Decision Tree–Based Classifier Combined with Neural-Based Predictor for Water-Stage Forecasts in a River Basin During Typhoons: A Case Study in Taiwan

Chia-Cheng Tsai,¹ Mi-Cheng Lu,² and Chih-Chiang Wei^{2,*}

¹Department of Marine Environmental Engineering, National Kaohsiung Marine University, Kaohsiung, Taiwan.

²Department of Information Management, Toko University, Pu-Tzu, Chia-Yi, Taiwan.

Received: April 26, 2011

Accepted in revised form: July 2, 2011

Abstract

To solve the complicated problem of water-stage predictions under the interaction of upstream flows and tidal effects during typhoon attacks, this article presents a novel approach to river-stage predictions. The proposed CART-ANN model combines both the decision trees (classification and regression trees [CART]) and the artificial neural network (ANN) techniques, which comprise the multilayer perceptron (MLP) and radial basis function (RBFNN). The combined CART-ANN model involves a two-step predicting process. First, the CART stage-level classifier can classify the river stages into higher, middle, and lower levels. Then, the ANN-based water-stage predictors are employed to predict the water stages. The proposed model was applied to the Tanshui River Basin in Taiwan. The Taipei Bridge, which is close to the estuary and affected by tidal effects, was taken as the study gauge. The mean square error and the mean absolute error were used for evaluating the variance and bias performances of the models. This study makes two contributions. First, the CART-MLP and CART-RBF were modeled to predict river stages under tidal effects during typhoons, and they were compared with three benchmark models, CART, back-propagation neural network, and RBFNN. Second, the CART-RBF successfully demonstrated that it achieved more accurate prediction than CART-MLP and three benchmark models.

Key words: water stage; prediction; typhoon; decision tree; neural network

Introduction

TYPHOONS, ALSO KNOWN AS tropical cyclones, often attack the western North Pacific region. Enormous flood damage in Taiwan is frequently caused by typhoons between May and October every year. Such flood hazards cause considerable economic losses and casualties (Hsu and Wei, 2007; Wei and Hsu, 2009). Therefore, accurate water-stage prediction is important to enable the authorities concerned to issue a forewarning of the impending flood and to implement early evacuation measures when required (Chau, 2006).

Tanshui River Basin, as can be seen in Fig. 1, is located in the northern part of Taiwan having steep terrain and excessive rainfall. In this basin, there is a population of over 6 million people with a drainage area of 2726 km². The most severe disaster in Tanshui River Basin is flood, which is caused by intensive rainfall and rapid flows during typhoon seasons (Wei and Hsu, 2008). Because the Taipei Bridge station is lo-

cated at the political and economic center of Taipei Metropolitan, it is chosen as the prediction target. As seen in Fig. 1, the studied stage (i.e., Taipei Bridge) is close to the estuary and strongly affected by tides. Therefore, the water level at this site is influenced by natural processing of rainfall runoff, control flow by two upstream reservoirs (i.e., Shihmen and Feitsui Reservoirs), and vibrations of the water body due to the sea tides at the same time. The intense interactions between upstream flows and tidal effects during typhoons simultaneously affect the stages of a gauge. During typhoon attacks, higher water stages might occur under a special combination of the above three factors (i.e., rainfall, runoff, and sea tide). In other words, a peak stage may appear under a typhoon surge, the maximal rainfall rate in a watershed, and/or maximum releases from reservoirs.

The major challenge of solving the problems of river-level forecasts under tidal effects is how to improve the accuracy of prediction. This study presented a CART-ANN model that combines both the decision trees (classification and regression trees [CART]) and the two artificial neural network (ANN) techniques (i.e., the multilayer perceptron [MLP] and radial basis function [RBFNN]) to conquer the accuracy problem. The two CART-ANN models (i.e., CART-MLP and CART-RBF)

*Corresponding author: Department of Information Management, Toko University, No. 51, Sec. 2, University Road, Pu-Tzu, Chia-Yi 61363, Taiwan. Phone: +886 5 3622889; Fax: +886 5 3622899; E-mail: d89521007@ntu.edu.tw

FIG. 1. Map of great Tanshui River Basin.

involve a two-step approach that can be employed to address the complicated problem of water-stage forecasts. In the first step, the river stages are classified into three levels, namely higher (HL), middle (ML), and lower (LL) levels. Second, the ANN-based water-stage predictors are separately trained to predict the water levels. In this study, the forecast performance of the CART-ANN models are compared with the results obtained from three benchmarks, the conventional CART, back-propagation neural network (BPNN), and RBFNN. The proposed model was applied to the Tanshui River Basin in Taiwan.

