CrossMark

ORIGINAL ARTICLE

Genetic algorithm and fuzzy neural networks combined with the hydrological modeling system for forecasting watershed runoff discharge

**Chih-Chieh Young · Wen-Cheng Liu ·
Chung-En Chung**

Abstract The accurate prediction of hourly runoff discharge in a river basin during typhoon events is of critical importance in operational flood control and management. This study utilizes three model approaches to predict runoff discharge in the Laonong Creek basin in southern Taiwan: the hydrological engineering center hydrological modeling system (HEC-HMS) model and two hybrid models which combine the HEC-HMS model with a genetic algorithm neural network (GANN) and an adaptive neuro-fuzzy inference system approach (ANFIS). Hourly runoff discharge data during seven heavy rainfall/typhoon events were collected for model calibration (training) and validation. Six statistical indicators [i.e., mean absolute error, root-mean-square error, coefficient of correlation, error of time to peak discharge, error of peak discharge, and coefficient of efficiency (CE)] were used to evaluate the prediction accuracy. The simulation results indicate that the HEC-HMS model cannot satisfactorily predict hourly runoff discharge during the typhoon events. Both hybrid approaches that use the HEC-HMS model in conjunction with the GANN and ANFIS models can significantly improve the prediction accuracy for the n-h-ahead runoff discharge.

C.-C. Young
Hydrotech Research Institute, National Taiwan University,
Taipei 10617, Taiwan

W.-C. Liu (✉) · C.-E. Chung
Department of Civil and Disaster Prevention Engineering,
National United University, Miao-Li 36063, Taiwan
e-mail: wcliu@nuu.edu.tw

W.-C. Liu
Taiwan Typhoon and Flood Research Institute, National Applied
Research Laboratories, Taipei 10093, Taiwan

1 Introduction

Accurate prediction of runoff discharges is a critical topic in hydrology and water resources. Hydrological problems are highly nonlinear and complicated, exhibiting a wide degree of variability in space and time [1]. Therefore, hydrological modeling has always been an important task for water resources planning and management [2].

Hydrological models provide a simplified representation of an actual watershed system to obtain a better understanding of hydrological processes in the study area, e.g., the characteristics of a watershed and its responses to different inputs [2–4]. In the past decades, continuous efforts have been made to develop accurate rainfall–runoff modeling techniques. They can be divided into two general classes: physically based and/or data-driven approaches. Physically based models express the physical behavior of the hydrological system by solving partial differential equations that represent the flow processes within the watershed to the best of our knowledge. Data-driven models treat the relationship between rainfall (i.e., input) and runoff (i.e., output) as a black box using historical measured results with the algorithms developed in the areas of statistics, soft computing, computational intelligence, machine learning, and data mining [5–8].

Artificial neural networks (ANNs) are one kind of data-driven models suitable for rainfall–runoff modeling as reported in the previous studies [9–11]. ANNs learning the rainfall–runoff relationship through the training process can quantitatively predict the runoff without requiring the catchment characteristics. ANNs also demonstrate the

superiority over most classical time series modeling techniques in terms of handling highly nonlinear processes between input and output variables. Models of this type are efficient and capable of giving excellent prediction results, in comparison with those physically based models. Rezaeianzadeh et al. [12] carried out a comparative study of rainfall–runoff prediction between a physically based model and a data-driven model. They showed that the hydrological engineering center hydrological modeling system (HEC-HMS) model described the runoff discharge response to rainfall events (to a certain degree). The multilayer perceptron neural network (MLPNN) model predicted peak flows and annual flood volumes more accurately than the HEC-HMS model.

For better prediction accuracy, different thought has been evolved. For example, the discharges simulated by a single rainfall–runoff model can be further updated with linear/nonlinear systems [13, 14]. Instead of switching from one model to another, an alternative approach is the combination of various forecasts produced from several models. The basic idea is that better solutions could be obtained since each model captures a certain aspect of the data. Shamseldin [15] was the pioneer constructing the multi-model combined predictions based upon five rainfall–runoff models with three kinds of combinations (i.e., simple averaging method, a regression-based scheme, and a neural network approach). Generally, the combined models could provide more accurate predictions than any single model. Afterward, the multi-model combination has been extensively implemented for hydrological research [16–21].

According to the complete literature review, it can be recognized that simulating physical processes in a watershed are of critical importance for water resources system management. The data-driven ANNs based upon sufficient training with appropriate input–output data sets can accurately predict the runoff discharge in the watershed [22, 23] but with the drawback of black-box feature hindering the simulation of physical processes [24]. The multi-model combination approach for hydrological research is promising. Recently, Young and Liu [25] proposed a physically based and ANN hybrid model. In their study, the performance of the physically based model, the data-driven ANN model, and the hybrid model was carefully examined. The main conclusion was that the n-step ahead prediction from the single ANN model showed apparent accumulated errors although model of this type could approximate most nonlinear functions demanded by practice. The hybrid model with the additional input effectively reduced the errors, revealing the role of the physically based model. While the hybrid model has improved the predicted runoff discharge in a watershed, a challenge to determine an appropriate structure as well as the interior parameters of the black-box model still remains. The genetic algorithm (GA) and/or

fuzzy logic approach can be a solution to the conventional used trial-and-error procedure.

The objective of the present study was to develop a new hybrid model for rainfall–runoff forecasting. The well-known HEC-HMS model was firstly applied to simulate the runoff discharges in the Laonong Creek basin of southern Taiwan. Two ANN models including genetic algorithm neural network (GANN) and adaptive neuro-fuzzy inference system (ANFIS) were subsequently adopted to improve the predictions from the HEC-HMS model under different leading time. Six quantitative statistical measures [the mean absolute error (MAE), the root-mean-square error (RMSE), the coefficient of correlation, the error of time to peak discharge, the error of peak discharge, and the CE] were used to assess the predictions of runoff discharges. The accuracy of these models is carefully discussed in this paper.

2 Study site and data collection

The Laonong Creek basin is a subbasin of the Kaoping River watershed in southern Taiwan. The total area of the Laonong Creek basin is 1,375 km^2, and the length of the main stream is 133 km. The mean elevation of the basin is 1,515 m, and the mean slope is 0.6574 m/m. The Laonong Creek basin is of six-stream order. There are four rain gauge stations in the basin: the Tengzhi, Meishan, Gaozhong, and Tianchi stations. A gauge for discharge measurement is also located at the Hsinfa Bridge station (Fig. 1). Data from these stations were collected from the Water Resources Agency, Taiwan. The measured rainfall and runoff discharge during the Bilis Typhoon in 2006 and the Sepat Typhoon in 2007 are shown in Fig. 2 for illustration. The rainfall data during these two typhoon events gave different patterns for four gauge stations, suggesting that measurements of precipitation in the basin are subject to regional and topographic effects.

3 Methodology

In this study, two hybrid models that combine the HEC-HMS with GANN and ANFIS models were presented to obtain an accurate rainfall–runoff prediction. In the following, a brief description of the well-known HEC-HMS model and the basic idea of the hybrid models will be given.

