spatial difference of PODs, with range of just 0.18. FAR values of CMORPH-BLD, PERSIANN-CDR, CHIRPS, and the four merged data all decreased from north to south. Additionally, the FAR of the four merged datasets was clearly worse than 3B42V7 and ERA5.

D. ACCURACY EVALUATION IN DIFFERENT SCENARIOS

According to [23], daily precipitation events can be divided into 3 categories according to precipitation intensities: light precipitation (0.1–1 mm/day), moderate precipitation (1–50 mm/day) and heavy precipitation (≥ 50 mm/day). We classified our data according to these categories and revaluated the accuracy of each approach. The error statistics are shown in Table 4. In the light precipitation scenario, the CCs of the five precipitation products and four merged datasets were all small. Among the five precipitation products, the accuracy of ERA5 (CC = 0.15, RMSE = 1.74 mm/day, RB = 56.58%) was the best. However, the CCs of the four merged datasets were all better than any of the non-fused products, but they overestimated the amount of precipitation more seriously than ERA5. M-RF and M-LSTM were the only two approaches that exhibited smaller RMSEs than ERA5. In the moderate precipitation scenario, the accuracy of CMORPH-BLD (CC = 0.66, RMSE = 7.76 mm/day, RB = 0.93%) was the best among the five precipitation products. The accuracy of M-RF (CC = 0.79, RMSE = 5.80 mm/day, RB = − 1.60%) was the best among the four merged datasets, with a CC that increased by 19.70% and an RMSE that decreased by 25.26% compared with

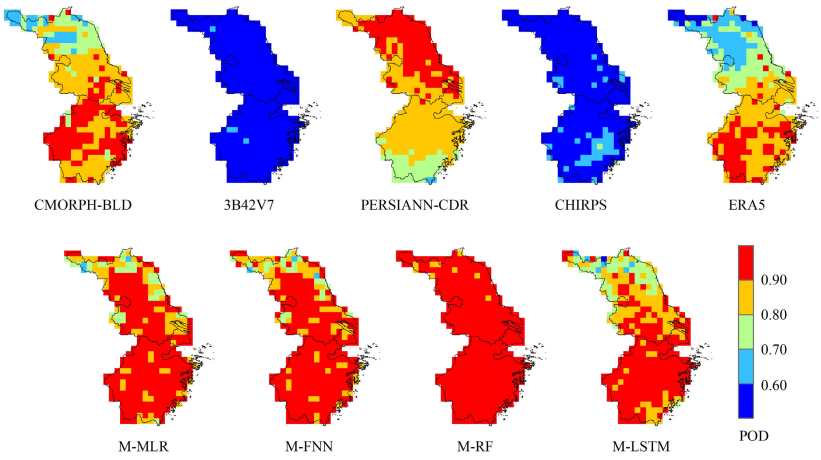


FIGURE 9. Spatial distribution of PODs for the five precipitation products and four merged data.

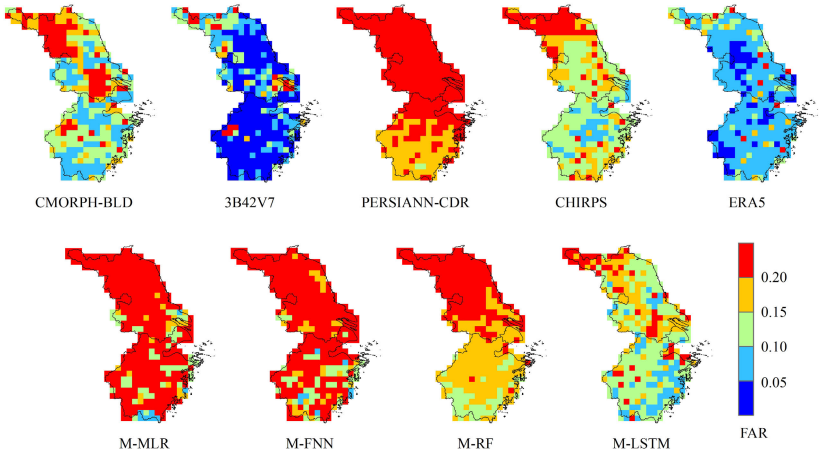


FIGURE 10. Spatial distribution of FARs for the five precipitation products and four merged data.

