



available at www.sciencedirect.com



journal homepage: www.elsevier.com/locate/jhydrol



Merging multiple precipitation sources for flash flood forecasting

Yen-Ming Chiang^a, Kuo-Lin Hsu^b, Fi-John Chang^{a,*}, Yang Hong^c,
Soroosh Sorooshian^b

^a Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan, ROC

^b Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, CA, USA

^c GEST/UMBC, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Received 24 August 2006; received in revised form 7 March 2007; accepted 3 April 2007

KEYWORDS

Recurrent neural networks;
Satellite-derived precipitation;
Merged precipitation;
Bias adjustment;
Flood forecasting

Summary We investigated the effectiveness of combining gauge observations and satellite-derived precipitation on flood forecasting. Two data merging processes were proposed: the first one assumes that the individual precipitation measurement is non-bias, while the second process assumes that each precipitation source is biased and both weighting factor and bias parameters are to be calculated. Best weighting factors as well as the bias parameters were calculated by minimizing the error of hourly runoff prediction over Wu-Tu watershed in Taiwan. To simulate the hydrologic response from various sources of rainfall sequences, in our experiment, a recurrent neural network (RNN) model was used.

The results demonstrate that the merged method used in this study can efficiently combine the information from both rainfall sources to improve the accuracy of flood forecasting during typhoon periods. The contribution of satellite-based rainfall, being represented by the weighting factor, to the merging product, however, is highly related to the effectiveness of ground-based rainfall observation provided gauged. As the number of gauge observations in the basin is increased, the effectiveness of satellite-based observation to the merged rainfall is reduced. This is because the gauge measurements provide sufficient information for flood forecasting; as a result the improvements added on satellite-based rainfall are limited. This study provides a potential advantage for extending satellite-derived precipitation to those watersheds where gauge observations are limited.
© 2007 Elsevier B.V. All rights reserved.

* Corresponding author. Tel.: +886 2 23639461; fax: +886 2 23635854.
E-mail address: changfj@ntu.edu.tw (F.-J. Chang).

Introduction

Forecasting of the rainfall-runoff process is recognized as one of the most important tasks for watershed and flood management in Taiwan, especially in typhoon season. Taiwan is located on the main typhoon track in the northwestern Pacific Ocean. On average, typhoons attack the island around four times per year and usually bring heavy rainfall. Owing to the steep surface slopes of watersheds over the island, the flood hydrographs usually have large peaks with fast-rising limbs that might cause great damage. Undoubtedly, a precise flood forecasting/warning system is needed, and this requires accurate and effective rainfall information since it is the most important driven force for a hydrological model during typhoon events.

Flood forecasting remains one of the most challenging tasks in operational hydrology. In the past, it was performed by using ground-based gauge observations. However, for the remote regions where gauge measurement is limited, the runoff forecast is not reliable from the high uncertainty of input precipitation forcing data. With the availability of improved measurements, such as radar measurements, satellite estimates, and numerical weather prediction output, precipitation observation for regions where gauge observation are lacking can be improved. In this study, we evaluate the potential merge of information for gauge and satellite observations. Four ways of combination based on two rainfall sources, as gauge measurements and satellite-derived observations, are evaluated. These four testing combinations include using (1) gauge observation only, (2) satellite observation only, and the weighted merging observations from satellite and gauge observations (3) with and (4) without further consideration of bias adjustment in each product. In order to find the effective bias and weighted merging parameters, heavy rainfall events during typhoon season were selected for the model calibration and validation. Optimal parameter sets of models were selected based on the improvement of the short-term forecasting of runoff of the watershed in the calibration events.

For the hydrologic model, an artificial neural network model is applied to generate the runoff time series from the individual and combined precipitation sources. Previous work has shown that ANNs can be appropriately applied to the field of hydrology and meteorology, such as precipitation estimation/prediction (Minns and Hall, 1996; Hsu et al., 1997; Luk et al., 2000; Chiang et al., 2007), flood forecasting (Hsu et al., 1995; Campolo et al., 1999; Sajikumar and Thandaveswara, 1999; Chang and Chen, 2003; Chen et al., 2006), and prediction of water quality parameters (Maier and Dandy, 1996), etc. The attractiveness of ANNs derives mainly from data processing characteristics such as non-linearity, noise tolerance, and learning and generalization capability. Given sufficient amount and complexity of training data, ANNs are capable of learning and modeling any relationship between a series of inputs and outputs. For their abilities of simplifying calculation and enhancing adaptability, neural networks have been effectively used in a wide variety of problems that are difficult to understand and define.

Rainfall is the major forcing variables to the hydrologic process. This study investigates the impacts from different rainfall sources on flash flood forecasting. We use a dynamic

recurrent neural network (RNN) with internal time-delay feedback loops in both hidden layer and output layer was implemented to convert the input rainfall to the streamflow. Unlike the feed-forward neural networks which feed the time-lagged runoff in the network inputs, the recurrent neural networks consist of internal time-delay feedback in the network without the need of feeding the time-lagged runoffs. Compared with feed-forward networks, the RNN networks usually contain fewer parameters; the performance, however, is comparable to the feed-forward networks. The RNNs have been a popular tool applied in time series analysis (Han et al., 2004; Ho et al., 2002; Zhang and Xiao, 2000), in economics (Binner et al., 2004; Kumar et al., 2003; Saad et al., 1998), in signal processing (Atiya et al., 2005; Parlos et al., 2001; Parlos et al., 2000), as well as in hydrology (Chang et al., 2004; Chiang et al., 2004; Coulibaly et al., 2001).

The organization of this paper is as follows. The next section presents the study area and hydrologic data including streamflow and precipitation observations during typhoon periods. Description of the merged methods of precipitation products, methodology of recurrent neural network, as well as the model calibration procedures is provided in the section "Methods and calibration procedures". The section "Results" provides the comparison of the forecasted hydrologic responses and the selection of optimal merging parameters. Additional case studies of the contributions from gauge and satellite information on flood forecasting are also discussed. Finally, the conclusions are given in the section "Conclusions".

Study area and data

Study area

The study area (Wu-Tu watershed) is located in northern Taiwan within the coverage of latitude 24.96–25.20° N and longitude 121.60–121.88° E. The drainage area of the watershed is about 204 km², and the mean annual precipitation and runoff depth are 2865 and 2177 mm, respectively (Pan and Wang, 2004). The Kee-Lung River wanders through the watershed and merges into the downstream of the Tanshui River which surrounds Taipei City. Owing to the rugged topography of the watershed, flash floods arrive rapidly in the middle- and downstream and often result in serious damage during typhoon periods. For example, on September 17, 2001, typhoon Nari hit North Taiwan and brought record-breaking rainfall to flood Taipei. This event caused great damage to the Taipei MRT (Mass Rapid Transit) system with a total of 16 metro stations being completely submerged. Therefore, accurate flood forecasting in this watershed plays a key role in determining the flood risk for the city. The watershed contains one streamflow and three rain gauge stations as shown in Fig. 1, where S₁ represents the streamflow gauge station, and R₁, R₂, and R₃ represent the three rain gauge stations.

Gauge measurements

The regular typhoon season in Taiwan is from July to October, but can be extended to December as records

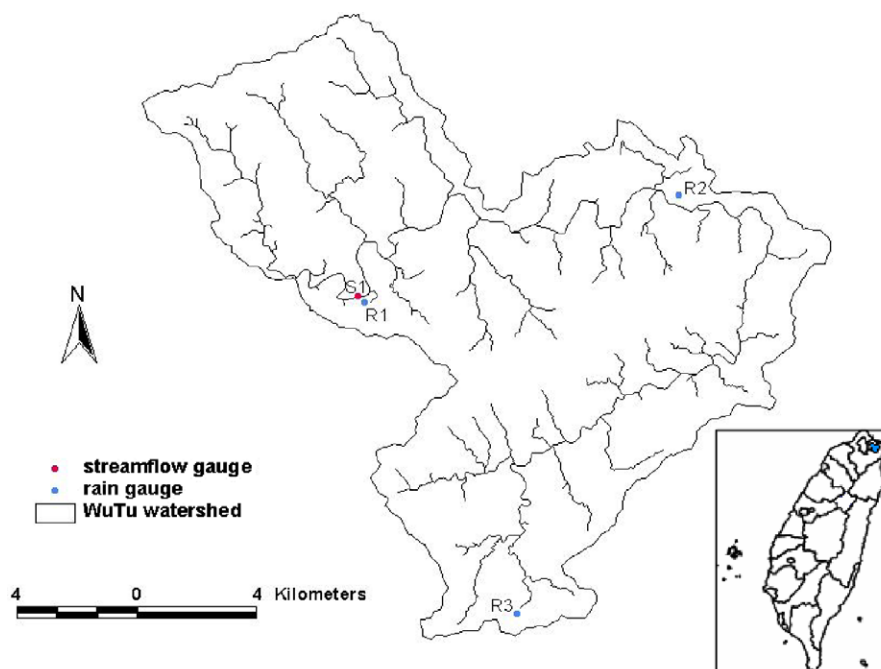


Figure 1 Location of study area and gauge stations.