3.1 The HEC-HMS rainfall–runoff model

The HEC-HMS model developed by the US Army Corps of Engineers [26] can be applied to various types of hydrological simulations and analyses [27].
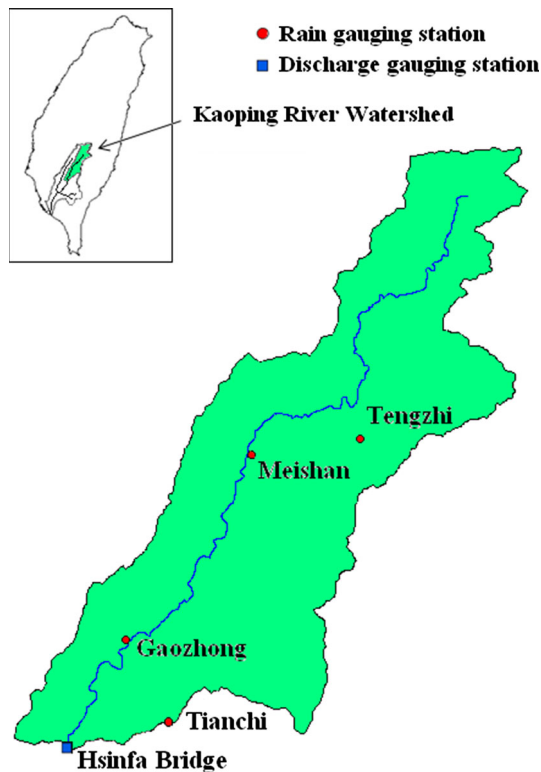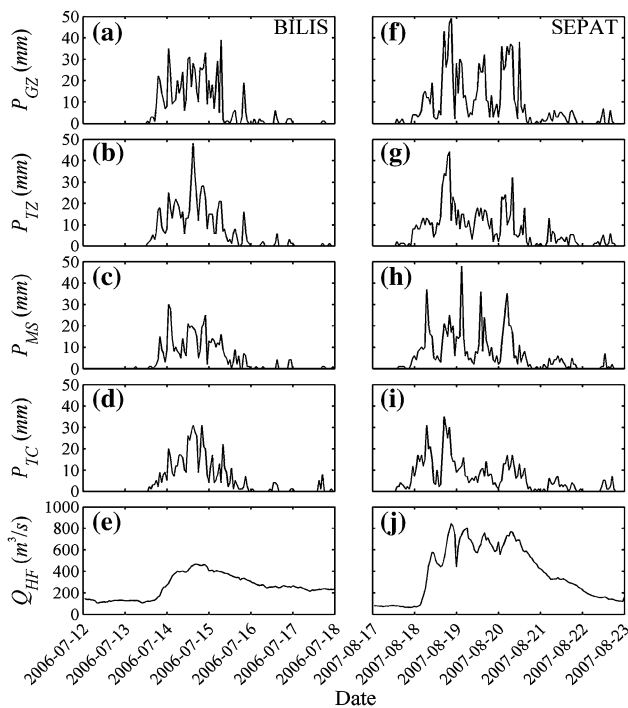
Fig. 1 Study site of Laonong Creek basin

Fig. 2 Measured rainfall and discharge data during the Bilis and Sepat Typhoons. **a**, **f** Rainfall at the Gaozhong station. **b**, **g** Rainfall at the Tengzhi station. **c**, **h** Rainfall at the Meishan station. **d**, **i** Rainfall at the Tianchi station. **e**, **j** Runoff (discharge) at the Hsinfa Bridge gauge station

In the HEC-HMS model, land and water bodies in a watershed can be classified as the directly connected impervious and pervious surfaces. On the directly connected impervious surface, precipitation runs off with no volume losses. For the pervious land surfaces, the contributing precipitation is subject to infiltration, which is simply modeled without the consideration of storage and movement of water within the soil layer. The overland flow and near-surface flow are implicitly combined to form the direct runoff.

In terms of the infiltration loss, partitioning of rainfall into infiltration and runoff is defined/derived from a set of empirical equations, i.e.,

$$P_e = \frac{(P - I_a)^2}{P - I_a + S} \tag{1}$$

where P and P_e are the accumulated rainfall depth and precipitation excess at time t, respectively, I_a is the initial abstraction (loss), and S is the potential maximum retention, i.e., the ability of a watershed to retain storm precipitation. Based upon analysis of many experimental watersheds, the Soil Conservation Service (SCS) gave an empirical relationship $I_a = 0.2\ S$, where the maximum retention S is determined using the following equation (SI unit system, cm)

$$S = \frac{2{,}540}{CN} - 25.4 \tag{2}$$

with the SCS curve number (CN) in the range between 1 and 100. The CN indicator determining storm runoff (or spatially distributed infiltration capability) based on the land use, land cover types, and hydrological soil group [28] can be obtained from the standard table developed by the SCS [29]. For example, the value of CN is 98 in an impervious area.

Translation and attenuation of spatially distributed precipitation excess to runoff are modeled using the modified Clark unit hydrograph (ModClark) algorithm [30]. The storm hydrograph (Q) is calculated by convolving the precipitation increments (P) with unit hydrograph ordinates (U), i.e., $Q_n = \sum_{m=1}^{n \leq M} P_m U_{n-m+1}$, where m increase from 1 to n. The time of concentration for a watershed t_c (as a function of basin length and slope) is used to calculate the travel time or translation lag for each cell in the basin, i.e., $t_{cell} = t_c (d_{cell}/d_{max})$, where d_{cell} and d_{max} are the distance from the cell to the outlet and the maximum traveling distance in the watershed, respectively [27]. In addition, the storage coefficient R is used to represent attenuation/reduction of discharge as excess is stored in a watershed. The coefficient value can be estimated as the flow divided by its time derivative at the inflection point on the recession limb of the hydrograph. The cell outflow hydrograph

is routed (using a linear reservoir concept) with the following equation [27]

$$Q(t) = \frac{\Delta t}{R + 0.5\Delta t} \times I(t) + \left[1 - \frac{\Delta t}{R + 0.5\Delta t} \times Q(t-1) \right] \tag{3}$$

where $Q(t)$ and $Q(t-1)$ are the outflow from storage, respectively, at current time t and previous time $t-1$; $I(t)$ is average inflow to storage at time t; R is the storage coefficient; and Δt is the time increment.

Baseflow that defines a minimum river depth also plays an important role in flood studies. Without consideration of baseflow, models may underestimate water levels and fail to identify inundated reaches. Typically, baseflow can be addressed using an exponential decrease function $Q = Q_0 e^{-kt}$, where Q_0 is an average value of the initial baseflow before a storm; k is an exponential decay constant; t is the time [31].

In this study, the hydrological modeling was carried out by adopting the SCS-CN loss method, Clark transform method, and recession baseflow method. Besides, the Muskingum–Cunge Standard Section method was employed for channel routing. The key parameters, CN, k, t_c, and R, are set to constant values which are 83.5, 0.35, 3.2, and 4.15, respectively. A sensitivity analysis is performed with these four parameters. We found that CN and t_c values are the most sensitive and less sensitive parameters to affect simulated results, respectively. The time step was set as 1 h based upon the sampling frequency of

measured rainfall and discharge data in the typhoon events. Note that the HEC-HMS model with given rainfall data from the field measurement or numerical weather predictions can be used to simulate or forecast the river runoff in the historical, present, or future typhoon events. The outflow at each sampling time t is computed based upon the information at the previous time $t-1$ (see Eq. 3). Direct computation form time $t-1$ to $t+1$ (or $t+n$) is not available. The whole hydrograph is obtained after the (recursive) time integration.

3.2 The HEC-HMS–GANN hybrid model

GA searches for the optimal value of a complex function through a natural process of gene evolution: selection, crossover, and mutation [32]. Initially, random population of chromosomes is generated as candidate solutions. Then, the fitness function (e.g., the MAE of predictions) of each chromosome is evaluated to determine the probability for the fitness-based selection. Subsequently, crossover operation combines a pair of selected chromosomes and alters the gene components at specific positions to create the next generation. The rarely happened mutation operator can change an arbitrary gene component, so that the offspring can have new (or even better) characteristics. The evaluation of next population and the genetic manipulation processes are repeated until the evolution has reached the stop or convergent criterion. It has been proved that the genetic manipulation (i.e., selection, crossover, and mutation)

Fig. 3 Architecture of the back-propagation neural network (BPNN) algorithm
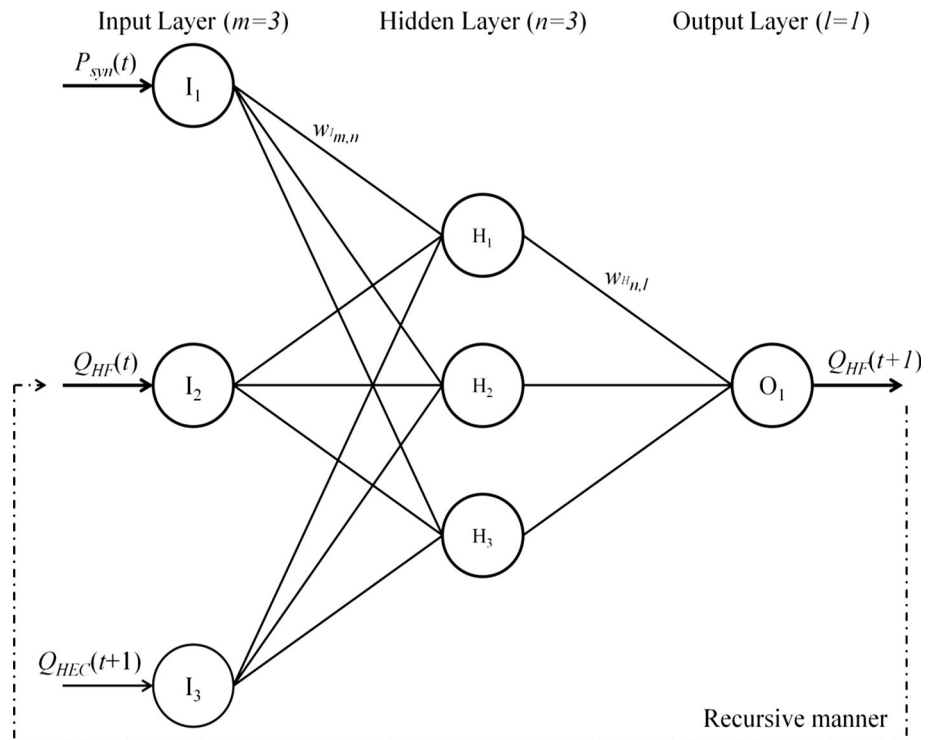
enables GA to explore virtually all regions of the state space, exploit the promising areas, and find out the global optima in a complicated problem [33].