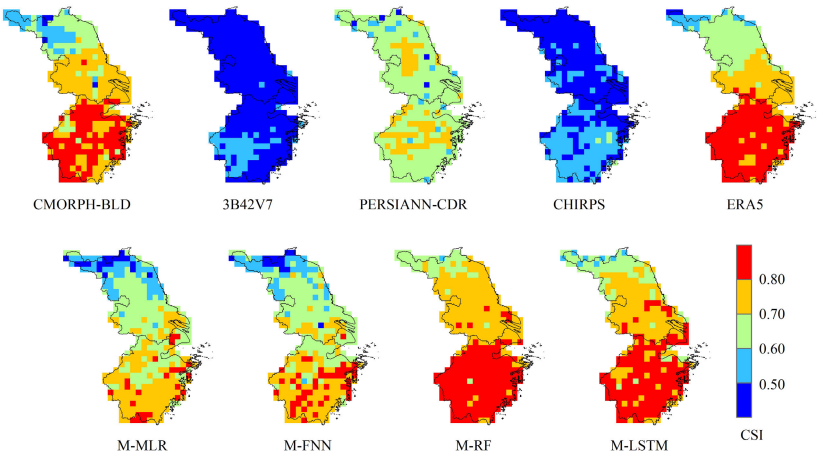


FIGURE 11. Spatial distribution of CSIs for the five precipitation products and four merged data.

CMORPH-BLD. All nine types of precipitation data studied slightly overestimated the actual precipitation. In the heavy precipitation scenario, CMORPH-BLD (CC = 0.48, RMSE = 33.92 mm/day, RB = 56.58%) was again the most

TABLE 4. Error statistics of the five gridded precipitation products and four merged data in different scenarios.

Data	Light Precipitation			Moderate Precipitation			Heavy Precipitation		
	CC	RMSE (mm/day)	RB (%)	CC	RMSE (mm/day)	RB (%)	CC	RMSE (mm/day)	RB (%)
CMORPH-BLD	0.13	3.48	283.75	0.66	7.76	0.93	0.48	33.92	-31.53
3B42V7	0.07	3.56	101.31	0.56	13.51	8.08	0.26	38.35	-11.83
PERSIANN-CDR	0.13	5.11	652.69	0.39	10.32	-1.43	0.24	50.77	-59.71
CHIRPS	0.08	5.26	395.52	0.39	12.57	-3.79	0.19	49.13	-53.83
ERA5	0.15	1.74	56.58	0.67	9.09	4.71	0.31	39.66	-17.63
M-MLR	0.16	1.84	167.72	0.78	6.00	-3.64	0.60	27.48	-23.95
M-FNN	0.18	1.82	161.77	0.78	6.03	-3.08	0.58	28.34	-23.83
M-RF	0.23	1.43	131.04	0.79	5.80	-1.60	0.54	31.51	-31.11
M-LSTM	0.19	1.37	81.35	0.77	6.08	-10.77	0.59	29.64	-27.45

TABLE 5. Error statistics and daily precipitation event detection capacity of merged data.

Data	CC	RMSE (mm/day)	RB (%)
N-MLR	0.80	6.75	-13.80
N-FNN	0.77	6.90	-12.58
N-RF	0.74	7.00	-10.84
N-LSTM	0.77	7.12	-42.00

accurate of the five precipitation products. Of the merged datasets, M-MLR (CC = 0.60, RMSE = 27.48 mm/day, RB = − 23.95%) was the most accurate, with a CC that increased by 20.83% compared to CMORPH-BLD. However, all of the nine studied types of precipitation data underestimated the actual amount of precipitation.