have shown in recent years. As shown in Fig. 1, the outlet of the watershed is at Wu-Tu station (S_1), where hourly runoff (cm s) observations were collected from 19 storm events during 1988–2000 and in the year of 2004. For the same observational periods, hourly precipitation (mm) at three rain gauges was collected (see Fig. 1). All these events belong to the same type of precipitation considered as tropical cyclonic with heavy rainfall intensity covering around 3–4 days duration. For the hydrologic forecasting of runoff, the represented precipitation from gauge measurements is calculated the basin-averaged rainfall (P_g).

Satellite estimations

Satellite-based rainfall estimation is provided by the PERSIANN CCS (Hong et al., 2004). The PERSIANN CCS is a cloud-patch-based algorithm which generates rainfall rates based on the longwave infrared images of Geostationary Operational Environmental Satellite (GOES). The spatial and temporal resolution of the data is $0.04^\circ \times 0.04^\circ$ and hourly, respectively. Instead of direct pixel-to-pixel fitting of infrared cloud images to the rain rate, the PERSIANN CCS takes into account image segmentation and classification methods to process cloud images into a set of disjointed cloud-patch regions. Each classified cloud-patch group is specified with a multivariate non-linear temperature-rainfall curve, and the parameters of the curve are calibrated based on rainfall estimates of the Next-Generation Weather Radar (NEXRAD) network (Hong et al., 2004). In this study, the satellite-derived precipitations in 2004 were generated over the study area by using the calibrated PERSIANN CCS. Fig. 2 shows the mapping of PERSIANN CCS in 4-km grids onto the Wu-Tu watershed. Accordingly, the hourly basin-averaged rainfall (P_s) hyetographs were calcu-

lated from 15 selected data (blue grids) among these PERSIANN CCS data grids.

Merged precipitation products

This study investigates the effectiveness of merging different rainfall sources on flash flood forecasting; therefore, the gauge observations and satellite estimations are evaluated at the first stage. Then, merged precipitation products were generated and evaluated for their potential improvement of flood forecasting from the non-merged scenarios. Merging weighting factors of each precipitation source are parameterized. The parameters were assumed from the sources with bias and non-bias conditions. In the non-bias case, weighting factors are parameters to be found, while in the bias scenario, parameters including weighting factor and bias of each data source are to be optimized. The details of the merging procedures are described in next section.

Methods and calibration procedures

Precipitation data merging

Two merged precipitation products (P_{m1} and P_{m2}) were derived from the basin-averaged rainfall (P_g and P_s) data. Case studies were made to the merging model with bias or non-bias data sources. For the non-bias scenario, each rainfall source was observed from gauge or satellite unbiased. The procedure of merged rainfall products is calculated as follows:

$$P_{m1}(t) = P_g(t) \times \theta + P_s(t) \times (1 - \theta)$$

where $0 \leq \theta \leq 1$; the $P_{m1}(t)$ represents the merged precipitation without bias adjustment and the θ means the merging

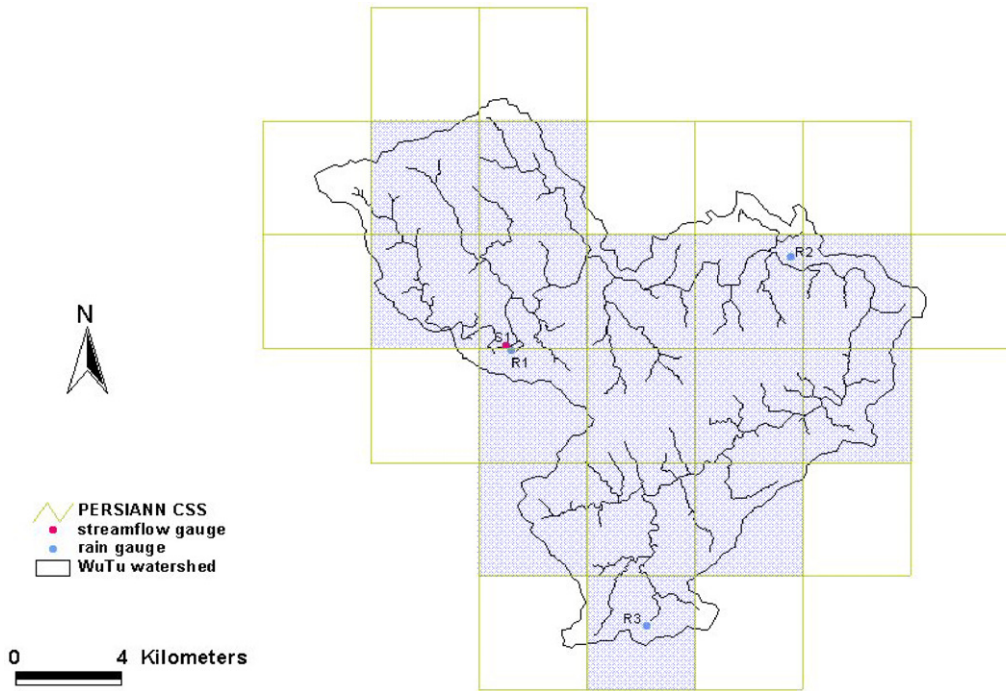


Figure 2 Mapping of the PERSIANN CCS data grids onto the study area.

weighting factor for the gauge source. The $P_g(t)$ and $P_s(t)$ represent the hourly basin-averaged rainfall calculated from gauge measurements and satellite estimates, respectively.

For bias adjustment, the bias is assumed to be proportional to the observed rainfall. The merging formula is then expressed as

$$P_{m2}(t) = (1 - \alpha_1)P_g(t) \times \theta + (1 - \alpha_2)P_s(t) \times (1 - \theta)$$

where $0 \leq \theta \leq 1$, $-1 \leq \alpha_1, \alpha_2 \leq 1$; the $P_{m2}(t)$ represents the merged precipitation with bias adjustment and the α_1 and α_2 represent the bias coefficients.

Recurrent neural network

As described in the previous sections, from merging the precipitation data to the runoff forecasting, two model modules were included: (1) a precipitation merging model and (2) a basin scale hydrologic model. The precipitation merging model is set to a linear model, while the hydrologic model is implemented by the recurrent neural network model architecture as shown in Fig. 3.

Depending on the precipitation forcing as input to the hydrologics model, four optional precipitation inputs are calculated: they are (1) gauge precipitation, (2) satellite precipitation, (3) merged gauge and satellite-based precipitation without bias adjustment and (4) merged gauge and satellite-based precipitation with bias adjustment. The hydrologic processing model, on the other hand, is represented by a three-layer RNN; the RNN consists of internal time-delay feedback loops in hidden and output layer, respectively (see Fig. 3).

