

Research Article

Forecasting Short-Term Traffic Flow by Fuzzy Wavelet Neural Network with Parameters Optimized by Biogeography-Based Optimization Algorithm

Jeng-Fung Chen ¹, Shih-Kuei Lo,¹ and Quang Hung Do ²

¹Department of Industrial Engineering and Systems Management, Feng Chia University, Taichung 40724, Taiwan

²Faculty of Information Technology, University of Transport Technology, Hanoi 100000, Vietnam

Correspondence should be addressed to Quang Hung Do; quanghung2110@gmail.com

Received 10 May 2018; Revised 6 July 2018; Accepted 2 August 2018; Published 8 October 2018

Academic Editor: Pedro Antonio Gutierrez

Copyright © 2018 Jeng-Fung Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Forecasting short-term traffic flow is a key task of intelligent transportation systems, which can influence the traveler behaviors and reduce traffic congestion, fuel consumption, and accident risks. This paper proposes a fuzzy wavelet neural network (FWNN) trained by improved biogeography-based optimization (BBO) algorithm for forecasting short-term traffic flow using past traffic data. The original BBO is enhanced by the ring topology and Powell's method to advance the exploration capability and increase the convergence speed. Our presented approach combines the strengths of fuzzy logic, wavelet transform, neural network, and the heuristic algorithm to detect the trends and patterns of transportation data and thus has been successfully applied to transport forecasting. Other different forecasting methods, including ANN-based model, FWNN-based model, and WNN-based model, are also developed to validate the proposed approach. In order to make the comparisons across different methods, the performance evaluation is based on root-mean-squared error (RMSE), mean absolute percentage error (MAPE), and correlation coefficient (R). The performance indexes show that the FWNN model achieves lower RMSE and MAPE, as well as higher R , indicating that the FWNN model is a better predictor.

1. Introduction

In the transportation area, attention is not only paid to construct physical system capacity but also to improve operational efficiency and integration. The intelligent transportation system (ITS) applying the advanced sensing, analysis, control, and communications technologies aims to ease traffic congestion, improve traffic management, and reduce environmental impact. As ITSs have been widely developed throughout the world, how to improve the ability to predict traffic flow in the short term (within the next one hour, e.g., 5 min, 10 min, and 15 min) has been getting much attention from researchers. Short-term traffic flow forecasts can support proactive transportation management and comprehensive traveler information service. The goal is to predict traffic conditions in a transport network based on its

past behavior. Several methods have been implemented for the short-term traffic flow forecasts. These methods can be grouped into two categories: (1) methods based on statistical techniques and (2) methods based on artificial intelligence techniques.

With the assumption that the characteristics of forecasting traffic flow data are similar to historical and current flow data, forecasting models based on statistical techniques utilize mathematical statistics to deal with the pervious and current measurements of traffic flow and forecast the future values of traffic flow. Several well-known classical time-series approaches are the Box-Jenkins method [1], autoregressive integrated moving average (ARIMA) model [2, 3], seasonal ARIMA (SARIMA) [4, 5], and a number of variant forms of ARIMA models. Afterwards, researchers found that ARIMA model cannot tackle the problem of forecasting the extreme

volume values [6, 7]. Moreover, the short-term traffic flow forecast is more easily affected by the stochastic interferential factors than the long-term one, the uncertainty is greater, and the disciplinary laws are less obvious. Therefore, using the short-term traffic forecasting models based on the classical mathematical methods such as statistical techniques, the precision of forecast cannot meet the requirement of real-time transportation management systems [8].

As for the use of artificial intelligence-based techniques, several approaches have been applied to the task of traffic forecasting. Artificial neural network (ANN) is certainly the most widely used one for forecasting the transportation data, especially the short-term traffic flow forecasting [9]. It is a supervised learning algorithm that can be trained to learn a function between input features and the output, which is represented by the target to be predicted. The most widely used ANN-based models in short-term traffic flow forecasting are multilayer perception (MLP), backpropagation neural networks (BPNN), and radial basis function neural networks (RBFNN) [10]. The pros and cons of these models have been addressed in the literature [7, 11, 12].