ANN methods are capable of recognizing patterns in complex data through a nonlinear mapping between inputs and outputs. An optimal network architecture and its associated weights play an important role in successful ANN modeling. In most cases, the network parameters (e.g., number of neurons and layers) are usually determined based upon trial-and-error procedure. The weights are then obtained by using the back-propagation algorithm. Actually, the optimality is not guaranteed due to a random type of searching in a small state space [34]. To tackle the above issue, the so-called GANN that applies GA to optimize both the architecture and weights of ANN has been proposed [35].

In this study, we construct the HEC-HMS–GANN hybrid model using a three-layer network structure with 3, 3, and 1 neurons in the input, hidden, and output layers, respectively (as shown in Fig. 3). Consistent with the computation procedure of a physically based model, the measured rainfall–runoff data at time t hour are given to forecast runoff discharge at time $t + 1$ h. Besides, the physically based prediction result at time $t + 1$ h serves as an additional input since the computation of whole hydrograph has been carried out over the time integration. Thus, the three nodes in the input layer correspond to $P_{syn}(t)$, $Q_{HF}(t)$, and $Q_{HEC}(t + 1)$, where $P_{syn}(t) = \sum_{i=1}^{4} w_i P_i(t)$ is the synthetic rainfall intensity at time t hour; $P_i(t)$ is the rainfall intensity at time t hour for different rain gauge stations ($i = 1, 2, 3,$ and 4); w_i is the weighting factor determined by the Thiessen polygon method; $Q_{HF}(t)$ is the observed discharge at time t hour at the Hsinfa Bridge gauge station; and $Q_{HEC}(t + 1)$ is the runoff discharge at time $t + 1$ h at the Hsinfa Bridge gauge station obtained from the HEC-HMS model. The node in the output layer is the hourly runoff discharge prediction at time $t + 1$ h at the Hsinfa Bridge gauge station, $Q_{HF}(t + 1)$. Notice that we have also considered the input rainfall–runoff data with more lagged times (which are typically incorporated in time series modeling) and found insignificant differences in the prediction results. For the hidden layer, both the GA optimization and trial-and-error procedure indicate an overfitting issue if more than three hidden nodes are applied. Afterward, GA is further used to optimize the weights for the determined network architecture through the random initialization, fitness calculation, and genetic manipulation (i.e., selection, crossover, and mutation) procedures (see the flowchart in Fig. 4). In terms of the GA parameters, the population size, maximum generation, and crossover probability are 100, 5,000, and 0.8, respectively. When the training stage is completed, the
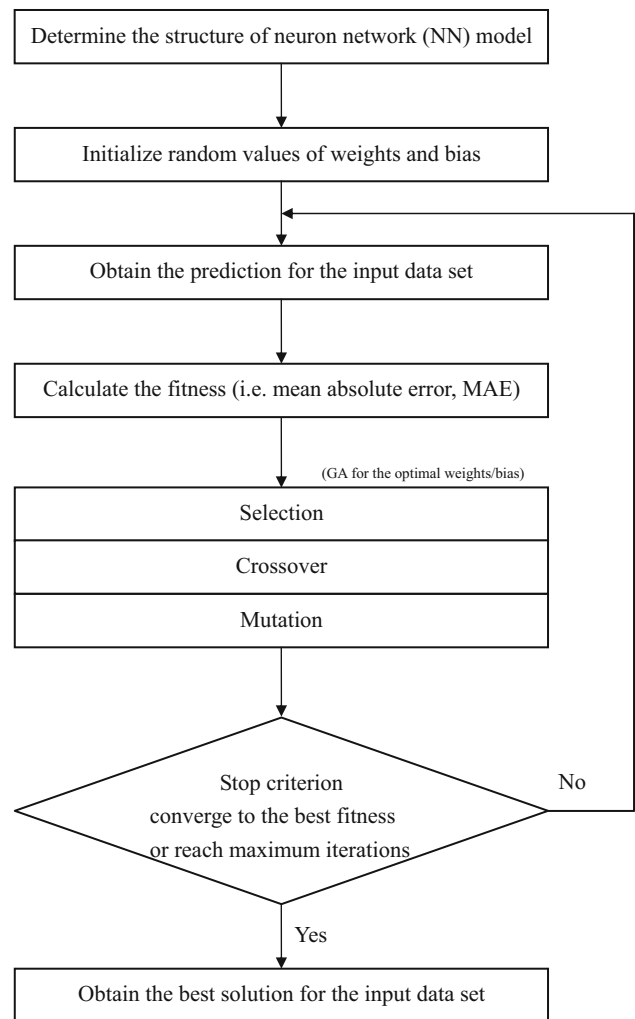


Fig. 4 Flowchart of the GANN model

HEC-HMS–GANN hybrid model can be applied for the 1-h-ahead prediction. For a different lead time, a recursive forecasting process is applied, i.e., the input node (for the measured runoff discharge at the Hsinfa Bridge gauge station) is updated from the output node (see Fig. 3).

3.3 The HEC-HMS–ANFIS hybrid model

An ANFIS is a kind of ANN based upon the fuzzy inference system (FIS), i.e., a representation of a FIS in an ANN architecture. By the combination of ANN and FIS, the ANFIS can learn the rules and membership functions from data [36]. In the multilayer feed-forward network of an ANFIS (see Fig. 5), each node performs a particular function on incoming signals. To obtain the desired input–output characteristics, the ANN learning capability (or algorithm) is utilized to determine the adaptive FIS parameters. We briefly describe the ANFIS network, parameters, and learning algorithm as follows.

Fig. 5 Architecture of the adaptive neuro-fuzzy inference system (ANFIS) model
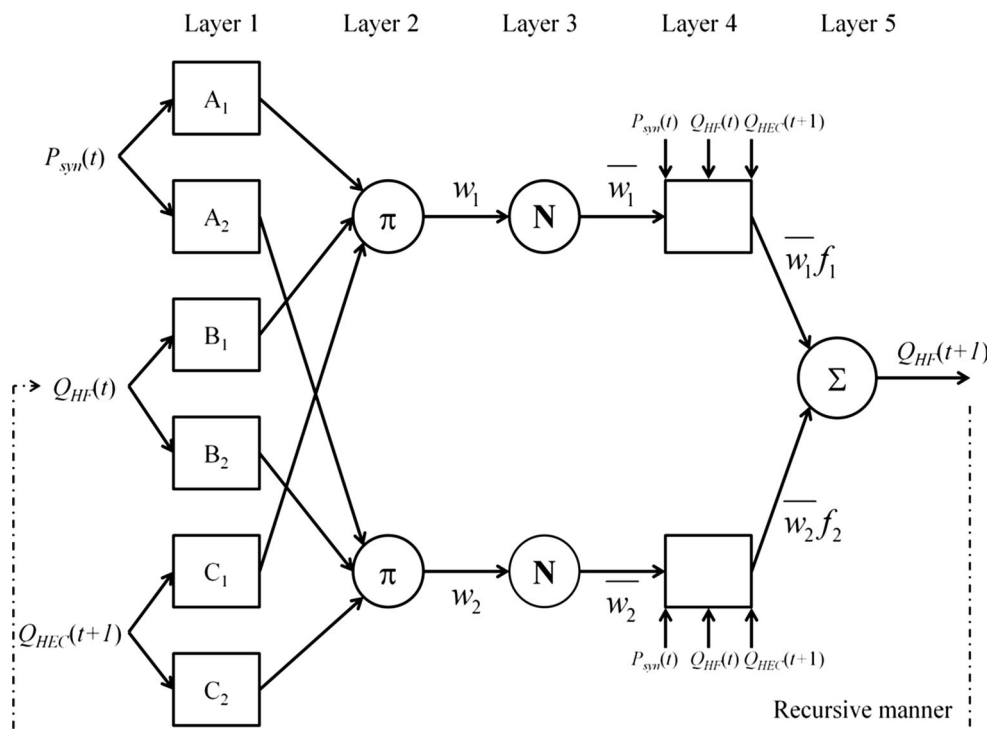
3.3.1 Network

Layer 1 The input nodes generate grades for the inputs based on the membership functions of the fuzzy sets

Layer 2 The rule nodes apply the AND or the OR operator to yield the firing strength for the antecedent part in the rule

Layer 3 The average nodes calculate the normalized ratio of the firing strength between the ith and all rules

Layer 4 The consequent nodes compute the contribution of each rule using the first-order Sugeno fuzzy model

Layer 5 The single output node sums all the incoming signals and computes the overall output according to the defuzzification process

3.3.2 Parameters and learning algorithm

The ANFIS mainly include the premise and consequent parameters. The former describes the shape of the membership functions, and the latter addresses the overall system output. The learning algorithm contains the gradient descent and least-squares methods for adjusting the premise and consequent parameters, respectively. One can further refer to Jang [36] and Nayak et al. [37] for the mathematical details of these algorithms.