Let $X_1(t)$ represent gauged precipitation ($P_g(t)$) and $X_2(t)$ represent satellite-based precipitation ($P_s(t)$). The merged

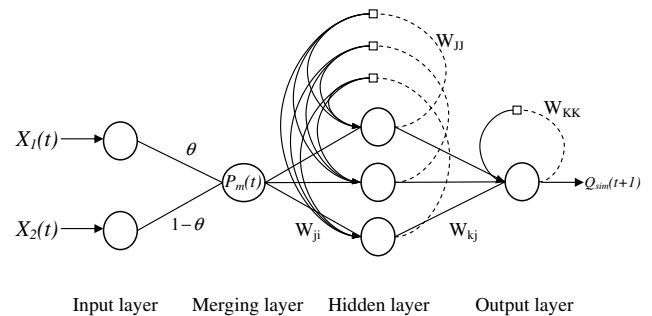


Figure 3 The architecture of a four-layer recurrent neural network.

rainfall is presented as $P_m(t)$. The merging parameters, θ , is in the range of $[0, 1]$, while θ is set to 1 for only taking gauge rainfall as its input, likewise, θ is set to 0 for only using satellite-based rainfall as its input. When merging from two separate sources, the RNN takes the combined rainfall as its input. For the non-bias condition, $X_1(t)$ represents $P_g(t)$ and $X_2(t)$ represents $P_s(t)$; 0 while for biased precipitation inputs, an additional bias adjustment parameter, α , is added. The bias is assumed to be proportional to the rainfall measurement and therefore, the input $X_1(t)$ is then assigned as $P_g(t)(1 - \alpha_1)$ and $X_2(t)$ is assigned as $P_s(t)(1 - \alpha_2)$. Hydrologic responses from rainfall sequence are generated from the recurrent neural network models. The RNN output is one-hour-ahead runoff prediction ($Q'(t + 1)$) of the watershed. The RNN model is denoted as RNN_s and RNN_g when only either gauge or satellite rainfall were used separately. Form the merged precipitation sources, the RNN models are assigned as RNN_{m1} and RNN_{m2} for non-bias and bias inputs, respectively. With various

settings of precipitation inputs provided, the performance of RNN models in flood forecasting is evaluated.

Several learning methods, such as real-time recurrent learning algorithm (Chang et al., 2002; Mandic and Chambers, 1998) and backpropagation through time algorithm (Han et al., 2004; Valdez-Castro et al., 2003) were proposed in the calibration of RNN models. Although these training algorithms have different titles, they are all gradient-based methods. Training is set to progressively update the network parameters such that will minimize the forecasting error. As shown in Fig. 3, the output of neuron J in the hidden layer is computed as

$$y_J(t) = f\left(\sum_{i=1}^I w_{Ji}P_i(t) + \sum_{j=1}^J w_{Jj}y_j(t-1)\right) \\ = f\left(P_m(t) + \sum_{j=1}^J w_{Jj}y_j(t-1)\right)$$

where $f(\cdot)$ is the neural transfer function, $y_J(t)$ is the output of the hidden unit J at time t , w_{Ji} is the connection weight from the input neuron i to the hidden layer neuron J , $P_i(t)$ is the model input (rainfall) and w_{Jj} is the time-delayed feedback weight from the hidden neuron j to the hidden neuron J . The output of neuron K in the output layer is calculated by

$$\hat{Q}_K(t+1) = f\left(\sum_{j=1}^J w_{Kj}y_j(t) + \sum_{k=1}^K w_{Kk}\hat{Q}_K(t)\right)$$

where w_{Kj} is the connection weight from hidden unit j to the output unit K , w_{Kk} is the time-delayed feedback weight from output unit k to output unit K .

The weight change for any particular weight W can be generalized by using the following formula through the chain rule with partial derivatives:

$$W_{\text{new}} = W_{\text{old}} - \eta \frac{\partial E_{\text{total}}}{\partial W}$$

where η denotes the learning rate and E_{total} represents the objective function

$$E_{\text{total}} = \sum_{t=1}^N E(t) = \frac{1}{2} \sum_{t=1}^N \sum_{k=1}^K [Q_k(t) - \hat{Q}_k(t)]^2$$

where $Q_k(t)$ is the target value of neuron k at time t ; and $\hat{Q}_k(t)$ is the network output of neuron k at time t .

Further details on the gradient methods can be found in previous studies (Ham and Kostanic, 2001; Haykin, 1999). In part of the search process, through a trial and error procedure, the number of hidden neurons was set to three. The precipitation merging parameters ($\theta, \alpha_1, \alpha_2$) were determined from the grid-based searching of the parametric domain.

Calibration procedures

In order to learn the effort of merging different source of precipitation information for flood forecasting and investigate the generalizing capability of the constructed networks, two calibration datasets with different quantity and quality of input information were adopted for separately training the recurrent neural networks (identical RNN structure). The relative performance of the con-

structed networks in flood forecasting derived from individual precipitation source would be detected through an identical validation dataset. The calibration data set #1 consists of 13 historical typhoon events selected from the period of 1988 to 2000. The RNN structures and parameters (weights) were only trained by using gauge (P_g) and streamflow data. After training, the constructed network and its calibrated weights were then directly implemented to four networks (RNN_g, RNN_s, RNN_{m1}, and RNN_{m2}), i.e. identical structure and weights, and were tested by using corresponding rainfall sequences (P_g , P_s , P_{m1} , and P_{m2}) in the validation dataset. Two typhoon events in 2004 were selected for the validation of model performance.

The second calibration experiment considered that the RNN networks were trained from the individual input sources, as for P_g , P_s , P_{m1} , and P_{m2} separately. Because only limited satellite-based rainfall data is available, we evaluated the merging parameters and calibrated the RNN models based on four typhoon events in 2004 and then applied the constructed RNN models to the validation events in 2004 (see Calibration set #2 in Table 1). Four optimal recurrent neural networks would be obtained by individual rainfall sequence and validated by the identical datasets. In sum, the constructed networks have two types of weights herein, termed as identical parameters for calibration set #1 and optimal parameters for calibration set #2. Besides, the merging parameter was varied from 0 to 1 for non-bias scenario; whereas both bias coefficients were varied from -1 to 1 for bias scenario. The merging weighting factor (θ) is stepped by 0.01 to search the optimal values. In other words, there are 101 sets of values for θ to find the best runoff calculation for non-bias scenario. For bias scenario, there are two searching phases for the determination of bias coefficients. In first phase, both bias coefficients (α_1 and α_2) are stepped by 0.1 which results in 44,541 ($21 \times 21 \times 101$) combinations. Based on the result obtained in first phase, both α_1 and α_2 are confined to 0.1 range and then stepped by 0.01 for the search of optimal values (another 44,541 combinations). In sum, there are about 8.9×10^4 combinations to be searched for bias scenario. Even though the grid-based

Table 1 Data configurations in different calibration procedures

Calibration set #1			Validation		
Data	# of events	Period	Data	# of events	Period
<i>Identical parameters</i>					
P_g	13	1988–2000	P_g	2	2004
<i>Calibration set #2</i>					
Data	# of events	Period	P_s		
<i>Optimal parameters</i>					
P_g	4	2004			
P_s			P_{m1}		
P_{m1}			P_{m2}		
P_{m2}					

searching method might be time-consuming, it provided a higher possibility to get a better solution than the simple weighted average method.

Table 1 shows the details of data configurations. The same validation dataset was used for testing their performances and investigating the efforts of two training strategies. These two training procedures are also used for analyzing the influence of amount and content of the calibration data sets. As shown in Table 1, calibration set #1 consisted of 13 events with the peak flow of 1090 (cm s). The value is merely two third of the peak flow (1570) in validation dataset. Although calibration set #2 only consisted of 4 events, the maximum value (1490) was similar as compared with that of validation.

Several statistical criteria were selected for the evaluation of precipitation estimates and flood forecasting. The quantitative accuracy of satellite precipitation estimations was assessed by relative bias (bias), and the agreement between observations and forecasted flood was evaluated based on mean absolute error (MAE), root mean square error (RMSE), correlation coefficient (CC), and coefficient of efficiency (CE). In addition, a skill score (SS) was used for evaluating the percentage improvement of model performance in target model with respect to forecasts from base model. These criteria are defined as follows:

$$\text{bias} = \frac{\sum_{i=1}^N \hat{P}(i) - \sum_{i=1}^N P(i)}{\sum_{i=1}^N P(i)} \times 100\%$$

$$\text{MAE} = \frac{\sum_{i=1}^N |\hat{Q}(i) - Q(i)|}{N}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{Q}(i) - Q(i))^2}{N}}$$

$$\text{CC} = \frac{\sum_{i=1}^N (Q(i) - \bar{Q})(\hat{Q}(i) - \bar{\hat{Q}})}{\sqrt{\sum_{i=1}^N (Q(i) - \bar{Q})^2 \sum_{i=1}^N (\hat{Q}(i) - \bar{\hat{Q}})^2}}$$

$$\text{CE} = 1 - \frac{\sum_{i=1}^N (Q(i) - \hat{Q})^2}{\sum_{i=1}^N (Q(i) - \bar{Q})^2}$$

$$\text{SS} = \left(\frac{E_{\text{base}} - E_{\text{target}}}{E_{\text{base}}} \right) \times 100\%$$

where \hat{P} is the estimated precipitation (mm/h), P is the observed precipitation (mm/h); \hat{Q} is the forecasted flood (cm s) and Q is the observed flood (cm s); \bar{Q} and $\bar{\hat{Q}}$ are the mean of observed and forecasted flood, respectively. E_{target} is the statistical error measurements in either RNN_{m1} or RNN_{m2} and E_{base} is statistical error measurements of RNN_g model herein. The SS was computed from RMSE statistics in this study. A positive SS means a better performance of the RNN_{m1} or RNN_{m2} forecasts relatively to the RNN_g model.