Another development of artificial intelligence based-techniques is the combination of the ANN and other methods. Xia and Zhang [12] combined the strengths of discrete wavelet transform and ANN processing to achieve strong nonlinear approximation ability and then applied them to the short-term traffic volume forecasting. A traffic flow prediction model based on wavelet transform and fuzzy neural network was proposed in optimal control of the intelligent traffic system [13]. Instead of using back-propagation algorithm, the master-slave particle swarm optimization (PSO) was used to optimize the parameters of the prediction model. A deep learning-based traffic flow prediction method (neural networks with many layers) was proposed to represent traffic flow features for prediction [14]. The method can successfully discover the latent traffic flow feature representation, such as the nonlinear spatial and temporal correlations from the traffic data. A traffic flow prediction model based on the fuzzy c -mean clustering method (FCM) and the neural network was proposed [15]. The FCM can improve the accuracy and robustness of the model, while ANN can optimize the generalization ability of the model.

Although various methodologies have been applied to the traffic forecasting problem, the ultimate objective remains the same to obtain the forecasting result with high accuracy and robustness. Attention has also focused on improving the existing methodologies and models. ANNs have been found to be more effective than traditional methods in various application areas [16–18]. Other than that, the hybrid intelligent system, which is based on the combination of artificial neural networks and other intelligent techniques, has been proposed to take the full advantages of ANNs. Fuzzy systems are appropriate if sufficient expert knowledge about the process is available, while neural networks are useful if sufficient process data are available or measurable. The fuzzy neural network can effectively solve nonlinear problems [19] and is particularly

useful in applications where classical approaches fail or too complicated to be used.

From the signal analysis point of view, the traffic flow can be considered as a linear combination of different traffic flow versus time frequencies. Every component of traffic flow corresponds to a range of frequencies. The wavelet transform is especially suitable for transient analysis because of its time-frequency characteristics with automatically adjusted time-window lengths. Recent studies have shown that the wavelet transform can be used as an effective tool for capturing important features and characteristics of the traffic flow.

On the other hand, new evolutionary algorithms, including biogeography-based optimization (BBO), inspired by the behavior of natural phenomena, were developed for solving optimization problems. Through the competitive results of benchmarking studies, these algorithms have been proven to be powerful and are considered to outperform the other well-known algorithms. The BBO, proposed by Simon [20], was inspired by the migration process of species. Since then, BBO has been used in solving various complicated problems and is considered to outperform other algorithms, such as genetic algorithms (GA), ant colony optimization algorithms (ACO) [21, 22].

Rather than choosing a single technique, it might be beneficial to take advantage of several individual techniques. The merit of BBO algorithm, the wavelet transform, the fuzzy system, and the success of ANNs have encouraged us to combine these techniques for forecasting traffic flow. The rest of the paper is organized as follows: Section 2 presents related techniques such as the fuzzy model, wavelet transform, BBO algorithm, and fuzzy wavelet neural network. Section 3 is devoted to the proposed fuzzy neural network trained by BBO. A case application is presented in Section 4. Section 5 reports results and discussion; finally, Section 6 gives the conclusion of the study.

2. Related Techniques

2.1. Fuzzy Model. The fuzzy logic model is an appropriate approach to model complex systems. It is a process of mapping from a given input to an output using the theory of fuzzy sets. Fuzzy logic systems have the ability to approximate any continuous function and deal with complex nonlinear systems with ill-defined conditions and uncertain factors [23]. Among the various fuzzy modeling techniques, Takagi–Sugeno–Kang (TSK) fuzzy model is one of the most popular ones because of its mathematical treatability. A TSK fuzzy model consisting of IF-THEN rules with fuzzy antecedents and a mathematical function at the consequent part of the form is as follows:

$$\begin{aligned}
 R^i : & \text{If } x_1 = A_1^i \text{ and } x_2 = A_2^i \text{ and } \dots \text{ and } x_r = A_r^i, \\
 & \text{Then } y^i = a_0^i + a_1^i x_1 + a_2^i x_2 + \dots + a_r^i x_r, \quad (1) \\
 & \text{for } i = 1, 2, \dots, K,
 \end{aligned}$$

where R^i represents the i th fuzzy inference rule, K is the number of rules, x_j ($j = 1, 2, \dots, r$) is the j th input, y^i is the

output of the fuzzy rule, $A_1^i, A_2^i, \dots, A_r^i$ are fuzzy sets with membership functions $A_j^i(x_j)$, and a_j^i s are real values. The output of the TSK fuzzy model is computed by

$$y = \frac{\sum_{i=1}^K \omega_i y_i}{\sum_{i=1}^K \omega_i}, \quad (2)$$

where ω_i is the firing strength of rule R^i , which is calculated by

$$\omega_i = A_1^i(x_1) \times A_2^i(x_2) \times \dots \times A_r^i(x_r). \quad (3)$$