V. DISCUSSION

A. ROBUSTNESS TEST

We conducted a robustness check to evaluate the four merging methods used. Goodfellow *et al.* [47] applied some perturbations to inputs of neural networks and successfully made neural networks output incorrect answers. Therefore, for each grid point, we applied perturbations to the validation data of one randomly selected day in 2015. The perturbations induced errors in one of the gridded precipitation products randomly (the error follows the standard normal distribution). Then, we used MLR, FNN, RF and LSTM to merge the data in case of error and generate four datasets: error data based on MLR (E-MLR), FNN (E-FNN), RF (E-RF) and LSTM (E-LSTM). The test was repeated ten times. The RMSEs between these data and the GID were calculated (see Fig. 12). In the robustness test, the mean value of the RMSEs of E-LSTM (8.71 mm/day) was the smallest, followed by that of E-RF (9.00 mm/day) and E-MLR (9.09 mm/day). The RMSE of the merged data generated by the error data based on the FNN (9.34 mm/day) was the greatest. Compared with M-MLR, M-FNN, M-RF and M-LSTM, the RMSEs of the corresponding data in the robustness test all became larger. Moreover, the RMSE of E-LSTM only increased 86.11%, which was smaller than that of E-MLR (97.18%), E-FNN (100.86%) and E-RF (95.23%). This result illustrates that the robustness of LSTM is better than that of MLR, FNN and RF when merging multiple precipitation data.

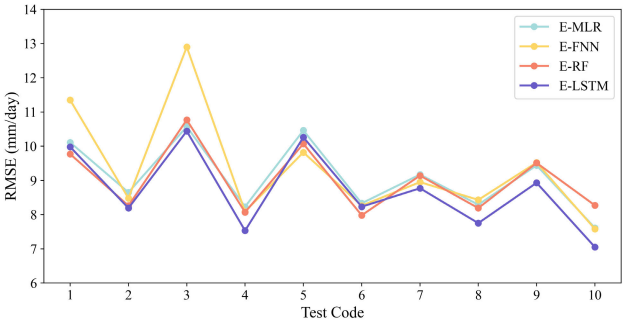


FIGURE 12. RMSEs of the data fused by error data and GID.

B. MERGING EXPERIMENT WITH SPARSE GAUGES

The accuracy of gridded precipitation data is affected by the density of the rain gauges [48], [49]. However, in some areas, the distribution of gauges is sparse, and thus the accuracy of the interpolated data in these areas is relatively low. To test the fusion effect of the four fusion methods used in this study in areas with sparse gauges, an experiment was run using only twelve international exchange stations for interpolation. New interpolated data with a spatial resolution of 0.25° × 0.25° was generated, based on the inverse distance weight interpolation method. These new interpolated data were first compared with GID data. The CC between them was 0.47, while the RMSE and RB were 7.79 mm/day and − 2.69%, respectively. Then, the new interpolated data were used as the output vectors of the training set to conduct a new merging experiment, and new merged data were generated. The new merged datasets were based on MLR (N-MLR), FNN (N-FNN), RF (N-RF), and LSTM (N-LSTM).

These merged data were also compared with GID (Table 5). Among these new merged datasets (N-MLR, N-FNN, N-RF and N-LSTM), N-MLR performs the best, containing the smallest RMSE and largest CC. Compared with the previous merged data (M-MLR, M-FNN, M-RF and M-LSTM), the CCs of the new merged data were all smaller and the RMSEs were all greater. Moreover, the previous merged

data performed better in RB. In terms of CC and RMSE, the differences between the previous data and new data based on MLR were the smallest. Compared with M-MLR, N-MLR's CC was only reduced by 0.07. Compared with the RMSE of M-MLR, the RMSE of N-MLR only increased by 2.14 mm/day. For RB, the difference between the previous data and new data based on RF was smallest among the four merging methods.

VI. CONCLUSION

Based on four main data fusion approaches, fusion experiments of four satellite precipitation products (CMORPH-BLD, 3B42V7, PERSIANN-CDR, and CHIRPS) and one reanalysis (ERA5) were performed. Subsequently, the quality of the different merged results was evaluated from multiple perspectives. The conclusions are as follows:

(1) All merging approaches can improve the accuracy of precipitation data, but only RF and LSTM can improve the daily precipitation event detection capacity. These approaches primarily reduced errors in scenarios with moderate precipitation, but did not effectively improve the accuracy when there was light or heavy precipitation.