In this study, we also construct the HEC-HMS–ANFIS hybrid model to forecast the hourly runoff discharge.

Similarly, we use three inputs and one output (see Fig. 5), i.e., the relation of $Q_{HF}(t + 1)$ to $P_{syn}(t)$, $Q_{HF}(t)$, and $Q_{HEC}(t + 1)$. For the model setting, two membership functions are applied for each input. Types of membership function are selected as Gaussian (or bell-shaped) and linear for inputs and output, respectively. Also, a subtractive fuzzy clustering that can automatically generate fuzzy inference systems by detecting clusters in input–output training data is applied to establish the rule-based relationship. The model is developed using the fuzzy logic toolbox of MATLAB [38].

3.4 Indices of simulation performance

To evaluate the performance of the HEC-HMS model, the HEC-HMS–GANN hybrid model, and the HEC-HMS–ANFIS hybrid model, six different criteria were used to compare the predicted results with the observed data, including the MAE, the RMSE, the coefficient of correlation (R), the error of time to peak discharge (ET_p), the error of peak discharge (EQ_p), and the CE, according to the following equations:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |(Q_m)_i - (Q_o)_i| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [(Q_m)_i - (Q_o)_i]^2} \tag{5}$$
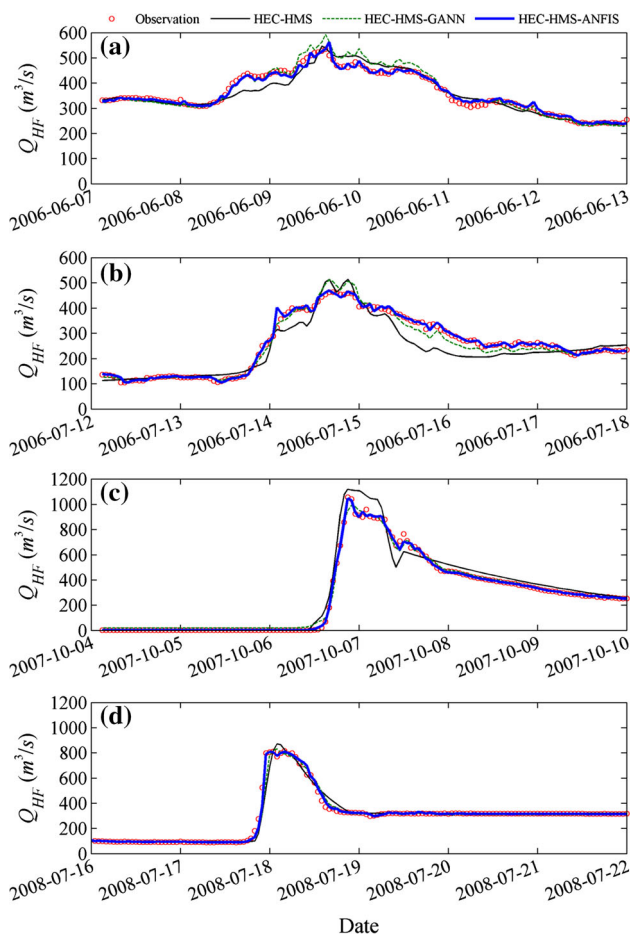
Fig. 6 Comparison of the observed and simulated runoff discharges at the Hsinfa Bridge gauge station during **a** June 9 Flood, **b** Typhoon Bilis, **c** Typhoon Krosa, and **d** Typhoon Kamegi for model calibration (training): observation (*symbols*), HEC-HMS (*black and thin solid lines*), HEC-HMS–GANN (*green and thin dashed lines*), and HEC-HMS–ANFIS (*blue and thick solid lines*)
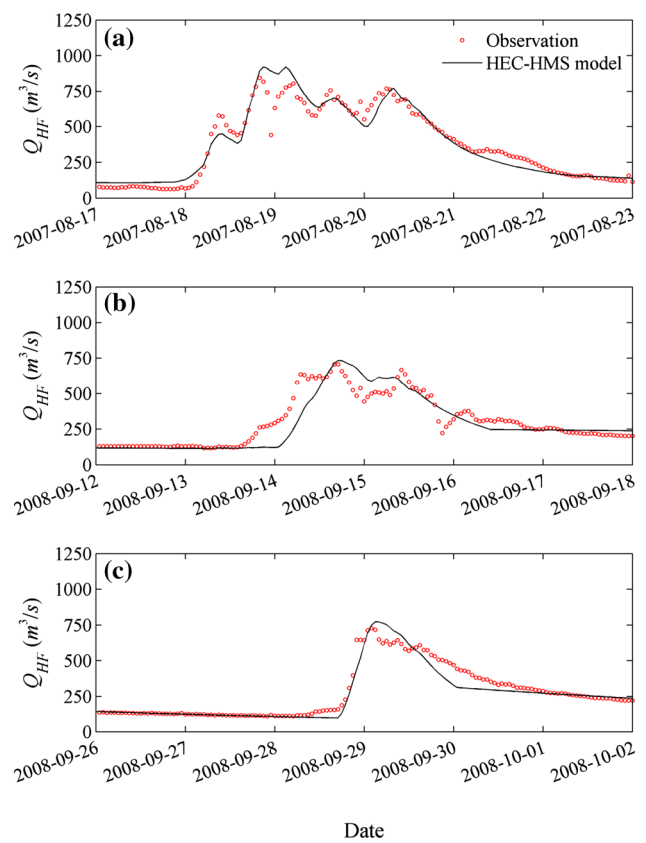

Fig. 7 Comparison of the observed and simulated runoff discharges at the Hsinfa Bridge gauge station for model validation using the HEC-HMS model. **a** Typhoon Sepat, **b** Typhoon Sinlaku, and **c** Typhoon Jangmi

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[(Q_\mathrm{m})_i - \overline{Q_\mathrm{m}}\right]\left[(Q_\mathrm{o})_i - \overline{Q_\mathrm{o}}\right]}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(Q_\mathrm{m})_i - \overline{Q_\mathrm{m}}\right]^2 \times \frac{1}{N}\sum_{i=1}^{N}\left[(Q_\mathrm{o})_i - \overline{Q_\mathrm{o}}\right]^2}} \tag{6}$$

$$ET_\mathrm{p} = \left|T_\mathrm{m,p} - T_\mathrm{o,p}\right| \tag{7}$$

$$EQ_\mathrm{p} = \frac{\left|Q_\mathrm{m,p} - Q_\mathrm{o,p}\right|}{Q_\mathrm{o,p}} \times 100\,\% \tag{8}$$

$$CE = 1 - \sum_{i=1}^{N}\left[(Q_\mathrm{o})_i - (Q_\mathrm{m})_i\right]^2 / \sum_{i=1}^{N}\left[(Q_\mathrm{o})_i - \overline{Q_\mathrm{o}}\right]^2 \tag{9}$$

where N is the total number of data points, Q_m is the predicted runoff discharge, Q_o is the observed runoff discharge, $T_\mathrm{m,p}$ and $T_\mathrm{o,p}$ are the peak time for the predicted peak runoff discharge and the observed peak runoff discharge, respectively, $Q_\mathrm{m,p}$ and $Q_\mathrm{o,p}$ are the predicted and observed peak runoff discharges, respectively; $\overline{Q_\mathrm{m}} = \frac{1}{N}\sum_{i=1}^{N}(Q_\mathrm{m})_i$; and $\overline{Q_o} = \frac{1}{N}\sum_{i=1}^{N}(Q_o)_i$.