Literature Reviews

In the fields of hydrology and water resource engineering, the ANN has been successfully applied to the prediction of rainfall, rainfall runoff, and river stage (Minns and Hall, 1996; Liong *et al.*, 2000; Maier and Dandy, 2000; Deka and Chandramouli, 2005; Park *et al.*, 2005; Wang *et al.*, 2006; Boughton, 2007; Kisi, 2008; Coulbaly, 2010; Alvisi and Franchini, 2011). The forecasting of river stage or streamflow by machine learning has currently received much attention.

Thirumalaiah and Deo (1998) depicted the use of a conjugate gradient ANN in real-time forecasting of water stages. Atiya *et al.* (1999) and Deka and Chandramouli (2005) applied neural networks to hydrologic flow routing. Cheng *et al.* (2002) presented an automatic rainfall-runoff model calibration based on genetic algorithm and fuzzy optimal model. Chang and Chen (2003) employed RBFNN for estuary water-stage forecasting. Bazartseren *et al.* (2003) showed that both the ANN and neuro-fuzzy systems outperformed the linear statistical models for predictions of daily water stages. Chau *et al.* (2005) compared

the linear regression, the genetic algorithm-based ANN, and the adaptive network-based fuzzy inference system for flood forecasting. Lin *et al.* (2006) employed the support vector machines for long-term hydrological prediction. Chau (2006) studied the predictions of river stage by a particle swarm optimization algorithm to improve the convergence rate. Cheng *et al.* (2008) presented a model based on the chaos optimization algorithm and genetic algorithm to address the hydropower reservoir operation problem. Wei and Hsu (2008) developed a linear ANN channel level routing model for predicting the water stages of gauging points that are affected by tidal effects. Wang *et al.* (2009) examined artificial intelligence techniques to develop hydrological forecasting model based on long-term observations of monthly river flow discharges. Wu *et al.* (2009) discussed the accuracy performance of monthly streamflow forecasts when using data-driven modeling techniques on the streamflow series. Hsu *et al.* (2010) proposed the flash flood routing model with the ANN predictions for stage profiles. Nourani and Kalantari (2010) developed an integrated ANN model for spatial and temporal forecasting of daily suspended sediment discharge at multiple gauging stations. Ni *et al.* (2010) presented a genetic programming technique to model the relationship between streamflow and the impact of climate change. The above studies showed that the application of machine learning techniques to water-stage or streamflow forecasts is less subjected to physical constraints. Further, it is good at identifying and learning correlated patterns between input data and the corresponding target values (Bazartseren *et al.*, 2003).

Currently, the neural tree algorithms have been widely applied in various fields (e.g., Tsujino and Nishida, 1995; Sethi and Yoo, 1997; Bhattacharya and Solomatine, 2005; Maji, 2008; Etamad-Shahidi and Mahjoobi, 2009; Su *et al.*, 2010; Selamat and Ng, 2011). This study proposes a neural-based decision tree method, which is one of the neural tree algorithms. The idea of the proposed method came from the M5 model tree algorithm put forward by Quinlan (1992) and Bhattacharya and Solomatine (2005). In the M5 algorithm, the variable space is split into subspaces by a decision tree with leaves. In each leaf there is a linear regression function, which can predict continuous numeric attributes (Bhattacharya and Solomatine, 2005).

Algorithm

The principles of CART-ANN model are described below.

CART classifier

The CART was developed by Breiman *et al.* (1984) from the information theory. It is a nonlinear discrimination method, which uses a set of variables to split a sample into progressively smaller subgroups (Michael and Gordon, 1997; Ture *et al.*, 2005). If the dependent variable is categorical, the CART produces a classification tree. When the dependent variable is continuous, it produces a regression tree (Mahjoobi and Etamad-Shahidi, 2008). The evaluation function used for splitting in the CART is the impurity GINI index (Kearns and Mansour, 1999; Tan *et al.*, 2006)

$$\text{GINI}(k) = 1 - \sum_i p_i^2 \quad (1)$$

where p_i is the probability of class i in node k .

FIG. 2. Example of splitting two-dimensional space by decision trees.

Figure 2 demonstrates a two-dimensional axis plane that is divided into five state-space regions. For each outcome, either a leaf or a decision node is assigned until all branches end in leaves of the tree (Bessler *et al.*, 2003). Further, the corresponding decision trees are depicted in Fig. 3. Generally speaking, decision trees are good at classification and easy to use (Breiman *et al.*, 1984; Apte and Weiss, 1997).