Results

Improvement of flood forecasting from merged rainfall

Fig. 4 shows the hourly basin-averaged rainfall from gauge and satellite-based PERSIANN CCS estimates for the six typhoon events in 2004. Each event covers around 3–4 day period of hourly measurements. Comparing these two measurements, the satellite-derived precipitation captures the trend and peak rainfalls, but slightly underestimates the light rainfalls, in particular the initial stage of the storm events. Table 2 shows the relative bias from the satellite-derived precipitation based on the gauge rainfall as its reference. Basically, the PERSIANN CCS rainfall underestimates

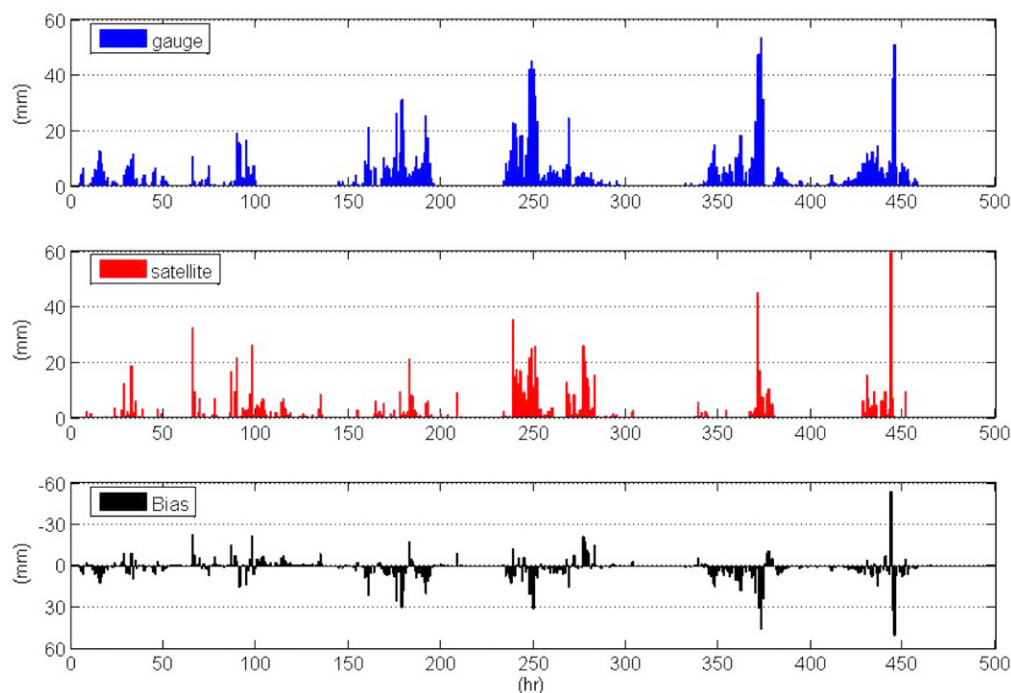


Figure 4 Basin-averaged rainfall estimated from gauge measurements and satellite-derived precipitation.

the amount of the listed events from 22% to 70%, except for event 2 with 70% of positive bias.

As described in the section "Recurrent neural network", two modules were included: a precipitation merging model and a basin scale hydrologic model. The precipitation merging model is set to a linear model, which consists of four set of combinations (P_g , P_s , P_{m1} , and P_{m2}) from two precipitation sources (P_g and P_s). The hydrologic model is implemented by an RNN model with model parameters being calibrated by historical gauged rainfall and runoff data from 13 events during 1988–2000. The validation statistics of runoff forecasts, in terms of RMSE, CC, CE, and MAE, based on four types of inputs, denoted as gauge rainfall, satellite-based rainfall, merged rainfall (non-bias parameters), and merged rainfall with bias parameters, respectively, are listed in Table 3. Because the RNN structure and parameters are carried out by using historical gauge rainfall data, it shows consistently in the evaluation events that the runoff forecasting based on gauge rainfall (RNN_g) as input to the RNN outperforms the runoff forecasting using satellite-based rainfall (RNN_s). Merging of precipitation with/without the bias parameters (RNN_{m1} and RNN_{m2}) consists of marginal improvement, as shown in Table 3. The results show that all models have similar performance except for the satellite data-driven RNN with

much higher bias in validation in terms of higher RMSE and MAE and lower CC and CE.

Table 4 presents the runoff forecasting statistics based on RNN that model parameters calibrated by four storm events in 2004. Compared with Table 3, the RNN model calibrated by an individual precipitation source provides a much improved forecast, which means the RNN model parameters calibrated by a specified source generate more effective result than that of the fixed parameters calibrated solely by historical gauge data. Fig. 5 shows the flood forecasts of four rainfall sources. The blue line means the forecasts were derived by using the parameter sets calibrated by historical gauge precipitation data, whereas the red dash line means the forecasts were derived from the parameter sets calibrated by four precipitation events in 2004. Based on the RNN model parameters found from long-term historical gauge data (Calibration set #1), the validation of forecasted runoff using satellite rainfall estimations seriously underestimated both peak flows and the rising limbs in the inception of storm periods (see Fig. 5b). Applied parameters of RNN found from satellite-based rainfall, the peak flows are improved, but still underestimated in one of the validation events (see Fig. 5b). Basin-averaged gauge rainfall performs well in the validation events (see Fig. 5a). In addition, as shown in Fig. 5c and d, both merged rainfall data sets also

Table 2 Comparison of accumulate precipitation between rain gauge measurements and satellite estimations

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
P_g (mm)	157	135	344	510	438	299
P_s (mm)	65	229	111	400	130	153
Relative bias (%)	-59	70	-68	-22	-70	-49

Table 3 Statistic performances of flood forecasting from different models with identical model parameters

	Calibration set #1				Validation			
	RMSE	CC	CE	MAE	RMSE	CC	CE	MAE
RNN_g	89.7	0.92	0.84	58.1	119.8	0.90	0.80	60.6
RNN_s	—	—	—	—	239.4	0.74	0.19	112.4
RNN_{m1}	—	—	—	—	118.4	0.91	0.80	58.2
RNN_{m2}	—	—	—	—	115.1	0.92	0.81	51.8

Table 4 Statistic performances of flood forecasting from different models with optimal model parameters

	Calibration set #2				Validation (case 1)			
	RMSE	CC	CE	MAE	RMSE	CC	CE	MAE
RNN_g	99.2	0.93	0.86	55.6	111.5	0.96	0.83	75.4
RNN_s	129.8	0.87	0.75	69.2	154.7	0.82	0.66	83.9
RNN_{m1}	99.0	0.93	0.86	55.3	109.7	0.96	0.84	75.6
RNN_{m2}	95.0	0.94	0.87	52.3	96.2	0.98	0.87	66.6

perform well in the validation events. Moreover, the merged rainfall product with bias adjustment has substantially improved RMSE, as well as the other validation statistics, such as CC, CE, and MAE (see Table 4).