The fuzzy membership functions of $A_j^i(x_j)$ are Gaussian functions calculated by

$$\mu_{ij}(x_j) = \exp\left(-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right), \quad (4)$$

where c_{ij} is the center and σ_{ij} represents the standard deviation for fuzzy membership function associated with rule i .

It has been shown that the TSK fuzzy model can separate the input space into local fuzzy regions and then approximates a system in every region by a linear equation.

2.2. Wavelet Transform. The wavelet transform is a recently developed mathematical tool for signal analysis. It has been applied successfully in a wide range of time-series analysis, such as in astronomy, data compression, signal and image processing, earthquake prediction, and so on [24, 25]. The fundamental idea in wavelet analysis is to select a suitable wavelet (mother wavelet), and then perform an analysis using its translated and dilated versions. There are several kinds of wavelets that can be used as a mother wavelet, such as the Haar wavelet, Meyer wavelet, Coiflet wavelet, Daubechies wavelet, and Morlet wavelet. Each wavelet has specific characteristics.

Wavelets are as in the following form:

$$\psi_{a,b} = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right), \quad (a, b \in \mathbb{R}, a \neq 0). \quad (5)$$

Wavelets are a family of functions derived from the function $\psi(x)$ by the operation of dilation and translation. $\psi(x) \in L^2(\mathbb{R})$ is a mother wavelet function that satisfies the following condition:

$$C_\psi = \int_0^{+\infty} \frac{|\widehat{\psi}(\omega)|^2}{\omega} d\omega < +\infty, \quad (6)$$

where $\widehat{\psi}(\omega)$ is the Fourier transform of $\psi(x)$.

The function $f(x)$ can be represented by the following equation:

$$f(x) = \frac{1}{C_\psi} \iint W f(a, b) |a|^{-1/2} \psi\left(\frac{x-b}{a}\right) \frac{1}{a^2} da db, \quad (7)$$

where $W f(a, b) = |a|^{-1/2} \int_{-\infty}^{+\infty} \psi(x-b/a) f(x) dx$ is the continuous wavelet transform of $f(x)$.

2.3. Fuzzy Wavelet Neural Network. The fuzzy wavelet neural network (FWNN) presented by Yilmaz and Oysal [26] and

Linhares et al. [27] is utilized in this work to construct the traffic flow forecasting model. The six-layer structure of the FWNN is presented in Figure 1.

Layer 1: the input layer transfers the input signal $x = \{x_1, x_2, \dots, x_n\}$ to the second layer.

Layer 2: in the fuzzification layer, each neuron in this layer gets fuzzy membership functions in the IF part of the rules. The membership functions are parameterized according to the specific applications. The outputs of the fuzzification layer are the values of membership functions. The Gaussian membership function is the most used one:

$$A_j^i = \exp\left(-\frac{1}{2} \left(\frac{x_j - \mu_{ij}}{\sigma_{ij}}\right)^2\right), \quad j = 1, 2, \dots, n \quad (8)$$

and $i_j = 1, 2, \dots, l_j$,

Layer 3: this layer is the fuzzy rule layer (inference layer). Each neuron has a fuzzy rule. The output of the l th node η_l is obtained by aggregating $A_j^i(x_j)$ using the AND (t-norm):

$$\eta_l = \prod_{j=1}^n A_j^{i_j}(x_j), \quad (9)$$

where $l = i_1, i_2, \dots, i_n; i_1 = 1, \dots, l_1, i_2 = 1, \dots, l_2; \dots; i_n = 1, \dots, l_n$.

Each possible combination of input membership functions denotes a fuzzy rule. All fuzzy rules are summed up to the node placed between layers 3 and 4.