CART coupled with ANN predictor

The CART-ANN is used for inducing a model tree. Figure 3 depicts the concept of combining both CART trees and MLP/RBF networks. The CART-ANN model uses the idea of splitting the parameter space into areas by the CART and builds in each of the area a local specialized MLP/RBF neural model. The specialized ANN-based models are then employed to forecast the prediction values.

MLP neural network. The MLP is a feed-forward neural network trained with the standard BP algorithm (Kurt *et al.*, 2008). The architecture of the MLP comprises an input layer,

one or more hidden layers, an output layer, and a connection system. The numbers of hidden layers and neurons in each layer are not constant and need to be optimized. In the MLP, the output is generated by passing signals from the input through the hidden layer to the output layer (Etemad-Shahidi and Mahjoobi, 2009). The MLP model is learned based on supervised BP learning algorithm. Although BP can provide very compact distributed representations of complex data-sets, it has its disadvantages such as slow learning, large training sets, easy sticking on local minima, and poor robustness (Song *et al.*, 1997).

RBF neural network. An RBF is a special kind of neural network. It consists of three layers: an input layer, a hidden layer (also called a receptor layer), and an output layer. The first layer consists of the input nodes. The hidden layer consists of many neurons that represent clusters of input patterns, similar to the clusters in a *k*-means model (Hwang and Bang, 1997). The output nodes are simple summations. This particular architecture of RBF has been proved to directly improve the training time at the expense of requiring a few times as many connection weights (Song *et al.*, 1997).

Model Application

Study area and data

The greater Tanshui River Basin (Fig. 1) comprises three major tributaries including Tahan River, Hsintien River, and Keelung River and two main reservoirs including Shihmen Reservoir and Feitsui Reservoir. The Taipei Bridge gauge strongly affected by tides is chosen as the prediction target. This study collected the hydrological data from rain gauges, water-stage gauges, and reservoir operation authorities. The data of hourly precipitations at Chuchiuh, Wudu, Ruifang, Shiding, Chungcheng, Tatungshang, Fushan, Tabau, Shihmen, and Shanshia gauges in the basin and the records of hourly river stages at Wudu, Taipei Bridge, Tudigong, and Hekou gauges were collected from the Water Resources Agency (WRA). The records of historic hourly releases from Shihmen and Feitsui Reservoirs were obtained from the WRA and the Taipei City Government. In this study, a total of 50 typhoon events from 1996 to 2007 were collected, as shown in Table 1. The average amount of hourly rainfall at all 11 rainfall gauges represented the hourly precipitations of Tanshui River watershed (excluding the drainage area in Shihmen and Feitsui Reservoirs).

Modeling

The steps of the CART-MLP model are described as follows.

Attributes and pattern classifications. According to the raw data, this study defined the attributes that can affect the water stages in a river during typhoon attacks. Table 2 lists the six attributes, their mean values, domain ranges, and notations. From these historic typhoon events, 3,530 hourly tuples (i.e., records) were available. These refined tuples were classified into training and validation.

Case design. As mentioned earlier, the water stages at the Taipei Bridge are significantly affected by the sea tides;

FIG. 3. Schematic of classification and regression trees-artificial neural network (CART-ANN) model configuration.

TABLE 1. TYPHOON EVENTS STUDIED

Year	Names of typhoons
1996	Herb
1997	Winnie, Amber, Ivan
1998	Otto, Yanni, Zeb, Babs
1999	Maggie, Sam, Dan
2000	Kai-Tak, Billis, Prapiroon, Bopha, Yagi, Xangsane
2001	Utor, Yutu, Toraji, Nari, Lekima, Haiyan
2002	Rammasun, Nakri, Sinlaku
2003	Soudelor, Imbudo, Morakot, Vamco
2004	Conson, Mindulle, Ranim, Aere, Haima, Nock-Ten, Nanmadol
2005	Haitang, Matsa, Talim, Khanun, Longwang
2006	Billis, Kaemi, Saomai, Shanshan
2007	Pabuk, Sepat, Wipha, Krosa

hence, this study focuses on forecasting the river stages during typhoon attacks. The travel time from upstream to downstream is often < 3 h in the basin (Wei and Hsu, 2008). To identify the suitable length of lag time, this study designed three cases with lag times (D) ranging from 1 to 3 h for predicting the channel stages.