(2) MLR had the cheapest computing cost in our study. Moreover, the accuracy of the merged data generated by MLR was even better than that generated by FNN and LSTM. Additionally, MLR performed better in situations with low gauge density than the other approaches. However, the daily precipitation event detection capacity of data generated by MLR was the worst of the four merged datasets.

(3) FNN performed moderately throughout the experiments.

(4) The merged data generated by RF had the best accuracy and daily precipitation event detection capacity. RF performed best in the moderate precipitation scenario. However, of the merging methods, RF had the worst performances in situations with low gauge density.

(5) LSTM performed best in the light precipitation scenario. The FAR of the data generated by LSTM was the smallest among the four merged datasets. In addition, LSTM was the most robust method in our study. However, compared to the other merging methods, LSTM also had the highest computing cost, and lowest accuracy of the merged data.

(6) In this study, FNN and LSTM often performed worse than MLR. However, one study [50] showed that RF performed better than geographically weighted regression (GWR) with more than 30,000 samples, but performed worse when sample size was only a few thousand or less. Therefore, we infer that the data size in our study might have been too small since the performance of the approaches of machine learning and deep learning can be affected by data size [51].

REFERENCES

[1] M. Pedro-Monzonís, A. Solera, J. Ferrer, T. Estrela, and J. Paredes-Arquiola, "A review of water scarcity and drought indexes in water resources planning and management," *J. Hydrol.*, vol. 527, pp. 482–493, Aug. 2015, doi: 10.1016/j.jhydrol.2015.05.003.

[2] W. Li, C. Liu, Y. Hong, M. Saharia, W. Sun, D. Yao, and W. Chen, "Rainstorm-induced shallow landslides process and evaluation—A case study from three hot spots, China," *Geomatics, Natural Hazards Risk*, vol. 7, no. 6, pp. 1908–1918, Nov. 2016, doi: 10.1080/19475705.2016.1179685.

[3] C. Kidd, A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum, "So, how much of the Earth's surface is covered by rain gauges?" *Bull. Amer. Meteorolog. Soc.*, vol. 98, no. 1, pp. 69–78, Jan. 2017, doi: 10.1175/BAMS-D-14-00283.1.

[4] F. M. Woldemeskel, B. Sivakumar, and A. Sharma, "Merging gauge and satellite rainfall with specification of associated uncertainty across australia," *J. Hydrol.*, vol. 499, pp. 167–176, Aug. 2013, doi: 10.1016/j.jhydrol.2013.06.039.

[5] E. E. Ebert, J. E. Janowiak, and C. Kidd, "Comparison of near-real-time precipitation estimates from satellite observations and numerical models," *Bull. Amer. Meteorolog. Soc.*, vol. 88, no. 1, pp. 47–64, Jan. 2007, doi: 10.1175/BAMS-88-1-47.

[6] Y. Li, Q. J. Wang, H. He, Z. Wu, and G. Lu, "A method to extend temporal coverage of high quality precipitation datasets by calibrating reanalysis estimates," *J. Hydrol.*, vol. 581, Feb. 2020, Art. no. 124355, doi: 10.1016/j.jhydrol.2019.124355.

[7] E. Kalnay and M. Cai, "Impact of urbanization and land-use change on climate," *Nature*, vol. 423, no. 6939, pp. 528–531, May 2003, doi: 10.1038/nature01675.

[8] C. Li, G. Tang, and Y. Hong, "Cross-evaluation of ground-based, multi-satellite and reanalysis precipitation products: Applicability of the triple collocation method across Mainland China," *J. Hydrol.*, vol. 562, pp. 71–83, Jul. 2018, doi: 10.1016/j.jhydrol.2018.04.039.

[9] Y. Shen, A. Xiong, Y. Wang, and P. Xie, "Performance of high-resolution satellite precipitation products over China," *J. Geophys. Res.*, vol. 115, no. D2, 2010, Art. no. D02114, doi: 10.1029/2009JD012097.

[10] A. K. Betts, M. Zhao, P. A. Dirmeyer, and A. C. M. Beljaars, "Comparison of ERA40 and NCEP/DOE near-surface data sets with other ISLSCP-II data sets," *J. Geophys. Res.*, vol. 111, no. D22, 2006, pp. 1–60, doi: 10.1029/2006JD007174.