Layer 4: in the normalization layer, each neuron calculates the normalization value for the l th rule by using the following equation:

$$\overline{\eta}_l = \frac{\eta_l}{\sum_{i=1}^m \eta_i} \quad (l = 1, \dots, m). \quad (10)$$

The output of this layer represents the contribution ratio of a rule to the final result.

Layer 5: the consequent layer calculates the weighted output value of a rule.

$$f_l = \overline{\eta}_l \Psi_l, \quad (l = 1, \dots, m). \quad (11)$$

In this study, the Mexican hat wavelet function is utilized in this layer as follows:

$$\Psi_l = \sum_{i=1}^n w_{il} \left(1 - \left(\frac{x_i - b_{il}}{c_{il}}\right)^2\right) \exp\left(-\frac{1}{2} \left(\frac{x_i - b_{il}}{c_{il}}\right)^2\right). \quad (12)$$

Layer 6: the output layer computes the overall output. All signals from the wavelet neurons are summed up.

$$y = \sum_{l=1}^m f_l. \quad (13)$$

2.4. Biogeography-Based Optimization (BBO) Algorithm. Biogeography is the science which studies the geographical distribution of living species. BBO is a new inspired algorithm that is based on biogeography [20, 28]. Simon [20] developed the mathematical models of biogeography to solve optimization problems. In BBO, variables that determine the quality of habitat are called suitability index

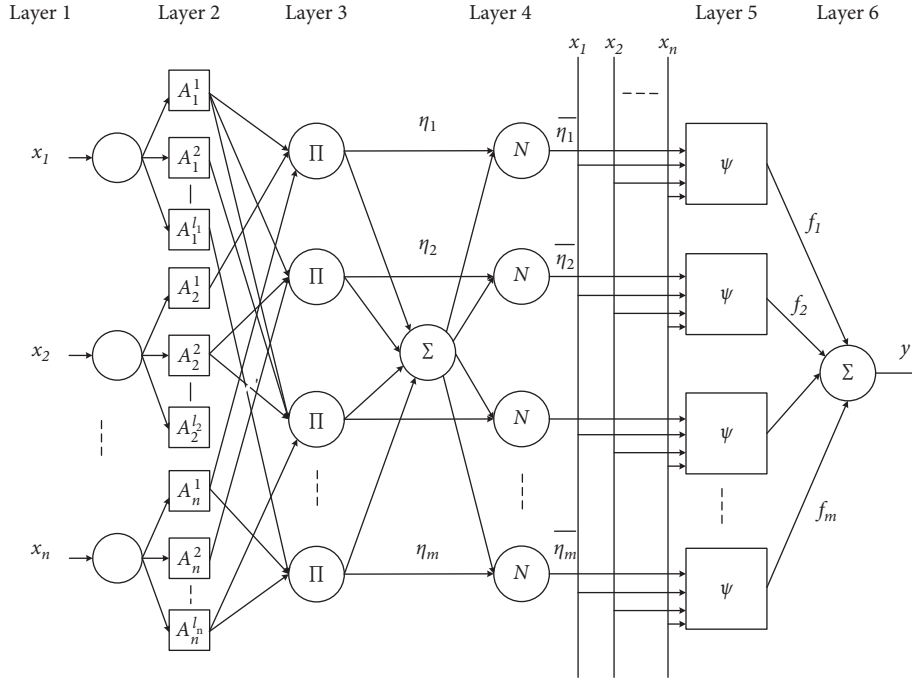


FIGURE 1: The structure of the FWNN.

variables (SIVs), and each habitat is considered as an individual and has its habitat suitability index (HSI). SIVs are independent variables, and HSI depends on SIVs. Habitats with large HSI accommodate more species which are suitable for species living, and, conversely, a low-HSI habitat contains fewer species which are not suitable for species living. When the number of species in a habitat increases, there is a strong tendency for species to emigrate from crowded habitats to find new ones with better life-supporting conditions and lower population density than the old habitats. Habitats with low population density may accept a lot of new species from high-HIS habitats by providing adequate life-supporting characteristics. The objective function can be considered as HSI, and the evolutionary procedure of BBO is to acquire the solutions which maximize the HSI by using the immigration and emigration features of the habitats. The pseudocode of the BBO algorithm can be described in Algorithm 1.