Model construction. The construction of the CART-ANN model comprises two phases, that is, training and validation. The flowchart of the training phase is depicted in Fig. 4. The training phase involves six steps as described below.

- Step 1: Rank tuples of the training dataset according to descending target values of $A_5(t+1)$, which represents the water stages at the Taipei Bridge for 1 h ahead forecast at time period t . Obtain the maximum value of $A_5(t+1)$ through the sorting processes above.
- Step 2: Divide the entire sorting dataset from Step 1 into three equal subdatasets.
- Step 3: Define the three levels of water stages for the above subdatasets. The first one-third of the dataset belongs to the HL , another one-third refers to the ML , and the remaining one-third is named as the LL .
- Step 4: Construct the CART level classifier model. From Step 3, an addition field, called $AC_5(t+1)$, is defined here as a new categorical attribute for water stages at the Taipei Bridge. It is derived from $A_5(t+1)$ and set to be HL , ML , or LL . Here, $AC_5(t+1)$ is regarded as the target

TABLE 2. CHARACTERISTICS OF HYDROLOGICAL ATTRIBUTES

Attribute/Notation	Unit	Mean	Domain range
Average rainfall/ A_1	mm/h	3.74	0–55.10
Release from Shihmen Reservoir/ A_2	cms	510.40	0–8510.36
Release from Feitsui Reservoir/ A_3	cms	242.07	0–3443.37
Water stage at Wudu/ A_4	m	7.12	3.98–19.09
Water stage at Taipei Bridge/ A_5	m	0.68	–1.36–5.79
Water stage at Hekou/ A_6	m	0.30	–1.57–2.50

FIG. 4. Flowchart of CART-ANN model construction. HL , higher level; ML , middle level; LL , lower level.

for the CART classifier model. In the model, the mapping function can be expressed as

$$AC_5(t+1) = f_{\text{CART}}\{(A_1(t-\Delta t), A_2(t-\Delta t), A_3(t-\Delta t), A_4(t-\Delta t), A_5(t-\Delta t), A_6(t-\Delta t)), \Delta t = 1, D\} \quad (2)$$

where Δt is the index of lag time and $A_i(t)$ refers to the attribute index i at time t . When constructing the CART level classifier model, the number of inputs is 6, 12, and 18 for D from 1 to 3 h.

Step 5: Construct the ANN-based water-stage predictor model. Three submodels, namely ANN_{HL} , ANN_{ML} , and ANN_{LL} , are trained using subdatasets HL , ML , and LL from Step 3, respectively. In ANN_{HL} , for example, the mapping function is expressed as

$$A_5(t+1) = f_{\text{ANN}_{HL}}\{(A_1(t-\Delta t), A_2(t-\Delta t), A_3(t-\Delta t), A_4(t-\Delta t), A_5(t-\Delta t), A_6(t-\Delta t)), \Delta t = 1, D\} \quad (3)$$

In the above equation, all attributes are of numeric type. Step 6: Integrate the CART classifier with the ANN-based predictors into the CART-ANN prediction model.

The trained CART-ANN model is validated by the four-step procedure depicted in Fig. 5.

- Step 1: Input tuples of the validation dataset into the CART level classifier model.
- Step 2: Generate the classifications of the three groups HL , ML , and LL using the CART classifiers.
- Step 3: Input the three predicted datasets of HL , ML , and LL to the ANN_{HL} , ANN_{ML} , and ANN_{LL} predictor models, respectively.
- Step 4: Obtain the forecast river stages from above step and compare with observations.

In both the training and validation phases, the cross-validation sampling approach was employed. The entire dataset

FIG. 5. Flowchart of CART-ANN model validation.

was randomly partitioned into five equal-sized subsets. This procedure was repeated five times so that each partition was used for validating exactly once.

Results and Discussion

Model parameters

The CART, MLP, and RBF are performed using the Clementine software (SPSS Inc., 2004). For CART, when using this software, the three important parameters are set, including the minimum records per child branch=1% of all records, where the sizes of subgroups can be used to limit the number of splits in any branch of the tree, the minimum change in impurity=0.00001, which creates a new split in the tree, and the maximum tree depth, which specifies the maximum number of levels below the root node and was chosen by the sensitivity analysis of tree depth in the interval of [1, 15]. The root mean square error (root MSE) was calculated and the results were plotted in the Fig. 6a. From the figure, the suitable tree depth was selected at the value of 9.

FIG. 6. Sensitivity of the parameters on the RMSE of (a) CART, (b) MLP, and (c) RBFNN models. MLP, multilayer perceptron; RBFNN, radial basis function; RMSE, root mean square error.