[11] M. G. Bosilovich, J. Chen, F. R. Robertson, and R. F. Adler, "Evaluation of global precipitation in reanalyses," *J. Appl. Meteorol. Climatol.*, vol. 47, no. 9, pp. 2279–2299, Sep. 2008, doi: 10.1175/2008JAMC1921.1.

[12] P. Xie and P. A. Arkin, "Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions," *J. Climate*, vol. 9, no. 4, pp. 840–858, Apr. 1996, doi: 10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO.2.

[13] A. Verdin, C. Funk, B. Rajagopalan, and W. Kleiber, "Kriging and local polynomial methods for blending satellite-derived and gauge precipitation estimates to support hydrologic early warning systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2552–2562, May 2016, doi: 10.1109/TGRS.2015.2502956.

[14] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, Feb. 15, 2017, doi: 10.1109/JLT.2016.2590989.

[15] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111716, doi: 10.1016/j.rse.2020.111716.

[16] A. C. Turlapaty, V. G. Anantharaj, N. H. Younan, and F. Joseph Turk, "Precipitation data fusion using vector space transformation and artificial neural networks," *Pattern Recognit. Lett.*, vol. 31, no. 10, pp. 1184–1200, Jul. 2010, doi: 10.1016/j.patrec.2009.12.033.

[17] O. M. Baez-Villanueva, M. Zambrano-Bigiarini, H. E. Beck, I. McNamara, L. Ribbe, A. Nauditt, C. Birkel, K. Verbist, J. D. Giraldo-Osorio, and N. X. Thinh, "RF-MEP: A novel random forest method for merging gridded precipitation products and ground-based measurements," *Remote Sens. Environ.*, vol. 239, Mar. 2020, Art. no. 111606, doi: 10.1016/j.rse.2019.111606.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[19] H. Wu, Q. Yang, J. Liu, and G. Wang, "A spatiotemporal deep fusion model for merging satellite and gauge precipitation in China," *J. Hydrol.*, vol. 584, May 2020, Art. no. 124664, doi: 10.1016/j.jhydrol.2020.124664.

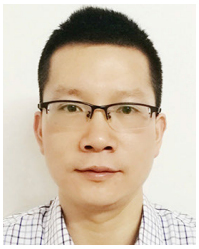
[20] I. V. Sideris, M. Gabella, R. Erdin, and U. Germann, "Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of switzerland," *Quart. J. Roy. Meteorolog. Soc.*, vol. 140, no. 680, pp. 1097–1111, Apr. 2014, doi: 10.1002/qj.2188.