Effectiveness of data merging from each precipitation source

This subsection discussed the variation of the merging parameters and the contributions of merging products

on flood forecasting. Fig. 6 shows the changes of model error (MSE) calculated based on normalized units under non-bias condition (P_{m1}) in calibration set #1. The upper small diagram illustrates the variation of errors with merging weighting factor (θ) from 0 to 1.0 and the bigger diagram represents the enlarged portion of θ from 0.9 to 1.0. Basically, the error was gradually reduced as the value of θ increased and the minimum error can be found when θ is 0.96. This means satellite-derived precipitation contributes only 4% to the merged rainfall, as compared

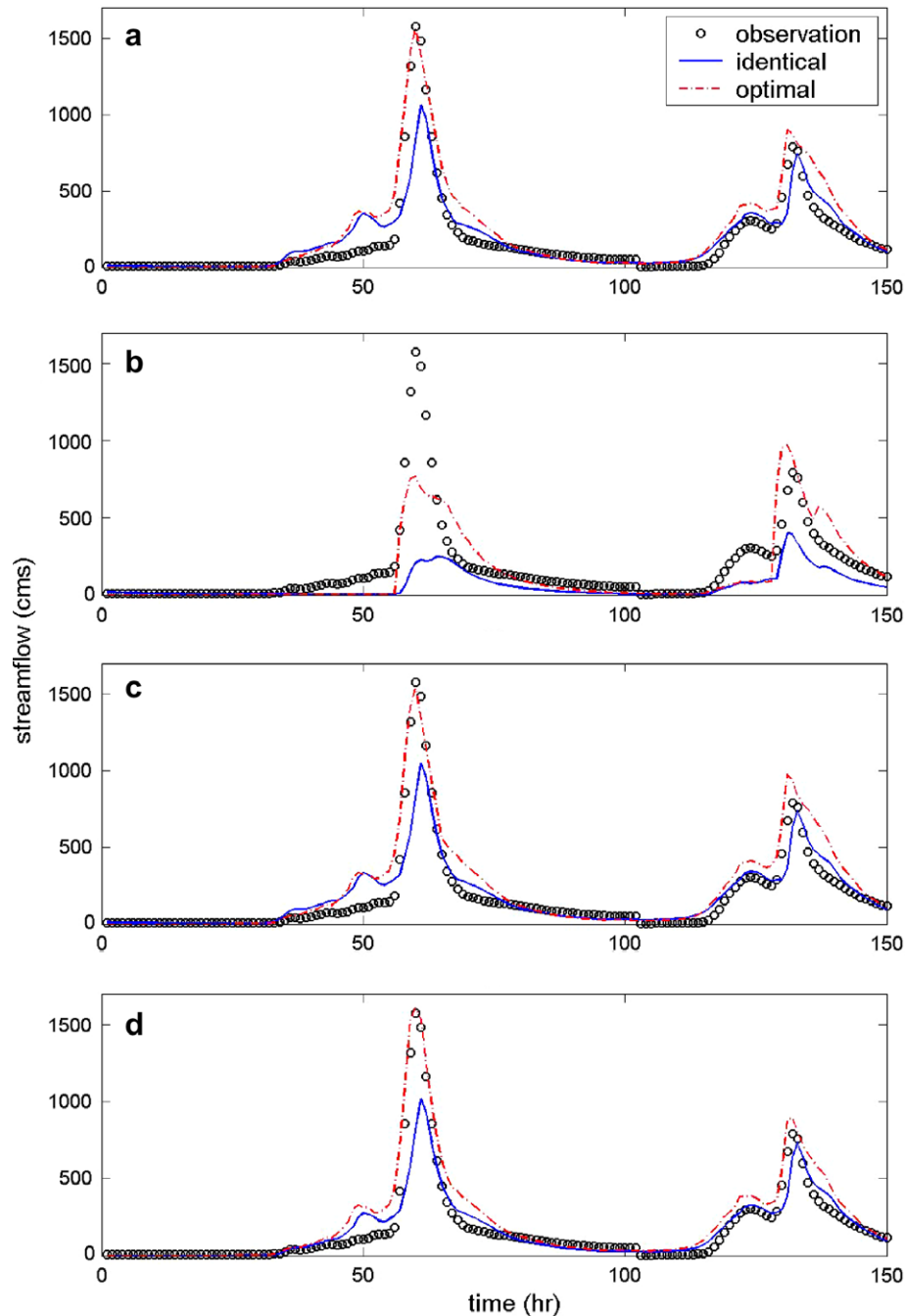


Figure 5 Observed versus forecasted floods with two parameter types that driven from (a) gauge data, (b) satellite estimations, (c) merged rainfall without bias adjustment and (d) merged rainfall with bias adjustment.

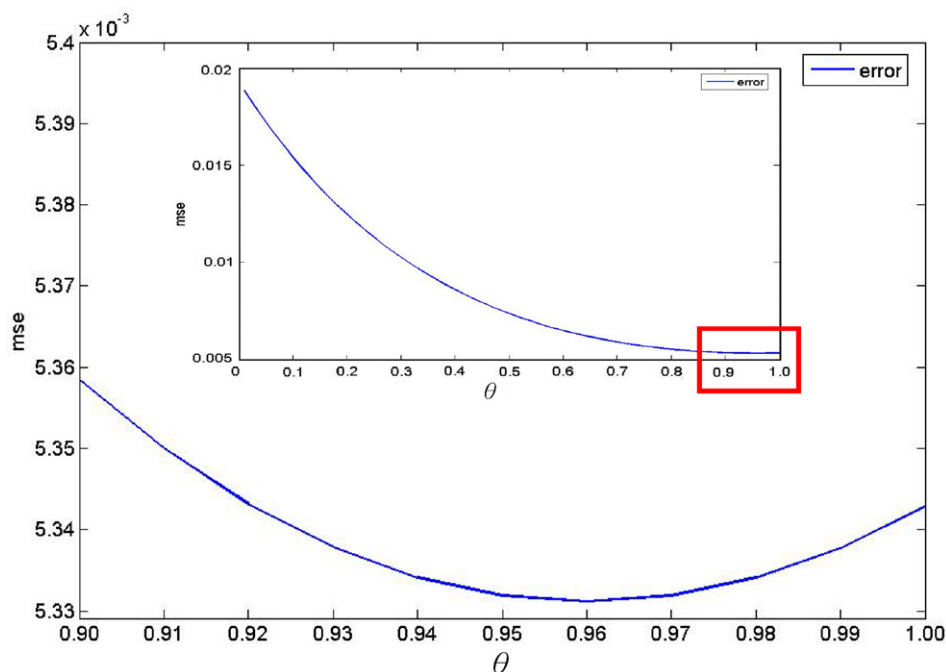


Figure 6 The optimal merging parameter of non-bias condition in calibration set #1.

with gauge observations. Fig. 7 illustrates the changes of model error (MSE) with bias adjustment in calibration set #1. The sensitivity of error with respect to weighting and bias parameters is shown in Fig. 7a. When θ is assigned to 1.0, the weighting factor of satellite-based rainfall ($1 - \theta$) is 0. Since the weighting factor is 0, no matter if the bias parameter of satellite-based rainfall (α_2) is changed, it has no impact on the normalized error. Similarly, when θ is assigned to 0, the weighting factor of gauge observed rainfall is 0. In this case, the bias parameter of gauge observed rainfall (α_1) is not sensitive to the normalized error. We could further find that when $\theta < 0.5$, the normalized error is more sensitive to the variation of α_2 , while when $\theta > 0.5$, the normalized error is more sensitive to the variation of α_1 . The minima error is found at $\{\theta, \alpha_1, \alpha_2\} = \{0.95, 0.04, 0.24\}$ (see Fig. 7b). Fig. 8 represents the variation of error and merging parameters under bias adjustment in calibration set #2. The errors shown in Fig. 8 did not change as smoothly as Fig. 7. However, the tendency of error changes towards the same direction in each sub-figure and the contribution of satellite-derived precipitation can be easily detected, which is 5% (Fig. 8b).

The optimal merging parameters in both calibration datasets show that the satellite-derived precipitation has limited contribution (5%) to merging procedures. However, the merged precipitation helps to improve the flood forecasting with improvement of RMSE for about 2–14%. The improvement was calculated by using the criterion of skill score (SS). For this case, SS was computed from RMSE statistics and the values of RMSE can be referred to Tables 3 and 4. Besides, the results also indicate that both of the gauge and satellite estimates have bias existing, especially in the bias condition of calibration set #2. The bias for gauge measure-

ments and satellite estimates is about -31% and 49%, respectively.

In this case study, the contribution of satellite-derived precipitation is insignificant due to the gauge measurements providing good quality for the flood forecasting. This is evidently true by inspection of Fig. 5a; the gauge measurements produce accurate performance on flood forecasting. Therefore, the improvement made by other rainfall sources is limited.

Merging of precipitation from poor quality gauge data

For further evaluating the effect of gauge quality to the optimal merging parameters, the basin-averaged rainfall calculated from one and two gauges are tested. The use of gauge was randomly selected for the test. Table 5 summarizes the validation performance merging options from using two gauges (case 2) and one gauge (case 3) to calculate the basin-averaged rainfall. For gauge data-driven models, the statistics clearly indicated that using one gauge resulted in worst forecasts. Besides, model efficiency in the three gauge case (case 1) was better than that of two gauges (case 2) and one gauge scenarios (case 3) (see Tables 4 and 5). The results demonstrated that the information provided by more gauges was superior to fewer gauges. Figs. 9–11 show the flood forecasted from gauge data, merged data without bias adjustment, and merged data with bias adjustment in all cases. The results shown in the figures were consistent with the tables that the information supported by one gauge was not good enough to produce accurate flood forecasting.