For MLP, parameters in the network include training cycles=2000, learning rate (α)=0.9, and sigmoid and linear activity function in the hidden layer and output layer, respectively. The number of nodes in the hidden layer was chosen by the sensitivity analysis in the interval of [1, 20]. In Fig. 6b, the suitable number of nodes occurred at the amount of 6. For RBFNN, the primary calculation in k -means is an iterative process of calculating cluster centers and assigning records to clusters. The parameter of learning rate (α) is 0.9. The number of centers (clusters) is the key in network construction. By the sensitivity analysis of clusters in the interval of [0, 700], as seen in Fig. 6c, the suitable clusters were selected at an amount of 250.

Analysis from CART-ANN

During classification, the classes assigned by the CART level classifier may be inconsistent or consistent with the true class of tuples. A confusion matrix can be employed to check the misclassification situations (Bessler *et al.*, 2003). As seen from the confusion matrices in Table 3, the validation datasets were classified into *HL*, *ML*, and *LL*. In the first case (i.e., $D=1$ h), just as anticipated, most of the correct classifications are in the diagonal cells. The other cases (i.e., $D=2$ and 3 h) show similar outcomes. Further, in all cases, if a misclassification occurs, the tendency is misclassification as a neighboring class rather than as a remote class. For example, the observed *HL* may be misclassified as *ML*, but not as *LL*. Table 4 lists the accuracy in each correct classification of each level at different lag times. According to Table 4, the maximal accuracy of *HL*, *ML*, and *LL* occurs at $D=3$ h. Moreover, of the three lag-time cases, the highest average accuracy of 89.94% appears in the $D=3$ case.

In the validation phase, the MSE was used for evaluating the performance of the models, defined as

$$MSE = \frac{1}{N-1} \sum_{j=1}^N (O_j^{\text{pre}} - O_j^{\text{obs}})^2 \quad (4)$$

where O_j^{pre} is the predicted water stages at tuple j , O_j^{obs} is the observed water stages at tuple j , and N is the number of hourly tuples. The MSE is often used to be a measure of variance. Therefore, the MSE was regarded as a generalization of the variance concept in this study. Generally, the smaller the criterion value, the better will be the performance of the predicted outcomes.

Figure 7 shows the bar charts for water-stage forecasts obtained from the $\{HL\}$, $\{ML\}$, and $\{LL\}$ classified datasets by ANN_{HL} , ANN_{ML} , and ANN_{LL} , respectively, as well as all the average MSE values. As seen in Fig. 7a, in CART-MLP for specific *HL* and *LL*, the suitable cases are $D=2$ h (MSE=0.022 and 0.033, respectively), and for *ML*, the suitable case is $D=3$ h (MSE=0.027). In terms of average performance for lag times,

TABLE 3. CONFUSION MATRIX FOR CLASSIFIED LEVELS AT VARIOUS LAG TIMES USING VALIDATION DATASETS

Classification trials		Classified level for lag-time cases								
		D=1 h			D=2 h			D=3 h		
		HL	ML	LL	HL	ML	LL	HL	ML	LL
Observed level	HL	191	41	0	198	34	0	206	26	0
	ML	40	159	31	30	188	12	17	204	9
	LL	0	29	215	0	26	218	0	19	225

HL, higher level; ML, middle level; LL, lower level.

the optimal case occurs at $D=3$ h (MSE=0.029). Also, in CART-RBF (Fig. 7b), for specific HL and ML, the suitable cases are of $D=3$ h (MSE=0.037 and 0.021, respectively), and for LL, the suitable case is of $D=2$ h (MSE=0.020). In the same way, the optimal case appears at $D=3$ h (MSE=0.027).

Comparisons with three benchmark models

In the CART-ANN models analysis, misclassification by the CART classifier model might propagate to the ANN-based predictors. Therefore, three benchmark models, CART, BPNN, and RBFNN, were employed for comparisons. In general, the mapping function $f'\{\cdot\}$ of these models can be written as

$$\begin{aligned} A_5(t+1) &= f'\{(A_1(t-\Delta t), A_2(t-\Delta t), A_3(t-\Delta t), \\ A_4(t-\Delta t), A_5(t-\Delta t), A_6(t-\Delta t)), \Delta t=1, D\} \end{aligned} \quad (5)$$

Note that the attributes of these benchmarks are of numeric type and the datasets used in both the training and validation phases are the same as those used in the CART-ANN models. Similarly, for the specific D ranging from 1 to 3 h, the numbers of inputs are 6, 12, and 18, respectively. In the model analysis, the tuples of water stages were not categorized into three levels of HL, ML, and LL; instead, the whole dataset was used.