[21] R. Hénin, M. Liberato, A. Ramos, and C. Gouveia, "Assessing the use of satellite-based estimates and high-resolution precipitation datasets for the study of extreme precipitation events over the iberian peninsula," *Water*, vol. 10, no. 11, p. 1688, Nov. 2018, doi: 10.3390/w10111688.
[22] E. Sharifi, J. Eitzinger, and W. Dorigo, "Performance of the state-of-the-art gridded precipitation products over mountainous terrain: A regional study over Austria," *Remote Sens.*, vol. 11, no. 17, p. 2018, Aug. 2019, doi: 10.3390/rs11172018.
[23] Q. Jiang, W. Li, Z. Fan, X. He, W. Sun, S. Chen, J. Wen, J. Gao, and J. Wang, "Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland," *J. Hydrol.*, Oct. 2020, Art. no. 125660, doi: 10.1016/j.jhydrol.2020.125660.
[24] H. Ashouri, K.-L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat, "PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies," *Bull. Amer. Meteorolog. Soc.*, vol. 96, no. 1, pp. 69–83, Jan. 2015, doi: 10.1175/BAMS-D-13-00068.1.
[25] G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker, "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales," *J. Hydrometeorol.*, vol. 8, no. 1, pp. 38–55, Feb. 2007, doi: 10.1175/JHM560.1.
[26] R. Joyce, J. Janowiak, P. Arkin, and P. Xie, "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution," *J. Hydrometeorol.*, vol. 5, no. 3, pp. 487–503, Jun. 2004, doi: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO.2.
[27] C. Funk, P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, and J. Michaelsen, "The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes," *Sci. Data*, vol. 2, no. 1, pp. 1–21, Dec. 2015, doi: 10.1038/sdata.2015.66.
[28] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, and A. Simmons, "The ERA5 global reanalysis," *Quart. J. Roy. Meteorolog. Soc.*, vol. 146, no. 730, pp. 1999–2049, Jul. 2020, doi: 10.1002/qj.3803.
[29] C. Miao, H. Ashouri, K.-L. Hsu, S. Sorooshian, and Q. Duan, "Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China," *J. Hydrometeorology*, vol. 16, no. 3, pp. 1387–1396, Jun. 2015, doi: 10.1175/JHM-D-14-0174.1.
[30] Z. Duan, J. Liu, Y. Tuo, G. Chiogna, and M. Disse, "Evaluation of eight high spatial resolution gridded precipitation products in adige basin (Italy) at multiple temporal and spatial scales," *Sci. Total Environ.*, vol. 573, pp. 1536–1553, Dec. 2016, doi: 10.1016/j.scitotenv.2016.08.213.
[31] T. Chen, L. Ren, F. Yuan, T. Tang, X. Yang, S. Jiang, Y. Liu, C. Zhao, and L. Zhang, "Merging ground and satellite-based precipitation data sets for improved hydrological simulations in the xijiang river basin of China," *Stochastic Environ. Res. Risk Assessment*, vol. 33, no. 10, pp. 1893–1905, Oct. 2019, doi: 10.1007/s00477-019-01731-w.
[32] V. K. Ojha, A. Abraham, and V. Snášel, "Metaheuristic design of feedforward neural networks: A review of two decades of research," *Eng. Appl. Artif. Intell.*, vol. 60, pp. 97–116, Apr. 2017, doi: 10.1016/j.engappai.2017.01.013.
[33] W. Sun, J. Tang, and C. Bai, "Evaluation of university project based on partial least squares and dynamic back propagation neural network group," *IEEE Access*, vol. 7, pp. 69494–69503, 2019, doi: 10.1109/ACCESS.2019.2919135.
[34] H. Xue and H. Cui, "Research on image restoration algorithms based on BP neural network," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 204–209, Feb. 2019, doi: 10.1016/j.jvcir.2019.01.014.
[35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[36] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.*, vol. 45, nos. 1–3, pp. 503–528, Aug. 1989, doi: 10.1007/BF01589116.
[37] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[38] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, Jun. 2016, doi: 10.1007/s11749-016-0481-7.
[39] J. Xie and M. Zhu, "Maneuver-based driving behavior classification based on random forest," *IEEE Sensors Lett.*, vol. 3, no. 11, pp. 1–4, Nov. 2019, doi: 10.1109/LSENS.2019.2945117.
[40] C. Gonzalo-Martin, M. Lillo-Saavedra, A. Garcia-Pedrero, O. Lagos, and E. Menasalvas, "Daily evapotranspiration mapping using regression random forest models," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5359–5368, Dec. 2017, doi: 10.1109/JSTARS.2017.2733958.
[41] X. He, N. W. Chaney, M. Schleiss, and J. Sheffield, "Spatial downscaling of precipitation using adaptable random forests," *Water Resour. Res.*, vol. 52, no. 10, pp. 8217–8237, Oct. 2016, doi: 10.1002/2016WR019034.
[42] X. Yuan, L. Li, Y. A. W. Shardt, Y. Wang, and C. Yang, "Deep learning with spatiotemporal attention-based LSTM for industrial soft sensor model development," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4404–4414, May 2021, doi: 10.1109/TIE.2020.2984443.
[43] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, "Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas," *J. Hydrol.*, vol. 561, pp. 918–929, Jun. 2018, doi: 10.1016/j.jhydrol.2018.04.065.
[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[45] Q. Jiang, W. Li, J. Wen, C. Qiu, W. Sun, Q. Fang, M. Xu, and J. Tan, "Accuracy evaluation of two high-resolution satellite-based rainfall products: TRMM 3B42 V7 and CMORPH in Shanghai," *Water*, vol. 10, no. 1, p. 40, Jan. 2018, doi: 10.3390/w10010040.
[46] Q. Yu, W. Wu, F. Zhou, and Q. Wang, "Climate modes of precipitation in eastern China and division of rainy season," *Trans. Atmos. Sci.*, vol. 37, no. 3, pp. 378–384, 2014, doi: 10.13878/j.cnki.dqkxxb.2014.03.012.
[47] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," 2014, *arXiv:1412.6572*. [Online]. Available: http://arxiv.org/abs/1412.6572
[48] C. Jing, M. Du, P. Dai, H. Wei, and H. Liu, "Research on accuracy assessment of urban rainfall spatial interpolation from gauges data," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 3121–3124, doi: 10.1109/IGARSS.2014.6947138.
[49] Q. Zeng, H. Chen, C.-Y. Xu, M.-X. Jie, J. Chen, S.-L. Guo, and J. Liu, "The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach," *J. Hydrol.*, vol. 563, pp. 106–122, Aug. 2018, doi: 10.1016/j.jhydrol.2018.05.058.
[50] Q. Yang, Q. Yuan, L. Yue, T. Li, H. Shen, and L. Zhang, "Mapping PM2.5 concentration at a sub-km level resolution: A dual-scale retrieval approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 165, pp. 140–151, Jul. 2020, doi: 10.1016/j.isprsjprs.2020.05.018.
[51] Y. Wu, L. Liu, C. Pu, W. Cao, S. Sahin, W. Wei, and Q. Zhang, "A comparative measurement study of deep learning as a service framework," *IEEE Trans. Services Comput.*, early access, Jul. 18, 2019, doi: 10.1109/TSC.2019.2928551.