In BBO, the probability to choose the solution H_i as the immigrating habitat depends on its immigration rate λ_i and the probability to choose the solution H_j as the emigration habitat depends on its emigration rate μ_j . Migration can be demonstrated as

$$H_i(\text{SIV}) \leftarrow H_j(\text{SIV}). \quad (14)$$

The immigration rate and emigration rate can be described as

$$\lambda_i = I \left(1 - \frac{k_i}{n} \right), \quad (15)$$

$$\mu_i = E \left(\frac{k_i}{n} \right),$$

where I and E are the maximum possible immigration rate and emigration rate, respectively. k_i represents the rank of habitat i after sorting all habitats according to their HSI and n is the number of solutions in the population. A better solution has higher emigration and lower immigration rates and vice versa.

The original BBO has several drawbacks including insufficient exploration capability and slow convergence speed. In order to improve the BBO, an attempt was made by combining the random ring topology and Powell's method [29]. The original BBO uses a global topology in which each pair of habitats can directly inform to the others. However, computing the distances between all pairs of habitats takes a high computational cost. A simplest form of local topology, called the ring topology, has been proposed to be used in BBO. In the ring topology, in order to reduce the computational cost and avoid premature convergence, each habitat is connected to only two other habitats, as shown in Figure 2 [30]. The pseudocode of the ring topology is presented in Algorithm 2.

Powell's method is utilized to effectively improve the solution precision. A perturbed best solution is used as the initial search point, and parameters ε and δ are the termination criteria and step size, respectively. The step size δ_j of the j th dimension can be derived from Equation 17:

$$\delta_j = \frac{0.0001 \cdot \sum_{i=1}^m (H_{i,j} - H_{best,j})}{0.1ps}, \quad (16)$$

where ps is the population size and $0.1ps$ is the number of solutions selected for calculation, H_i presents the i th solution, and H_{best} is the best solution. The step size δ_j decreases when the number of iterations increases.

```

Initialize the BBO parameters
Generate a set of habitats, corresponding to the potential solutions
Evaluate the fitness value or HSI for each habitat
While Stopping criterion is not satisfied do
    Decide immigration rate  $\lambda$  and emigration rate  $\mu$  for each habitat
    Modify habitats based on  $\lambda$  and  $\mu$ 
    For  $i = 1$  to Population size do
        Use  $\lambda$  to probabilistically decide whether to modify a habitat
        If  $\text{rand}(0,1) < \lambda_i$ 
            Select habitat  $H_j$  to immigration
            Perform migration on  $H_i$  and  $H_j$ 
            Evaluate the fitness value or HSI for newly generated solution
            Replace the new solution with  $H_i$ 
        End If
        If  $\text{rand}(0,1) < P_{Mutation}$ 
            Apply mutation on  $H_i$ 
            Evaluate the fitness value or HSI for newly generated solution
        End If
    End For
    Update habitats' population
End While

```

ALGORITHM 1: The BBO algorithm.

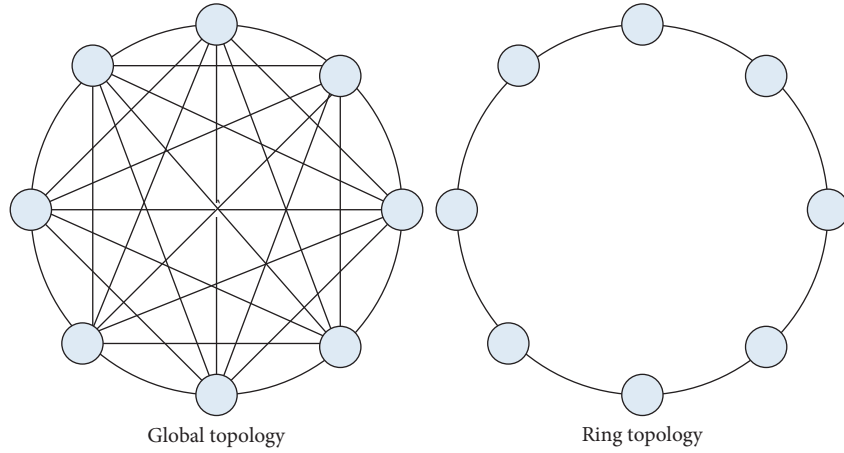


FIGURE 2: The global and local ring topologies.