Figure 8a compares the prediction performances of CART, BPNN, RBFNN, CART-MLP, and CART-RBF models in terms of MSE obtained using the validation dataset. In the CART, the MSE values of the three cases are 0.139, 0.106, and 0.104, respectively, and the best lag-time case appears at $D=3$ h. Also, in the BPNN and RBFNN, the best cases occur at $D=3$ h, where MSE=0.044 and 0.043, respectively. Comparing all the cases of the five models shows that the optimal case occurs at $D=3$ h, where the MSE obtained using the CART-RBF equals to 0.027.

To explore the bias, another criterion, the mean absolute error (MAE), was employed. The MAE was defined as

$$MAE = \frac{1}{N-1} \sum_{j=1}^N |O_j^{\text{pre}} - O_j^{\text{obs}}| \quad (6)$$

The MAE is computed through a term-by-term comparison of the error in the prediction with respect to the actual value of the variable. Thus, the MAE is an unbiased statistic for measuring the predictive capability of a model (Hu *et al.*, 2001; Wang *et al.*, 2009). Figure 8b gives the MAE results of the five models. As can be seen, the MAE results in the case of $D=3$ h was better than other lag times. Here, the MAE values in the $D=3$ h of five models were (0.226, 0.150, 0.129, 0.128, 0.116 m), respectively. Apparently, the CART-RBF had the smallest bias measurement. Interestingly, we found that the case of $D=1$ h was the worst of all model cases in terms of MSE and MAE. The reason might be, as mentioned earlier, the travel time of the studied basin was >1 h. Therefore, the hydrological inputs of $D=1$ h may not provide sufficient information for model predictions because of the delay time of upstream flood wave.

FIG. 7. Forecast water-stage levels for different lag-time cases: Results of (a) CART-MLP and (b) CART-RBF. MSE, mean square error.

TABLE 4. ACCURACY FOR CORRECTLY CLASSIFIED LEVELS AT VARIOUS LAG TIMES

Accuracy (%)		Lag time (h)		
		D=1	D=2	D=3
Level of correct classification	HL	82.68	86.84	92.38
	ML	69.43	75.81	81.93
	LL	87.40	94.78	96.15
Average of accuracy		80.03	85.55	89.94

FIG. 8. Performances (a) MSE and (b) MAE of five models for three lag-time cases. BPNN, back-propagation neural network.

Figure 9 depicts the scattered plots of observation versus simulation according to the five model results for Case 3. As can be seen, the CART-RBF achieved more accurate prediction than the traditional CART, BPNN, RBFNN, and CART-MLP. In other words, the CART-RBF demonstrated good prediction ability.

The limitations of the CART-ANN need to be mentioned. Similar to other machine learning techniques, CART-ANN

models have a limited domain of applicability and are mostly case dependent. Therefore, their generalizations are limited and only applicable in the range of training data (Etemad-Shahidi and Mahjoobi, 2009). It is suggested to use this method in other geographical locations with different hydrological conditions to have a more comprehensive evaluation.

Conclusions

The purpose of this article is to develop a model for water-stage prediction during typhoon attacks. This study presented a CART-ANN model that combines both the two neural network techniques (i.e., MLP and RBFNN) and the decision trees (i.e., CART) to overcome the accuracy problem of water-stage forecasts. The two CART-ANN models (i.e., CART-MLP and CART-RBF) involve a two-step predicting approach that can be employed to address the complicated problem.

The proposed model was applied to the Tanshui River Basin in Taiwan. The Taipei Bridge, which is close to the estuary and affected by tidal effects, was taken as the study gauge. Data of 50 typhoons from 1996 to 2007 were collected. When using CART-ANN, first step, we classified the levels of water stages into the *HL*, *ML*, and *LL*. Second, the three sub-ANN-based water-stage predictors (namely ANN_{HL} , ANN_{ML} , and ANN_{LL}) were separately trained to predict water levels in *HL*, *ML*, and *LL*. To identify the suitable length of lag time for predicting the channel stages from various hydrological sources, three cases with lag time ranging from 1 to 3 h were designed. The forecast performances of CART-ANN were compared with the results obtained from three

FIG. 9. Validation results of observation versus prediction for four lag-time cases: (a) CART, (b) BPNN, (c) RBFNN, (d) CART-MLP, and (e) CART-RBF.

benchmark models, including CART, BPNN, and RBFNN. The MSE and MAE were used for evaluating the variance and bias performances of the models. Results showed that the CART-ANN achieved more accurate prediction than the three benchmarks in all cases in terms of MSE and MAE. Consequently, the CART-ANN successfully demonstrated good accuracy of prediction and offered a practical solution to water-level prediction problems.