ZEDONG FAN received the B.S. degree in geographic information science from Shanghai Normal University, Shanghai, China, in 2018. He is currently pursuing the M.S. degree with the School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai. His current research interests include machine learning and precipitation.

WEIYUE LI received the B.S. degree in geography science from Shandong Normal University, Jinan, China, in 2006, the M.S. degree in photogrammetry and remote sensing from Liaoning Technical University, Huludao, China, in 2010, and the Ph.D. degree in cartography and geography information engineering from Tongji University, Shanghai, China, in 2014. He is currently working as an Associate Professor with the Shanghai Normal University, Shanghai. His research interests include feature extraction of hyperspectral imagery and LiDAR data, and the hazard analysis of landslide.

QIN JIANG received the M.S. degree in physical geography from Shanghai Normal University, Shanghai, China, in 2019. She is currently pursuing the Ph.D. degree with the School of Geographic Sciences, East China Normal University, Shanghai. Her research interests include quantitative precipitation estimation, data fusion, and remote sensing.



WEIWEI SUN received the B.S. degree in surveying and mapping and the Ph.D. degree in cartography and geographic information engineering from Tongji University, Shanghai, China, in 2007 and 2013, respectively. From 2014 to 2016, he held a postdoctoral position with the State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, China, where he was involved in intelligent processing in hyperspectral imagery. He is currently an Associate Professor with Ningbo University, Ningbo, China, and a Visiting Scholar with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. He has published more than 50 journal articles. His research interests include hyperspectral image processing with manifold learning, anomaly detection, and target recognition of remote sensing imagery using compressive sensing.

JIAHONG WEN received the M.S. degree in geography and the Ph.D. degree in geography from East China Normal University, Shanghai, China, in 1989 and 1998, respectively. He is currently working as a Professor with the Shanghai Normal University, Shanghai. His research interests include disasters, geographic information systems, floods, risk analysis, and risk management.

JUN GAO received the B.S. degree in biology from Shanghai Normal University, Shanghai, China, in 1984, and the Ph.D. degree in ecology from East China Normal University, Shanghai, in 2000. He is currently working as a Professor with the Shanghai Normal University, Shanghai. His research interests include urban ecology, landscape ecology, sustainable science, as well as national park and ecotourism.