Moreover, the modified mutation proposed by Lohokare et al. [31] is also utilized to increase the population diversity. The detail of the improved BBO is represented in Algorithm 3.

3. The Proposed Fuzzy Wavelet Neural Network (FWNN) Trained by the Improved BBO

A fuzzy wavelet neural network (FWNN) with parameters trained by the improved BBO (hereinafter referred to as FWNN-iBBO-based model) was developed for forecasting the traffic flow. The parameters in the FWNN structure that need to be updated are as follows:

- (1) μ_{i_j} and σ_{i_j} (the center and the standard deviation for Gaussian fuzzy membership function associated with rule i in the layer 2, resp.)

- (2) The translation parameters b_{il} and dilation parameters c_{il} of wavelet functions
- (3) The weight parameters w_{il} in the consequent part the rules
- (4) The parameter vector $\Theta = (\mu_{i_j}, \sigma_{i_j}, b_{il}, c_{il}, w_{il})$

The FWNN parameters are updated according to the performance index of root-mean-squared error (RMSE) given in the following:

$$\text{RMSE}(\Theta, \tilde{x}) = \sqrt{\frac{1}{K} \sum_{i=0}^K (y_i - \hat{y}_i)^2}, \quad (17)$$

where y_i is the actual (desired) value and \hat{y}_i represents the forecasted value.

```

Randomly organize all the habitats in a ring topology
For  $i = 1$  to Population size
  For  $j = 1$  to dimension of solution
    If  $\text{rand}(0,1) < \lambda_i$ 
      Uniformly generate a random integer value  $k_0 \in \{1, 2, \dots, \text{Population size}\}$ ;
      Determine two neighbor habitats  $H_{r_1}$  and  $H_{r_2}$  of  $H_{k_0}$  in the ring topology;
       $\alpha = 2 * \text{rand}(0, 1) - 1$ ;
      If  $\text{rand}(0, 1) < (\mu_{r_1} / (\mu_{r_1} + \mu_{r_2}))$ 
         $H_{i,j} = H_{r_1,j} + \alpha \cdot (H_{r_1,j} - H_{i,j})$ ;
      else
         $H_{i,j} = H_{r_2,j} + \alpha \cdot (H_{r_2,j} - H_{i,j})$ 
      End If
    End If
  End For
End For

```

ALGORITHM 2: Pseudocode of the ring topology.

```

Set the initial parameters: ps (population size);  $\varepsilon$  (tolerance), G (generation) and limit
Generate the population Pop with ps habitats randomly
Evaluate the HSI (fitness) for each habitat  $H_i$  in Pop
Gen = 0
While the stopping criteria are not satisfied do
  Gen = Gen + 1
  Sort all the habitats according to their HSI
  Calculate the immigration rate and emigration rate
  Randomly sort all the solutions
  Perform migration operator RMO
  Perform mutation operator
  Assure that each habitat legal based on boundary constraints
  Assess the HSI (fitness) for trial habitat  $M_i$  in the new population Pop
  For  $i = 1$  to ps
    If  $f(M_i) < f(H_i)$ 
       $H_i = M_i, \text{trial}_i = 0$ 
    Else
       $\text{trial}_i = \text{trial}_i + 1$ 
    End If
  End For
  If  $\text{mod}(\text{Gen}, G) = 0$  do
     $\varepsilon = (\varepsilon / \text{FES})$ 
    Calculate  $\delta$  according to Powell's method
    Randomly choose  $k \in \{1, 2, \dots, ps\}$ 
     $x_0 = H_{\text{best}} + \text{rand}(0, 1) \cdot (H_{\text{best}} - H_k)$ 
    Perform Powell's procedure with  $x_0$  as starting point and obtain point  $y$ .
    Use  $y$  to replace the worst habitat in the population
  End If
  If  $\max(\text{trial}_i) > \text{limit}$ 
    Replace  $H_i$  with a randomly generated solution
  End If
End While

```

ALGORITHM 3: Pseudocode of the improved BBO.