It should be noticed that although gauge observation is used as input, the model performance is sensitive to the

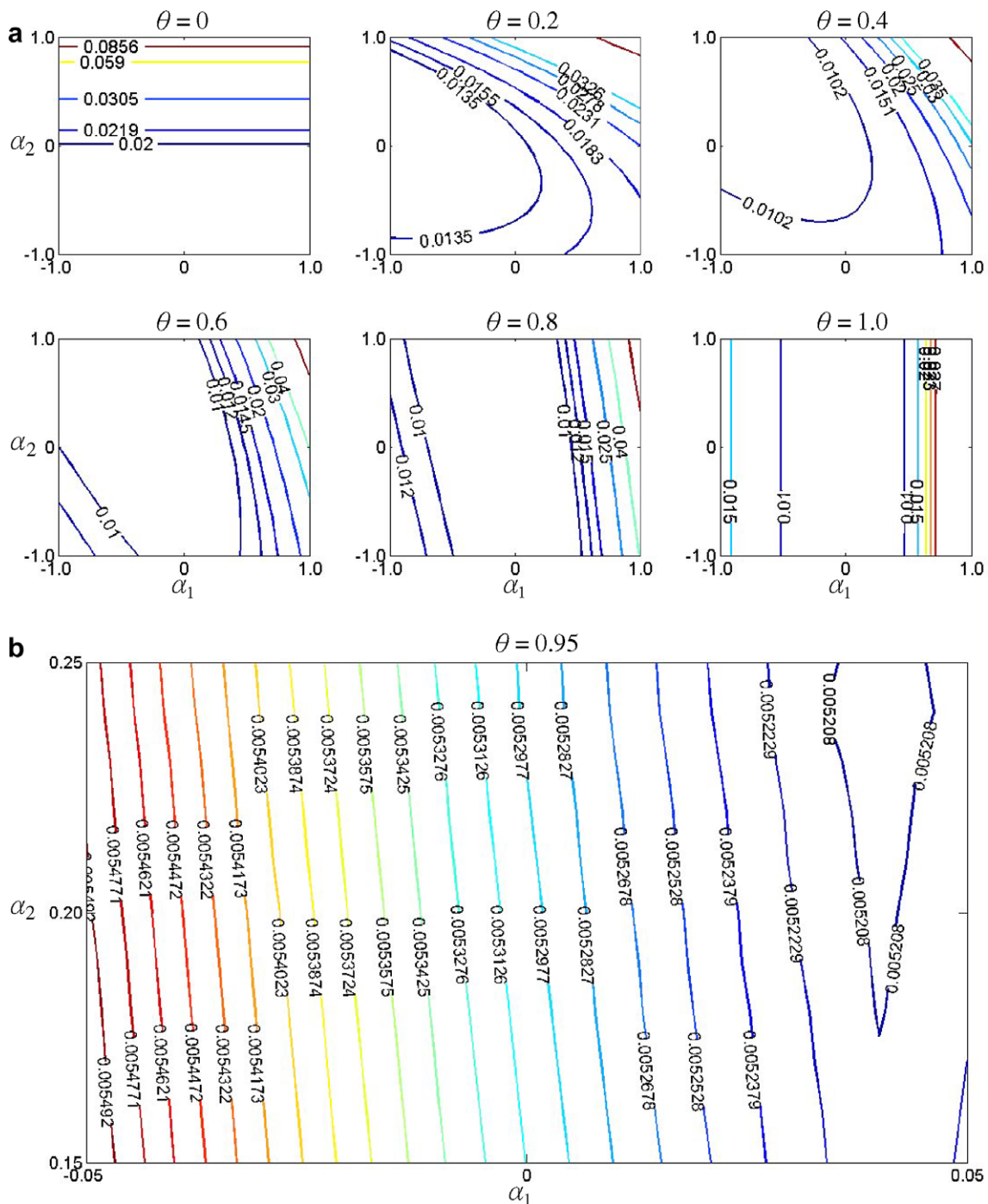


Figure 7 The error changes of bias condition in calibration set #1: (a) the pattern of error with θ varied from 0 to 1 and (b) optimal merging parameters.

gauge size included, and the performance of both merged data-driven models was similar. For example, the RMSE calculated from the merged rainfalls was about 110, 111, and 113 (cm s) in cases 1, 2, and 3 under non-bias condition. In bias condition, the RMSE calculated from merged rainfalls was about 96, 98, and 100 (cm s) in case 1, 2, and 3, respectively. This was mainly because the contribution from satellite-derived precipitation increased when the reliability

from gauge data was poor. This can be demonstrated by inspection of Table 6. Table 6 presents the optimal merging parameters and the improvements made by two merged precipitation products in all three cases. As mentioned above, the contributions made from satellite-derived precipitation in case 1 are 2% and 5% with and without bias adjustment, respectively. When two gauges were used, the contributions from satellite-based rainfall increased to

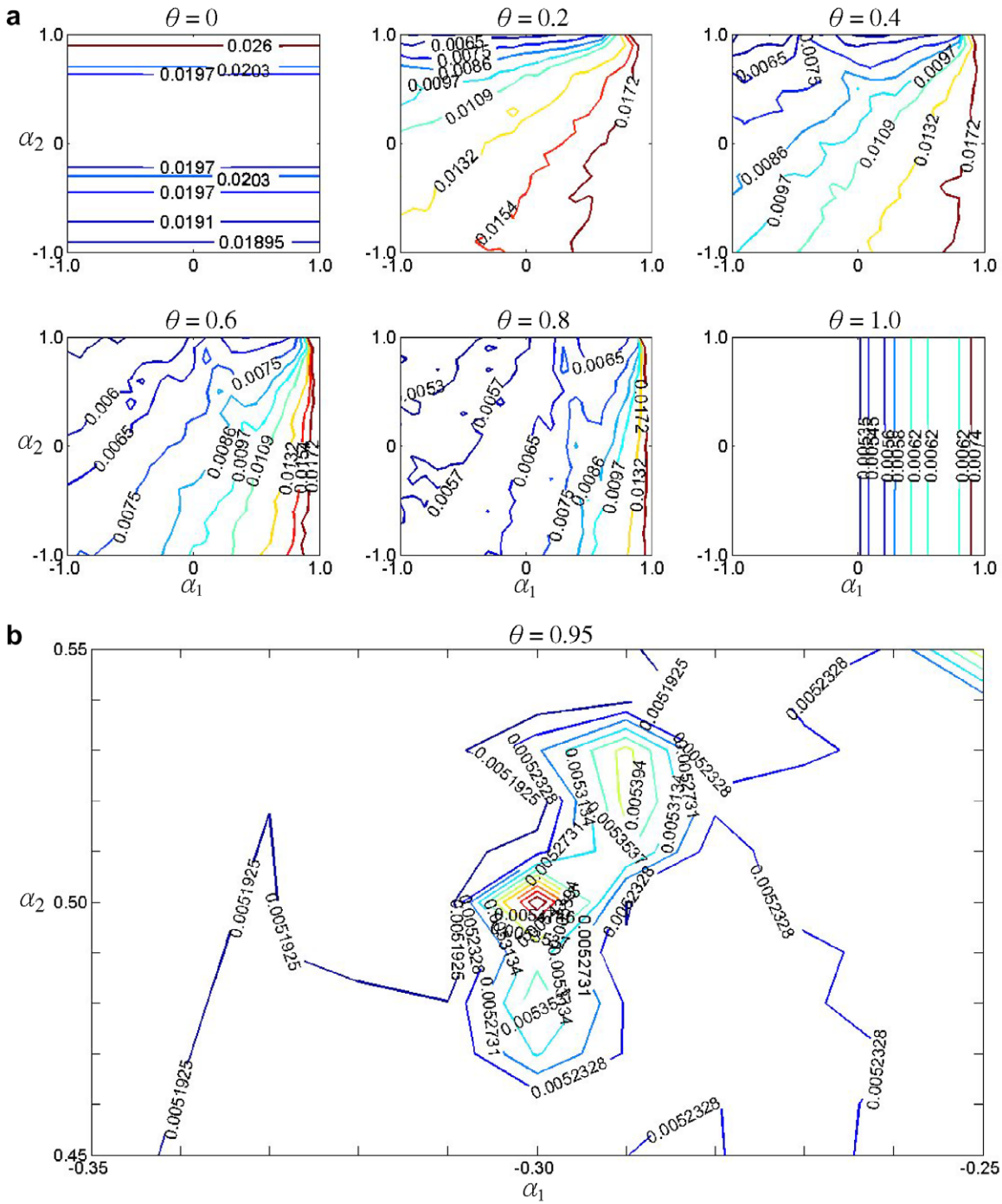


Figure 8 The error changes of bias condition in calibration set #2: (a) the pattern of error with θ varied from 0 to 1 and (b) optimal merging parameters.