4 Results and discussion

In this study, we apply three different approaches to predict runoff discharge in the Laonong Creek basin in southern Taiwan using the measured rainfall data during the historical typhoon events (or numerical weather predictions when considering the real operations). Notice that the original HEC-HMS model is not modified. In the hybrid models, the main idea is to treat the HEC-HMS simulation as a predictor, similar to the commonly used predictor–corrector concept in numerical modeling (e.g., see the cited references in Young and Wu [39]). Due to the operational nature of this work, the prediction lead time is further extended (from $t + 1$ to $t + 6$) based upon a recursive manner. In the following, the prediction results from the HEC-HMS and two hybrid models are presented and discussed.

Table 1 Performance assessment of runoff discharge prediction for model calibration

Typhoon event	Model	Statistical parameter					
		MAE (m³/s)	RMSE (m³/s)	R	ET_P (h)	EQ_P (%)	CE
0609 flood	HEC-HMS	17.71	24.35	0.957	1	2.34	0.908
	HEC-HMS–GANN	15.52	20.37	0.989	0	10.91	0.935
	HEC-HMS–ANFIS	7.35	10.66	0.991	1	5.39	0.982
Bilis	HEC-HMS	32.63	42.65	0.938	0	10.79	0.846
	HEC-HMS–GANN	14.93	19.75	0.987	0	10.48	0.967
	HEC-HMS–ANFIS	6.42	10.49	0.996	0	1.27	0.991
Krosa	HEC-HMS	37.29	63.04	0.984	0	5.96	0.955
	HEC-HMS–GANN	18.06	25.12	0.997	1	6.21	0.993
	HEC-HMS–ANFIS	5.47	12.28	0.999	0	0.93	0.998
Kamegi	HEC-HMS	12.9	36.33	0.980	2	6.71	0.961
	HEC-HMS–GANN	11.79	28.55	0.988	2	1.82	0.976
	HEC-HMS–ANFIS	5.96	16.5	0.996	0	0.85	0.992

Table 2 Performance assessment of runoff discharge prediction for model validation

| Typhoon event	Model	Lead time (h)	Statistical parameter					
			MAE (m³/s)	RMSE (m³/s)	R	ET_P (h)	EQ_P (%)	CE
Sepat	HEC-HMS	–	42.15	62.96	0.968	1	9.16	0.933
	HEC-HMS–GANN	1	19.58	33.22	0.991	1	2.51	0.981
		2	27.00	39.20	0.987	0	2.20	0.974
		4	36.97	49.95	0.979	0	1.39	0.958
		6	41.01	54.71	0.975	0	1.27	0.950
	HEC-HMS–ANFIS	1	18.74	37.70	0.989	0	8.75	0.976
		2	23.47	43.63	0.985	0	16.73	0.968
		4	32.30	55.83	0.975	0	22.53	0.948
		6	37.76	63.81	0.967	0	23.91	0.932
Sinlaku	HEC-HMS	–	42.25	68.82	0.915	1	3.48	0.819
	HEC-HMS–GANN	1	16.76	28.31	0.986	1	0.77	0.970
		2	23.76	37.18	0.976	1	1.17	0.947
		4	31.93	48.03	0.961	1	1.49	0.912
		6	35.16	52.52	0.954	0	1.91	0.895
	HEC-HMS–ANFIS	1	13.41	27.45	0.987	1	7.15	0.971
		2	16.52	32.67	0.982	0	7.69	0.959
		4	21.88	40.44	0.970	0	7.68	0.938
		6	26.30	47.43	0.959	0	7.68	0.915
Jangmi	HEC-HMS	–	25.54	45.26	0.963	1	5.97	0.923
	HEC-HMS–GANN	1	11.29	20.48	0.993	1	0.87	0.984
		2	18.42	28.73	0.987	0	1.33	0.969
		4	26.57	41.24	0.973	0	2.19	0.936
		6	29.95	47.13	0.965	0	2.40	0.917
	HEC-HMS–ANFIS	1	12.55	28.08	0.988	1	8.16	0.971
		2	16.26	34.51	0.980	1	7.61	0.955
		4	21.67	42.29	0.969	1	7.62	0.933
		6	26.13	48.19	0.959	1	7.62	0.913
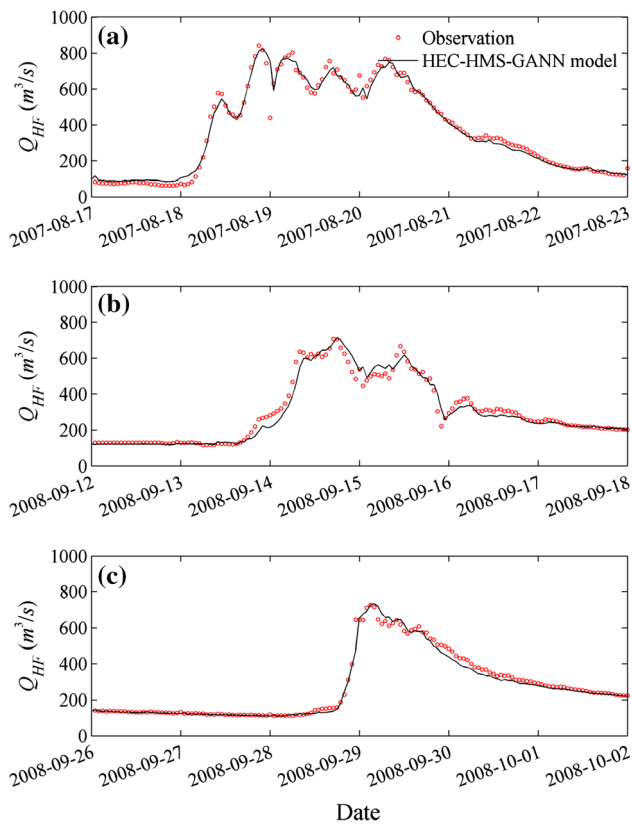
Fig. 8 Comparison of the observed and simulated runoff discharge at the Hsinfa Bridge gauge station for the validation phase using the HEC-HMS–GANN model. **a** Typhoon Sepat, **b** Typhoon Sinlaku, and **c** Typhoon Jangmi

4.1 Event-based rainfall–runoff modeling using HEC-HMS

For the HEC-HMS model application in the Laonong Creek basin, hourly rainfall and runoff discharge data were collected for model calibration and verification. Data during four heavy rainfall events—the June 9 Flood (2006), Typhoon Bilis (2006), Typhoon Krosa (2007), and Typhoon Kamegi (2008)—were used for model calibration. Data during three typhoon events—Typhoon Sepat (2007), Typhoon Sinlaku (2008), and Typhoon Jangmi (2008)—were used for model validation.