For this research, the CART-ANN model addresses successfully the problem of water-level prediction. In a larger basin such as the Tanshui River, when reservoir operators need to release from reservoirs for real-time flood control in a multireservoir basin, they may face with complex considerations such as channel level routing under tidal effects. Future studies are likely to focus on embedding the CART-ANN model into the multireservoir flood control operation model to expand the capability and applicability of the model.

Acknowledgments

The support by the National Science Council, Taiwan (Grant No. NSC 98-2111-M-464-001 and NSC 100-2111-M-464-001) is greatly appreciated. The authors are grateful for the constructive comments of the referees.

Author Disclosure Statement

No competing financial interests exist.

References

- Alvisi, S., and Franchini, M. (2011). Fuzzy neural networks for water level and discharge forecasting with uncertainty. *Environ. Model. Softw.* 26, 523.
- Apte, C., and Weiss, S. (1997). Data mining with decision trees and decision rules. *Future Generation Comput. Sys.* 13, 197.
- Atiya, A.F., El-Shoura, S.M., Shaheen, S.I., and El-Sherif, M.S. (1999). A comparison between neural-network forecasting techniques—Case study: river flow forecasting. *IEEE Trans. Neural Netw.* 10, 402.
- Bazartseren, B., Hildebrandt, G., and Holz, K.P. (2003). Short-term water level prediction using neural networks and neuro-fuzzy approach. *Neurocomputing* 55, 439.
- Bessler, F.T., Savic, D.A., and Walters, G.A. (2003). Water reservoir control with data mining. *J. Water Res. Plann. Manage.* 129, 26.
- Bhattacharya, B., and Solomatine, D.P. (2005). Neural networks and M5 model trees in modelling water level-discharge relationship. *Neurocomputing* 63, 381.
- Boughton, W.C. (2007). Effect of data length on rainfall-runoff modeling. *Environ. Model. Softw.* 22, 406.
- Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and Regression Trees*. Belmont: Wadsworth.
- Chang, F.J., and Chen, Y.C. (2003). Estuary water-stage forecasting by using radial basis function neural network. *J. Hydrol.* 270, 158.
- Chau, K.W. (2006). Particle swarm optimization training algorithm for ANN in stage prediction of Shing Mun River. *J. Hydrol.* 329, 363.
- Chau, K.W., Wu, C.L., and Li, Y.S. (2005). Comparison of several flood forecasting models in Yangtze River. *J. Hydrol. Eng.* 10, 485.
- Cheng, C.T., Ou, C.P., and Chau, K.W. (2002). Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration. *J. Hydrol.* 268, 72.
- Cheng, C.T., Wang, W.C., Xu, D.M., and Chau, K.W. (2008). Optimizing hydropower reservoir operation using hybrid genetic algorithm and chaos. *Water Res. Manage.* 22, 895.
- Coulilaly, P. (2010). Reservoir computing approach to Great Lakes water level forecasting. *J. Hydrol.* 381, 76.
- Deka, P., and Chandramouli, V. (2005). Fuzzy neural network model for hydrologic flow routing. *J. Hydrol. Eng.* 10, 302.
- Etemad-Shahidi, A., and Mahjoobi, J. (2009). Comparison between M5' model tree and neural networks for prediction of significant wave height in Lake Superior. *Ocean Eng.* 36, 1175.
- Hsu, M.H., Lin, S.H., Fu, J.C., Chung, S.F., and Chen, A.S. (2010). Longitudinal stage profiles forecasting in rivers for flash floods. *J. Hydrol.* 388, 426.
- Hsu, N.S., and Wei, C.C. (2007). A multipurpose reservoir real-time operation model for flood control during typhoon invasion. *J. Hydrol.* 336, 282.
- Hu, T.S., Lam, K.C., and Ng, S.T. (2001). River flow time series prediction with a rangedependent neural network. *Hydrol. Sci. J.* 46, 729.
- Hwang, Y.S., and Bang, S.Y. (1997). Recognition of unconstrained handwritten numerals by a radial basis function neural network classifier. *Pattern Recognit. Lett.* 18, 657.
- Kearns, M., and Mansour, Y. (1999). On the boosting ability of top-down decision tree learning algorithms. *J. Comput. Syst. Sci.* 58, 109.
- Kisi, O. (2008). The potential of different ANN techniques in evapotranspiration modeling. *Hydrol. Processes* 22, 2449.
- Kurt, I., Ture, M., and Kurum, A.T. (2008). Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Syst. Appl.* 34, 366.
- Lin, J.Y., Cheng, C.T., and Chau, K.W. (2006). Using support vector machines for long-term discharge prediction. *Hydrol. Sci. J.* 51, 599.
- Liong, S.Y., Lim, W.H., and Paudyal, G.N. (2000). River stage forecasting in Bangladesh: neural network approach. *J. Comput. Civil Eng.* 14, 1.
- Mahjoobi, J., and Etemad-Shahidi, A. (2008). An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Appl. Ocean Res.* 30, 172.
- Maier, H.R., and Dandy, G.C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environ. Model. Softw.* 15, 101.
- Maji, P. (2008). Efficient design of neural network tree using a new splitting criterion. *Neurocomputing* 71, 787.
- Michael, J.A., and Gordon, S.L. (1997). *Data Mining Technique for Marketing, Sales and Customer Support*. New York: Wiley.
- Minns, A.W., and Hall, M.J. (1996). Artificial neural networks as rainfall-runoff models. *Hydrol. Sci. J.* 41, 399.
- Ni, Q., Wang, L., Ye, R., Yang, F., and Sivakumar, M. (2010). Evolutionary modeling for streamflow forecasting with minimal datasets: a case study in the West Malian River, China. *Environ. Eng. Sci.* 27, 377.
- Nourani, V., and Kalantari, O. (2010). Integrated artificial neural network for spatiotemporal modeling of rainfall-runoff-sediment processes. *Environ. Eng. Sci.* 27, 411.
- Park, J., Obeysekera, J., and VanZee, R. (2005). Prediction boundaries and forecasting of non linear hydrologic stage data. *J. Hydrol.* 312, 79.
- Quinlan, J.R. (1992). Learning with continuous classes. In *Proceedings of the Australian Joint Conference on Artificial Intelligence*. World Scientific, Singapore, pp. 343–348.