Table 5 Validation performances of flood forecasting from three models in different cases

	Case 2 (two gauges- R_2 and R_3)				Case 3 (one gauge- R_3)			
	RMSE	CC	CE	MAE	RMSE	CC	CE	MAE
RNN _g	117.7	0.91	0.81	70.5	126.4	0.90	0.78	69.8
RNN _{m1}	111.1	0.92	0.83	54.7	113.1	0.92	0.83	58.5
RNN _{m2}	98.0	0.95	0.87	67.8	99.5	0.94	0.86	62.9

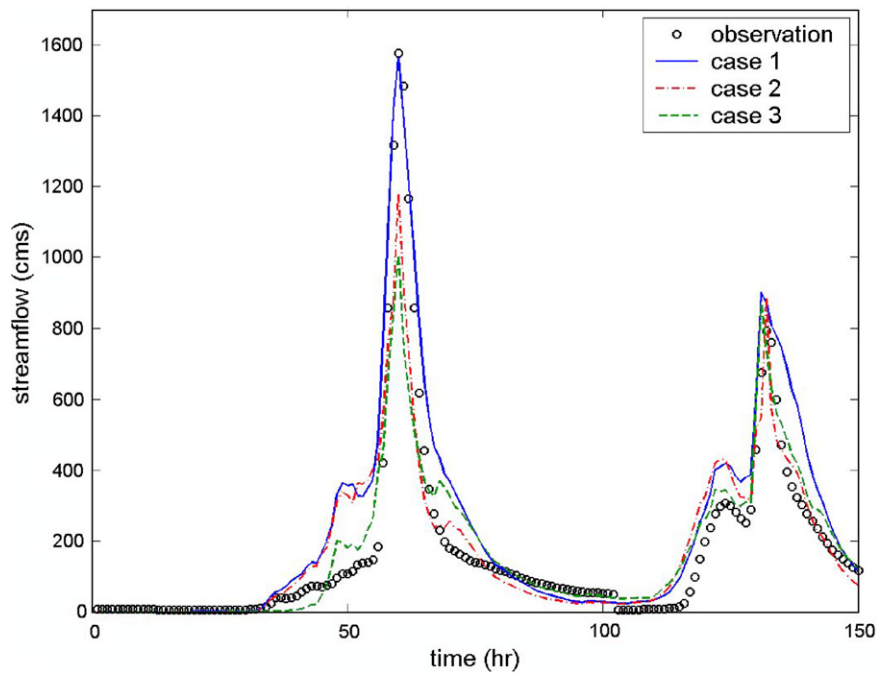


Figure 9 Flood forecasting from gauge data-driven in different cases.

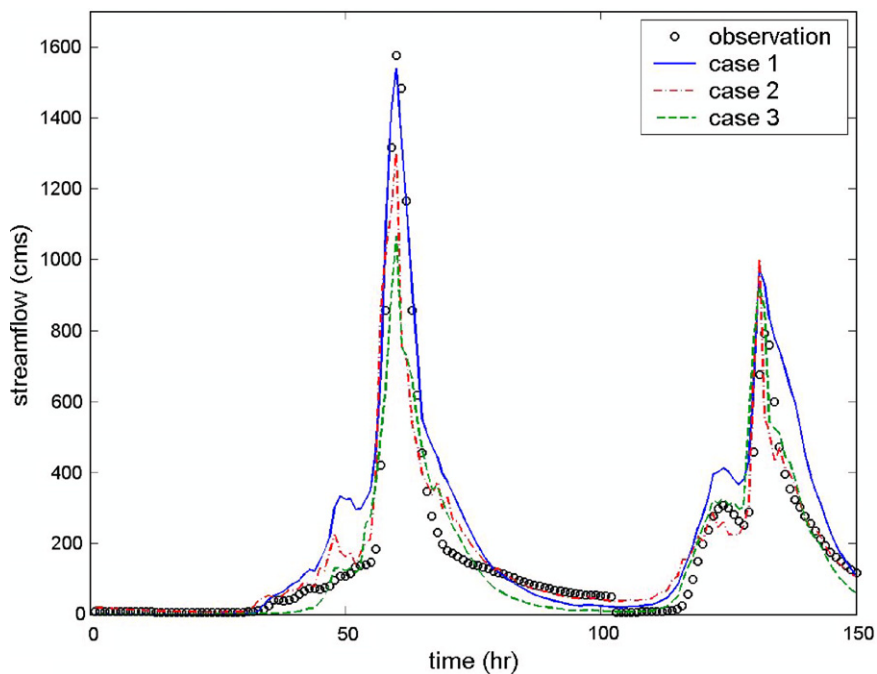


Figure 10 Flood forecasting from merged data-driven without bias adjustment in different cases.

15% in both bias and non-bias settings. Likewise, when one gauge was used, the contribution of satellite-derived precipitation was largely increased to 44% in non-bias merging setting and to 49% in the bias merging setting. This might be due to larger measurement error within R_3 gauge because the total precipitation of R_3 gauge (2500 mm) has a significant difference as compared with other two gauges

(1546 mm in R_2 gauge and 1631 mm in R_1 gauge). To note that the value of optimal merging parameter also reflects a positive bias of $\alpha_1 = 0.52$ in case 3. Overall, these results indicate that the satellite-derived precipitation provides relative effectiveness to the merging product, depending on the quality of gauge measurement, for the hydrological forecasting.

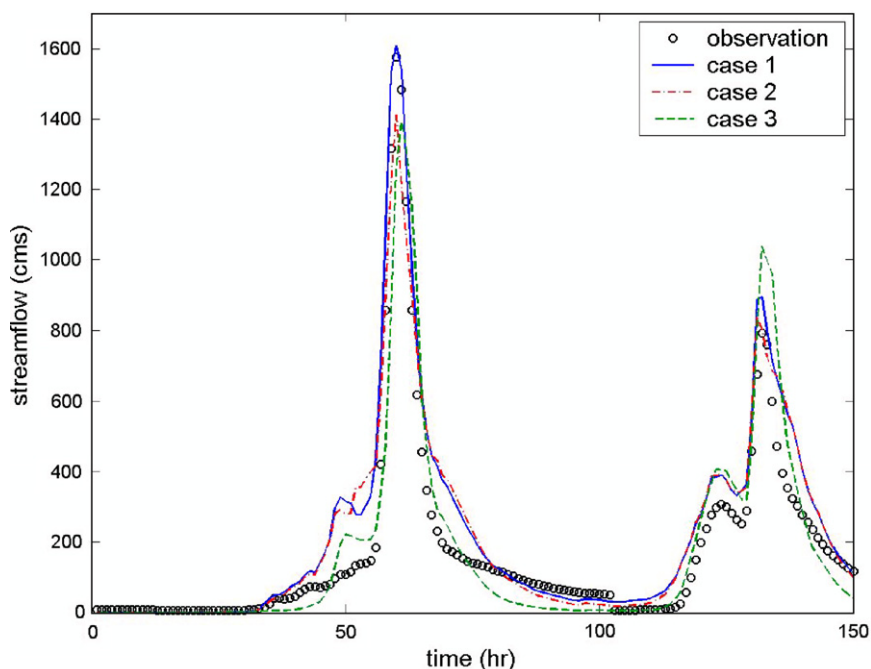


Figure 11 Flood forecasting from merged data-driven with bias adjustment in different cases.