Based on the measured rainfall and discharge data, a 1-h time step was selected for use in the event hydrological modeling. The model calibration and validation results for the HEC-HMS model are shown in Figs. 6 and 7, respectively. Further, Tables 1 and 2 present their performance assessment. Overall, the model predicts the runoff reasonably well during the calibration phase. For example, the simulation of Typhoon Kamegi gives MAE = 12.9 m^3/s, RMSE = 36.33 m^3/s, R = 0.980, ET$_p$ = 2 h, EQ$_p$ = 6.71 %, and CE = 0.961. In terms of the validation events, the observed flow discharge cannot be totally captured. The
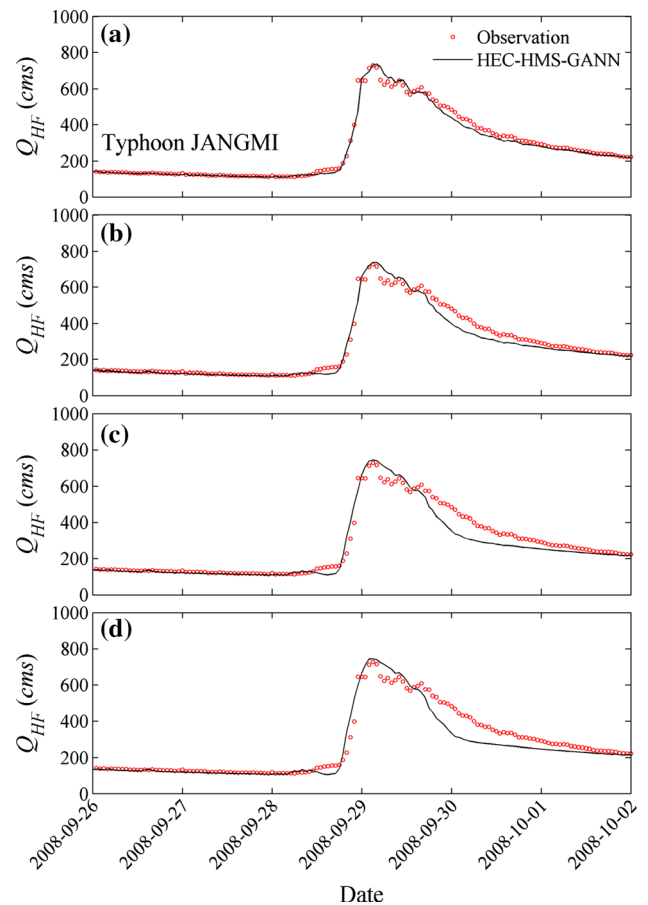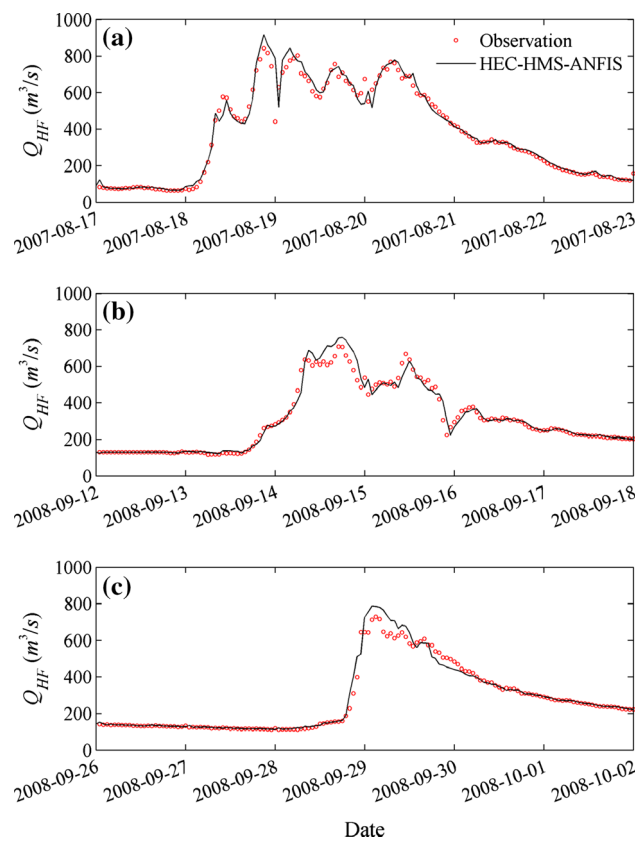

Fig. 9 Comparison of the observed and simulated runoff discharge at the Hsinfa Bridge gauge station during Typhoon Jangmi using the HEC-HMS–GANN model with **a** 1-h-, **b** 2-h-, **c** 4-h-, and **d** 6-h-ahead predictions

HEC-HMS model notably fails in predicting the runoff discharge pattern during Typhoon Sinlaku, although the predicted time and discharge of peak flow are acceptable. The MAE, RMSE, R, ET$_p$, and EQ$_p$, and CE values under the worst prediction are 42.45 m^3/s, 68.82 m^3/s, 0.915, 1 h, 3.48 %, and 0.819, respectively.

4.2 Runoff prediction using modified/hybrid models

4.2.1 The HEC-HMS–GANN model prediction

Figures 6 and 8 show the runoff discharge predicted by the HEC-HMS–GANN hybrid model. For both training and validation phases, the 1-h-ahead forecasting results are in good agreement with the observation data. The predictions of the HEC-HMS model are apparently improved within the presented combination framework. Tables 1 and 2 provide the model performance assessment. An excellent training of the hybrid model has been achieved according to the statistical parameters. For model validation phase,
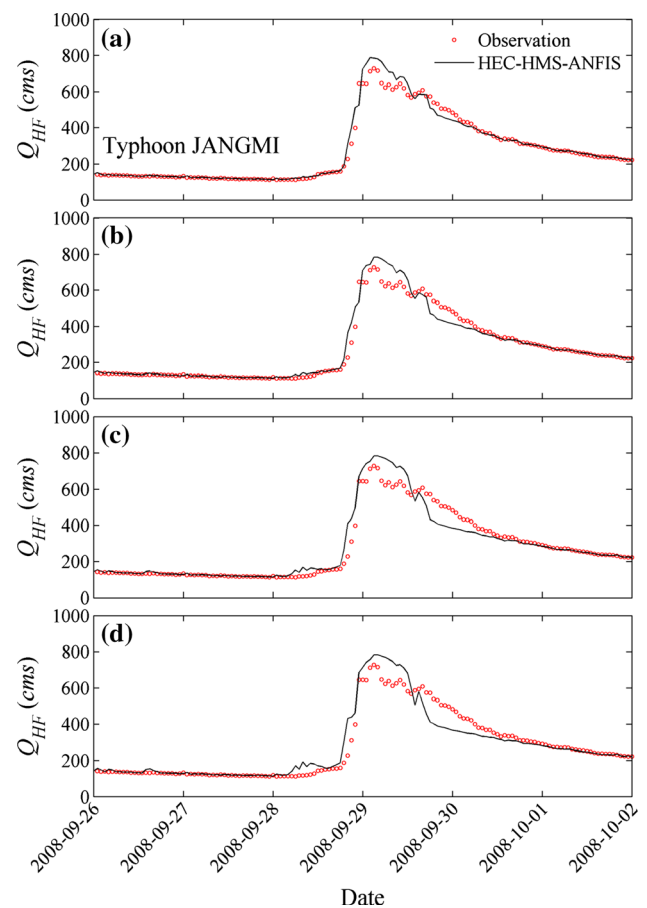
Fig. 10 Comparison of the observed and simulated runoff discharge at the Hsinfa Bridge gauge station for the validation phase using the HEC-HMS–ANFIS model. **a** Typhoon Sepat, **b** Typhoon Sinlaku, and **c** Typhoon Jangmi

Fig. 11 Comparison of the observed and simulated runoff discharge at the Hsinfa Bridge gauge station during Typhoon Jangmi using the HEC-HMS–ANFIS model with **a** 1-h-, **b** 2-h-, **c** 4-h-, and **d** 6-h-ahead predictions

the 1-h-ahead predictions yield MAE, RMSE, R, ET_p, EQ_p, and CE in the ranges of 11.29–19.58 m³/s, 20.48–33.22 m³/s, 0.986–0.993, 0–1 h, 0.77–2.51 %, and 0.970–0.984, respectively.

Further, the n-h-ahead predictions based on a recursive manner were carried out to examine the capability of the trained HEC-HMS–GANN model. Figure 9 compares the 1-h-, 2-h-, 4-h-, and 6-h-ahead runoff predictions with the measured data during Typhoon Jangmi. Overall, the predicted patterns remain quite close to the measured data, capturing discharge and time of the peak flow. The performance assessment in Table 2 also reveals that the accuracy is influenced by the lead time because there is some accumulated error. As the lead time increases, the performance gets worse. For the 6-h-ahead runoff prediction during Typhoon Jangmi, the MAE, RMSE, R, ET_p, EQ_p, and CE values are 29.95 m³/s, 47.13 m³/s, 0.965, 0 h, 2.40 %, and 0.917, respectively.
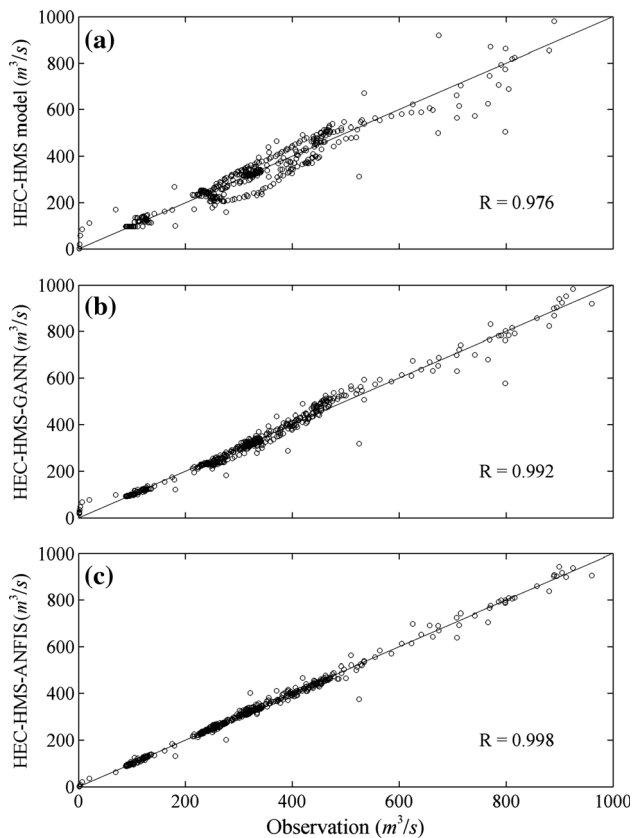
4.2.2 The HEC-HMS–ANFIS model prediction

Figures 6 and 10 also compare the measured and predicted runoff discharge for the HEC-HMS–ANFIS hybrid model.