- Selamat, A., and Ng, C.C. (2011). Arabic script web page language identifications using decision tree neural networks. *Pattern Recognit.* 44, 133.
- Sethi, I.K., and Yoo, J.H. (1997). Structure-driven induction of decision tree classifiers through neural learning. *Pattern Recognit.* 30, 1893.
- Song, Y.H., Xuan, Q.X., and Johns, A.T. (1997). Comparison studies of five neural network based fault classifiers for complex transmission lines. *Electric Power Syst. Res.* 43, 125.
- SPSS Inc. (2004). *Clementine 10 User's Guide*. Chicago: SPSS.
- Su, M.C., Lo, H.H., and Hsu, F.H. (2010). A neural tree and its application to spam e-mail detection. *Expert Syst. Appl.* 37, 7976.
- Tan, P.N., Steinbach, M., and Kumar, V. (2006). *Introduction to Data Mining*. Boston: Addison Wesley.
- Thirumalaiah, K., and Deo, M.C. (1998). River stage forecasting using artificial neural networks. *J. Hydrol. Eng.* 3, 26.
- Tsujino, K., and Nishida, S. (1995). Implementation and refinement of decision trees using neural networks for hybrid knowledge acquisition. *Artif. Intell. Eng.* 9, 265.
- Ture, M., Kurt, I., Kurum, A.T., and Ozdamar, K. (2005). Comparing classification techniques for predicting essential hypertension. *Expert Syst. Appl.* 29, 583.
- Wang, W.C., Chau, K.W., Cheng, C.T., and Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol.* 374, 294.
- Wang, Y.C., Han, D., Yu, P.S., and Cluckie, I.D. (2006). Comparative modelling of two catchments in Taiwan and England. *Hydrol. Processes* 20, 4335.
- Wei, C.C., and Hsu, N.S. (2008). Multireservoir flood-control optimization with neural-based linear channel level routing under tidal effects. *Water Res. Manage.* 22, 1625.
- Wei, C.C., and Hsu, N.S. (2009). Optimal tree-based release rules for real-time flood control operations on a multipurpose multireservoir system. *J. Hydrol.* 365, 213.
- Wu, C.L., Chau, K.W., and Li, Y.S. (2009). Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. *Water Resour. Res.* 45, W08432.