Table 6 Optimal merging parameters in different cases and their improvements

	Assumption	$1 - \theta$	α_1	α_2	RMSE	SS (%)
Case 1	Non-bias	0.02	—	—	109.7	2
	Bias	0.05	-0.31	0.49	96.2	14
Case 2	Non-bias	0.15	—	—	111.1	6
	Bias	0.15	-0.11	0.54	98.0	17
Case 3	Non-bias	0.44	—	—	113.1	10
	Bias	0.49	0.52	0.53	99.5	21

Conclusions

We have investigated the effectiveness of merged satellite-derived precipitation and rain gauge measurements in flood forecasting during typhoon periods in Taiwan. The satellite precipitation estimations over Taiwan were generated through Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN CCS) at grid resolution of 4 km. Four basin-averaged rainfalls were calculated from satellite estimates, gauge observations and two merged rainfall products; these four rainfalls were used to investigate their impacts and contributions on flood forecasting.

The experiments show that the validation performance is dependent on the data sources used in the calibration phase. Because each precipitation source is subject to error of many factors, better forecasting performance may be obtained by suitable combination of several precipitation observations. The results also exemplify the need for satellite-derived precipitation on merging procedures when the gauge information is insufficient and for a detailed investigation of the techniques that use the merged rainfall products into hydrological models. From the merged weighting

parameters, it shows that satellite-derived precipitation contributes around 5% on merged product in case 1. The improvements of runoff forecasting in terms of RMSE of gauged simulation (111.5 cm s) can be as high as 14% from the merged data sources (96.2 cm s). For case 2, the contributions made from satellite-derived precipitation increased to 15% and resulted in 17% improvement of runoff forecasting. For case 3, it was surprising that the contributions of satellite-derived precipitation were almost equal to gauge measurements (44% without bias adjustment and 49% with bias adjustment) and helped to improve the accuracy of flood forecasting by 10% and 21% improvements in non-bias and bias adjustment, respectively. This demonstrates that the merged method used in this study can efficiently combine the information from both rainfall sources to improve the accuracy of flood forecasting during typhoon periods.

In this study, there are two reasons for using the merging procedure instead of directly inputting two precipitation sources to the RNN. First, the contribution from specific source to the merged product can be explicitly presented. By contrast, it would have difficulty in getting the precise contribution (weight) if these two sources are directly used as model inputs because the relations of connected weights

between input variables and different hidden neurons can not be clearly explained. Secondly, there is a need to take into account if the observational error in gauge measurements and estimated error in PERSIANN-CCS outputs are existent. Results obtained in this study indicated that both precipitation sources are biased.

From the case studies, although using gauge rainfall measurement could give good forecasting of runoff time series, when the gauge measurements are poor, satellite-derived precipitation may show its add-on value to the short-term hydrologic forecasting. Besides, the satellite information is also important to extend the time-ahead of flood prediction. In this case study, the mean lag-time between the raining event and runoff production is about 2–3 h which highly depend on the intensity and pattern of raining event. This physical lag-time limits the lead time of flood forecasting to less than 3 h if rainfall observations are used as model inputs. For long-term forecasting (lead time greater than 3 h), it is necessity of using precipitation prediction as inputs for the hydrologic model. Investigation on this issue is currently being explored by the authors.

Acknowledgements

This study was partially supported by the National Science Council (NSC), Taiwan ROC for Graduate Students Study Abroad Program and by the Center for Hydrometeorology and Remote Sensing (CHRS), UC Irvine.

References

- Atiya, A.F., Aly, M.A., Parlos, A.G., 2005. Sparse basis selection: New results and application to adaptive prediction of video source traffic. *IEEE Transactions on Neural Networks* 16 (5), 1136–1146.
- Binner, J.M., Elger, T., Nilsson, B., Tepper, J.A., 2004. Tools for non-linear time series forecasting in economics – An empirical comparison of regime switching vector autoregressive models and recurrent neural networks. *Applications of Artificial Intelligence in Finance and Economics. Advances in Econometrics: a Research Annual*, pp. 71–91.
- Campolo, M., Andreussi, P., Soldati, A., 1999. River flood forecasting with a neural network model. *Water Resources Research* 35 (4), 1191–1197.
- Chang, F.J., Chen, Y.C., 2003. Estuary water-stage forecasting by using radial basis function neural network. *Journal of Hydrology* 270, 158–166.
- Chang, F.J., Chang, L.C., Huang, H.L., 2002. Real-time recurrent learning neural network for stream-flow forecasting. *Hydrological Processes* 16 (13), 2577–2588.
- Chang, L.C., Chang, F.J., Chiang, Y.M., 2004. A two-step-ahead recurrent neural network for stream-flow forecasting. *Hydrological Processes* 18 (1), 81–92.
- Chen, S.H., Lin, Y.H., Chang, L.C., Chang, F.J., 2006. The strategy of building a flood forecast model by neuro-fuzzy network. *Hydrological Processes* 20, 1525–1540.
- Chiang, Y.M., Chang, L.C., Chang, F.J., 2004. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. *Journal of Hydrology* 290 (3–4), 297–311.
- Chiang, Y.M., Chang, F.J., Jou, B.J.D., Lin, P.F., 2007. Dynamic ANN for precipitation estimation and forecasting from radar observations. *Journal of Hydrology* 334 (1–2), 250–261.
- Coulibaly, P., Anctil, F., Bobee, B., 2001. Multivariate reservoir inflow forecasting using temporal neural networks. *Journal of Hydrologic Engineering* 6 (5), 367–376.
- Ham, F.M., Kostanic, I., 2001. *Principles of Neurocomputing for Science & Engineering*. McGraw-Hill, New York.
- Han, M., Xi, J.H., Xu, S.G., Yin, F.L., 2004. Prediction of chaotic time series based on the recurrent predictor neural network. *IEEE Transactions on Signal Processing* 52 (12), 3409–3416.
- Haykin, S., 1999. *Neural Network: A Comprehensive Foundation*, second ed. Prentice Hall.
- Ho, S.L., Xie, M., Goh, T.N., 2002. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers & Industrial Engineering* 42 (2–4), 371–375.
- Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X.G., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology* 43 (12), 1834–1852.
- Hsu, K.L., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research* 31, 2517–2530.
- Hsu, K.L., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology* 36, 1176–1190.
- Kumar, A., Agrawal, D.P., Joshi, S.D., 2003. Study of Canada/US dollar exchange rate movements using recurrent neural network model of FX-market. *Advances in Intelligent Data Analysis V. Lecture Notes in Computer Science*, pp. 409–417.
- Luk, K.C., Ball, J.E., Sharma, A., 2000. A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *Journal of Hydrology* 227, 56–65.
- Maier, H.R., Dandy, G.C., 1996. The use of artificial neural networks for the prediction of water quality parameters. *Water Resources Research* 32 (4), 1013–1022.
- Mandic, D.P., Chambers, J.A., 1998. A posteriori real-time recurrent learning schemes for a recurrent neural network based nonlinear predictor. *IEEE Proceedings-Vision Image and Signal Processing* 145 (6), 365–370.
- Minns, A.W., Hall, M.J., 1996. Artificial neural networks as rainfall-runoff models. *Hydrological Sciences* 41 (3), 399–417.
- Pan, T.Y., Wang, R.Y., 2004. State space neural networks for short term rainfall-runoff forecasting. *Journal of Hydrology* 297, 34–50.
- Parlos, A.G., Rais, O.T., Atiya, A.F., 2000. Multi-step-ahead prediction using dynamic recurrent neural networks. *Neural Networks* 13 (7), 765–786.
- Parlos, A.G., Parthasarathy, S., Atiya, A.F., 2001. Neuro-predictive process control using on-line controller adaptation. *IEEE Transactions on Control Systems Technology* 9 (5), 741–755.
- Saad, E.W., Prokhorov, D.V., Wunsch, D.C., 1998. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. *IEEE Transactions on Neural Networks* 9 (6), 1456–1470.
- Sajikumar, N., Thandaveswara, B.S., 1999. A non-linear rainfall-runoff model using an artificial neural network. *Journal of Hydrology* 216, 32–55.
- Valdez-Castro, L., Baruch, I., Barrera-Cortes, J., 2003. Neural networks applied to the prediction of fed-batch fermentation kinetics of *Bacillus thuringiensis*. *Bioprocess and Biosystems Engineering* 25 (4), 229–233.
- Zhang, J.S., Xiao, X.C., 2000. Predicting chaotic time series using recurrent neural network. *Chinese Physics Letters* 17 (2), 88–90.