The 1-h-ahead predictions mimic the observations in the training and validation phases. Tables 1 and 2 show the performance assessment. The overall training results are better than those obtained from the HEC-HMS–GANN hybrid model. For the most concerned validation procedure, the MAE, RMSE, R, ET_p, EQ_p, and CE parameters are about 12.55–18.74 m³/s, 27.45–37.70 m³/s, 0.987–0.989, 0–1 h, 7.15–8.75 %, and 0.971–0.976 in the 1-h-ahead prediction, respectively.
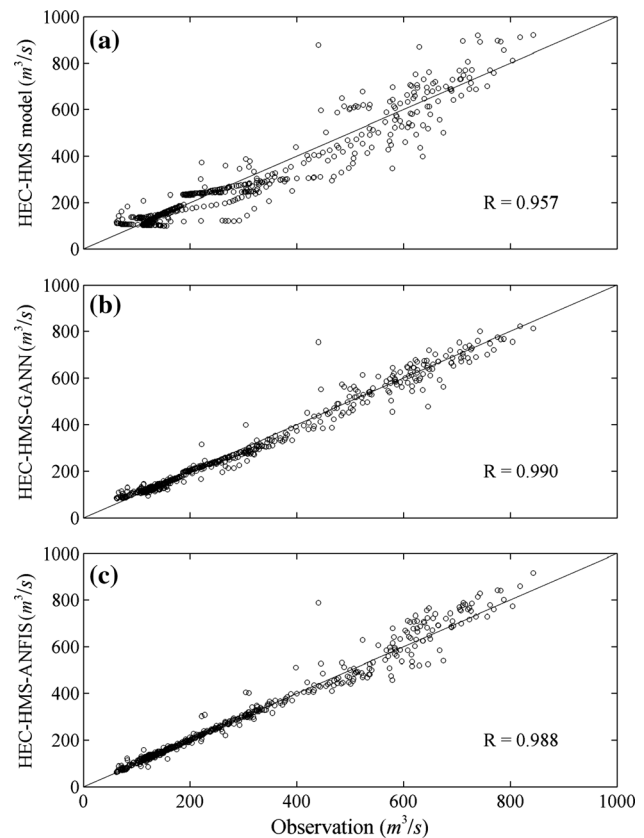
Similarly, the n-h-ahead predictions based upon the recursive manner (shown in Fig. 5) were conducted using the HEC-HMS–ANFIS model. Figure 11 compares the 1-h-, 2-h-, 4-h-, and 6-h-ahead runoff predictions with the measured data during Typhoon Jangmi. The predicted patterns remain quite close to the measured data. However, the peak discharge is slightly overestimated in the simulation. For the 6-h-ahead runoff prediction during Typhoon Jangmi, the MAE, RMSE, R, ET_p, EQ_p, and CE values are 26.13 m³/s, 48.19 m³/s, 0.959, 1 h, 7.62 %, and 0.913, respectively. Likewise, the performance gets worse as the lead time increases.

Fig. 12 Scatter plots of predicted and measured runoff discharges for the model calibration/training. **a** HEC-HMS model, **b** HEC-HMS–GANN model, and **c** HEC-HMS–ANFIS model

Fig. 13 Scatter plots of predicted and measured runoff discharges for the model validation. **a** HEC-HMS model, **b** HEC-HMS–GANN model, and **c** HEC-HMS–ANFIS model

4.3 Comparison of the runoff discharge predictions

To evaluate the prediction performance of these three modeling approaches, the observation–prediction pairs of hourly runoff discharge during the calibration (training) and validation phases are presented using scatter plots in Figs. 12 and 13, respectively. The HEC-HMS model did not yield a good prediction of the runoff discharge at the outlet of the watershed (Figs. 12a, 13a). The 1-h-ahead runoff discharge predictions yielded by the HEC-HMS–GANN model (Figs. 12b, 13b) and HEC-HMS–ANFIS model (Figs. 12c, 13c) are superior to that obtained with the HEC-HMS model. Figures 14 and 15 further compare the scatter plots for the *n*-h-ahead forecasting results from the HEC-HMS–GANN and HEC-HMS–ANFIS models, respectively. The two hybrid models both exhibit accumulated errors as the lead time increases (see the statistical parameters RMSE and *R* in Table 2). Notice that the detailed comparison raises an interesting point for a deeper discussion. While these two models are well trained, the results for ANFIS are overall better than GANN. However, the fuzzy logic approach (ANFIS) and genetic algorithm (GANN) show quite different generalization capability in the validation phase. The HEC-HMS–ANFIS model tends

to overpredict the peak flow and to underpredict the falling limb of the hydrograph (see Figs. 11d, 15c). By contrast, the HEC-HMS–GANN model can capture the peak flow with a similar underestimation of the recession (see Figs. 9d, 14c). For the EQ_p value that physically explains the error of peak flow, therefore, the performance of the HEC-HMS–GANN model is better than that of the HEC-HMS–ANFIS model.
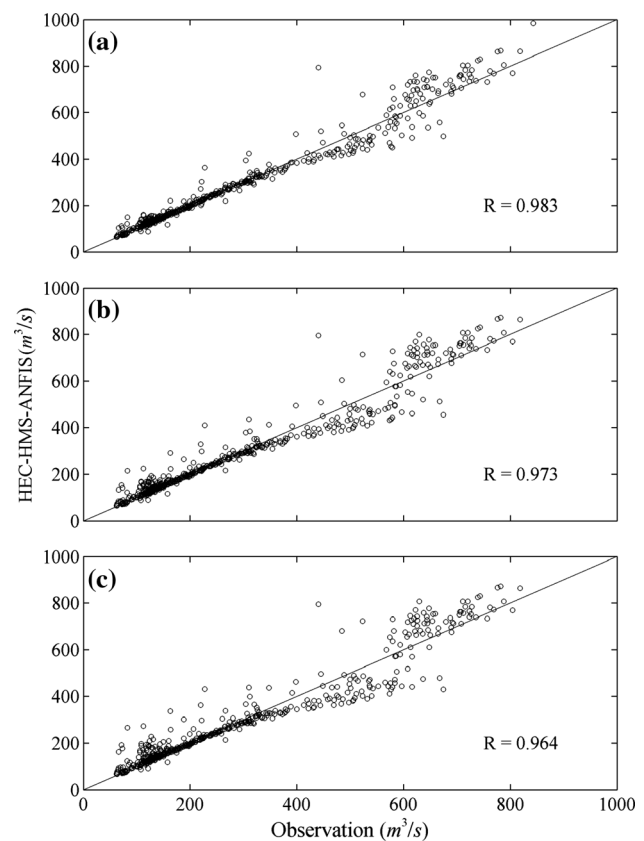
5 Conclusions

Three modeling approaches (i.e., the HEC-HMS model, the HEC-HMS–GANN hybrid model, and the HEC-HMS–ANFIS hybrid model) were used to predict runoff discharges in the Laonong Creek basin in southern Taiwan. Four heavy rainfall events were used for model calibration (training), and three typhoon events were used for model validation. The relative performance of these models was comprehensively evaluated using various statistical indices (i.e., MAE, RMSE, *R*, ET_p, EQ_p, and CE).

The calibrated and validated results indicated that the HEC-HMS model cannot satisfactorily predict runoff
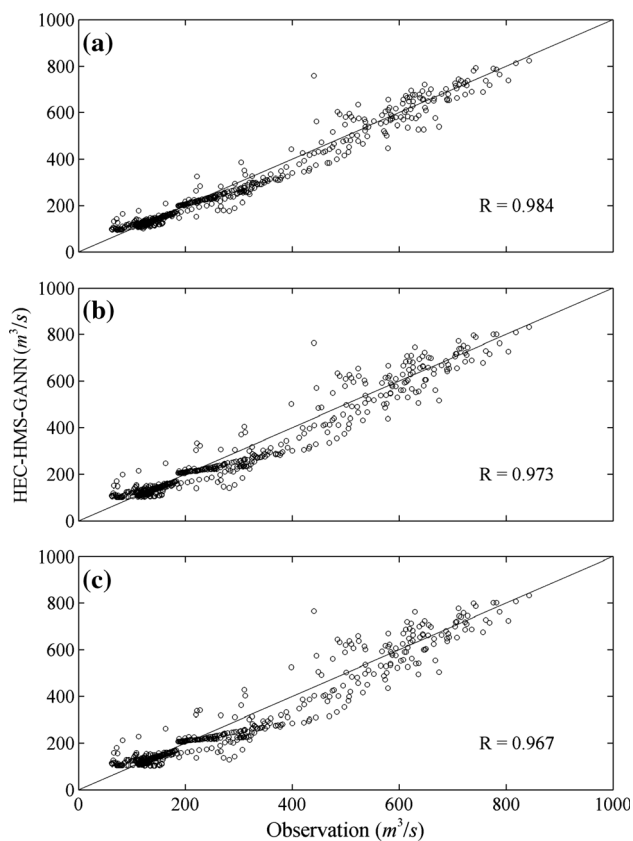
Fig. 14 Scatter plots of predicted and measured runoff discharges for the validation of HEC-HMS–GANN model. **a** 2-h-, **b** 4-h-, and **c** 6-h-ahead predictions

Fig. 15 Scatter plots of predicted and measured runoff discharges for the validation of HEC-HMS–ANFIS model. **a** 2-h-, **b** 4-h-, and **c** 6-h-ahead predictions

discharges during typhoon events. Two novel alternative approaches combining the physically based model (i.e., HEC-HMS) and the black-box models (i.e., GANN and ANFIS) were proposed to predict the runoff discharge with improved accuracy. The results indicated that the 1-h-ahead prediction obtained by the HEC-HMS–GANN and HEC-HMS–ANFIS hybrid models was better than that from HEC-HMS model. It demonstrates that the combination of physically based and black-box models can yield the excellent prediction results of runoff discharge during the typhoon events. When the forecasting lead time increases, the performance of the HEC-HMS–GANN and HEC-HMS–ANFIS models gets worse. We also found that the HEC-HMS–GANN and HEC-HMS–ANFIS models show quite similar performance in n-h-ahead predictions of runoff discharge according to statistical indices RMSE and R values. In terms of the EQ_p value, the performance of the

HEC-HMS–GANN model is generally better than that of the HEC-HMS–ANFIS model.

References

1. Wang W, Ding J (2003) Wavelet network model and its application to the prediction of hydrology. Nat Sci 1(1):67–71
2. Arabi M, Frankenberger JR, Engel BA, Arnold JG (2008) Representation of agricultural practices with SWAT. Hydrol Process 22(16):3042–3055
3. Pektas AO, Cigizoglu HK (2013) ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient. J Hydrol 500:21–36

4. Joo J, Kjeldsen T, Kim HJ (2014) A comparison of two event-based flood models (ReFH-rainfall runoff model and HEC-HMS) at two Korean catchments, Bukil and Jeungpyeong). KSCE J Civ Eng 18(1):330–343

5. Dawson CW, Wilby RL (2001) Hydrological modelling using artificial neural networks. Prog Phys Geogr 25(1):80–108

6. Abrahart RJ, Kneale PE, See L (2004) Neural networks for hydrological modelling. Taylor & Francis, New York, p 304

7. Tabari H, Marifi S, Abyaneh HZ, Sharifi MR (2010) Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in Samsami basin of Iran. Neural Comput Appl 19(4):625–635

8. Talei A, Chua LHC, Quek C, Jansson PE (2013) Runoff forecasting using a Takagi–Sugeno neuro-fuzzy model with online learning. J Hydrol 488:17–32

9. Machado F, Mine M, Kaviski E, Fill H (2011) Monthly rainfall–runoff modelling using artificial neural networks. Hydrol Sci J 56(3):349–361

10. Nourani V, Komasi M, Alami MT (2012) Hybrid wavelet-genetic programming approach to optimize ANN modeling of rainfall–runoff process. J Hydrol Eng ASCE 17(6):724–741

11. Rezaeianzadeh M, Tabari H, Yazdi AA, Isik S, Kalin L (2014) Flood flow forecasting using ANN, ANFIS and regression models. Neural Comput Appl 25(1):25–37

12. Rezaeianzadeh M, Stein A, Tabari H, Abghari H, Jalalkamali N, Hosseinipour EZ, Singh VP (2013) Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. Int J Environ Sci Technol 10(6):1181–1192

13. Shamseldin AY, O'Connor KM (2001) A non-linear neural network technique for updating of river flow forecasts. Hydrol Earth Syst Sci 5(4):577–597

14. Xiong L, O'Connor KM (2002) Comparison of four updating models for real-time river flow forecasting. Hydrol Sci J 47(4):621–639

15. Shamseldin AY (1997) Application of a neural network technique to rainfall–runoff modelling. J Hydrol 199(3–4):272–294

16. See L, Abrahart RJ (2001) Multi-model data fusion for hydrological forecasting. Comput Geosci 27(8):987–994

17. Ajami NK, Duan Q, Gao X, Sorooshian S (2006) Multimodel combination techniques for analysis of hydrological simulations: application to distributed model intercomparison project results. J Hydrometeorol 7(4):755–768

18. Shamseldin AY, O'Connor KM, Nasr AE (2007) A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall–runoff models. Hydrol Sci J 52(5):896–916

19. Viney NR, Bormann H, Breuer L, Croke BFW, Frede H, Graff T, Hubrechts L, Huisman JA, Jakeman AJ, Kite GW, Lanini J, Leavesley G, Lettenmaier DP, Lindstrom G, Seibert J, Sivapalan M, Willems P (2009) Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). II: ensemble combinations and predictions. Adv Water Resour 32(2):147–158

20. Fernando AK, Shamseldin AY, Abrahart RJ (2012) Use of gene expression programming for multimodel combination of rainfall–runoff models. J Hydrol Eng 17(9):975–985

21. Phukoetphim P, Shamseldin AY, Melville BW (2014) Knowledge extraction from artificial neural networks for rainfall–runoff model combination system. J Hydrol Eng 19(7):1422–1429

22. Panda RK, Pramanik N, Bala B (2010) Simulation of river stage using artificial neural network and MIKE 11 hydrodynamic model. Comput Geosci 36(6):735–745

23. Napolitano G, See L, Calvo B, Savi F, Heppenstall A (2010) A conceptual and neural network model for real-time flood forecasting of the Tiber River in Rome. Phys Chem Earth 35(3–5):187–194

24. Talei A, Chua LHC, Wang TSW (2010) Evaluation of rainfall and discharge inputs used by adaptive network-based fuzzy inference systems (ANFIS) in rainfall–runoff modeling. J Hydrol 391(3–4):248–262

25. Young CC, Liu WC (2015) Prediction and modelling of rainfall–runoff during typhoon events using a physically-based and artificial neural network hybrid model. Hydrol Sci J (in press)

26. Feldman AD (2000) Hydrologic modeling system HEC-HMS. Technical reference manual. U.S. Army Corps of Engineers, Davis

27. US Army Corps Engineers (2008) Hydrologic modeling system (HEC-HMS) application guide: version 3.1.0. Institute for Water Resources, Davis

28. US Soil Conservation Service (1986) Urban hydrology for small watershed. Technical release 55. US Department of Agriculture

29. McCuen RH (1998) Hydrologic analysis and design. Prentice-Hall, Upper Saddle River, pp 155–163

30. Clark CO (1945) Storage and the unite hydrograph. Trans Am Soc Civil Eng 110:1419–1488

31. Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw-Hill, New York

32. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Longman, New York

33. Karimi H, Yousefi F (2012) Application of artificial neural network-genetic algorithm (ANN-GA) to correlation of density in nanofluids. Fluid Phase Equilib 336:79–83

34. Ferentinos KP (2005) Biological engineering applications of feed forward neural networks designed and parameterized by genetic algorithms. Neural Netw 18(7):934–950

35. Karzynski M, Mateos A, Herrero J, Dopazo J (2003) Using a genetic algorithm and a perceptron for feature selection and supervised class learning in DNA microarray data. Artif Intell Rev 20(1–2):39–51

36. Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685

37. Nayak PC, Sudheer KP, Ragan DM, Ramasastri KS (2004) A neuro fuzzy computing technique for modeling hydrological time series. J Hydrol 291(1–20):52–66

38. MATLAB (2008) Fuzzy logic toolbox-user guide. The MathWorks, MA, USA

39. Young CC, Wu CH (2009) An efficient and accurate non-hydrostatic model with embedded Boussinesq-type like equations for surface wave modeling. Int J Numer Meth Fluids 60(1):27–53