In this study, the iBBO algorithm is utilized to train the forecasting model. The best parameters $\Theta = (\mu_i, \sigma_i, b_{il}, c_{il}, w_{il})$ are selected based on the performance criteria. First, the whole data set is grouped into the training set and the testing set. After the training process (as shown in Figure 3), the trained FWNN based on the training set is applied to the testing set, and the

performance criteria are recorded. The performance criteria are then applied to the trained model to estimate how well the trained model works. These criteria are used to compare forecasting values and actual values. They are as follows:

- (1) Root-mean-squared error (RMSE): this index calculates the residual between the actual value and

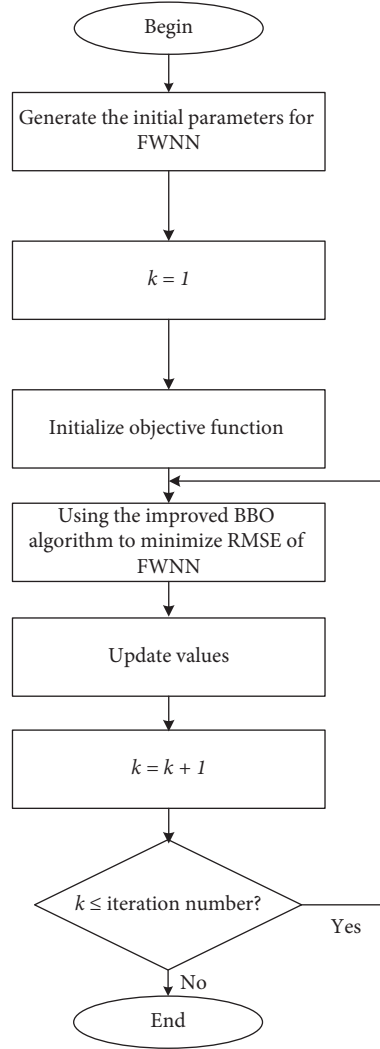


FIGURE 3: Training FWNN by iBBO.

predicted value. A model has better performance if it has smaller RMSE. RMSE equal to zero means perfect fit.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2}, \quad (18)$$

where y_i is the actual value, \hat{y}_i is the predicted value produced by the model, and n is the total number of observations.

- (2) Mean absolute percentage error (MAPE): this index indicates an average of the absolute percentage errors. A model with the lower MAPE achieves the better performance:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (19)$$

- (3) Correlation coefficient (R): this criterion indicates the strength of relationships between actual value and predicted value. The correlation coefficient has

a range from 0 to 1, and a model with the higher R means it has better performance.

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}, \quad (20)$$

where $\bar{y} = 1/n \sum_{i=1}^n y_t$ and $\bar{\hat{y}} = 1/n \sum_{i=1}^n \hat{y}_t$ are the average values of y_i and \hat{y}_i .

4. A Case Application

In this section, we use the traffic flow data from Ho Chi Minh City, Vietnam, to evaluate our developed model and compare the performance of our model with other models. Like other cities in Vietnam, Ho Chi Minh City is dealing with traffic problems characterized by mixed traffic flow including different categories of vehicles such as motorized and nonmotorized vehicles with the wide variation in sizes. All the vehicles including cars, buses, trucks, motorbikes, and bicycles are grouped into different categories, as shown in Table 1 [32]. Our study is

TABLE 1: Vehicles and their dimensions.

Category	Vehicles included	Average dimension (m)		Projected rectangular on ground (m ²)
		Length	Width	
Car	Car, jeep	3.72	1.44	5.39
Bus	Bus	10.1	2.43	24.74
Truck	Truck	7.5	2.35	17.62
Light commercial vehicle	Mini bus, van	6.1	2.10	12.81
Tractor	Tractor trailer	7.4	2.20	16.28
Two-wheeler	Scooter, motorbike	1.87	0.64	1.2

TABLE 2: Vehicles and their dimensions.

Category	Vehicles included	Projected rectangular on ground (m ²)	Area ratio car (reference vehicle)
Car	Car, jeep	5.39	1.00
Bus	Bus	24.74	0.22
Truck	Truck	17.62	0.31
Light commercial vehicle	Mini bus, van	12.81	0.42
Tractor	Tractor trailer	16.28	0.33
Two-wheeler	Scooter, motorbike	1.2	4.49

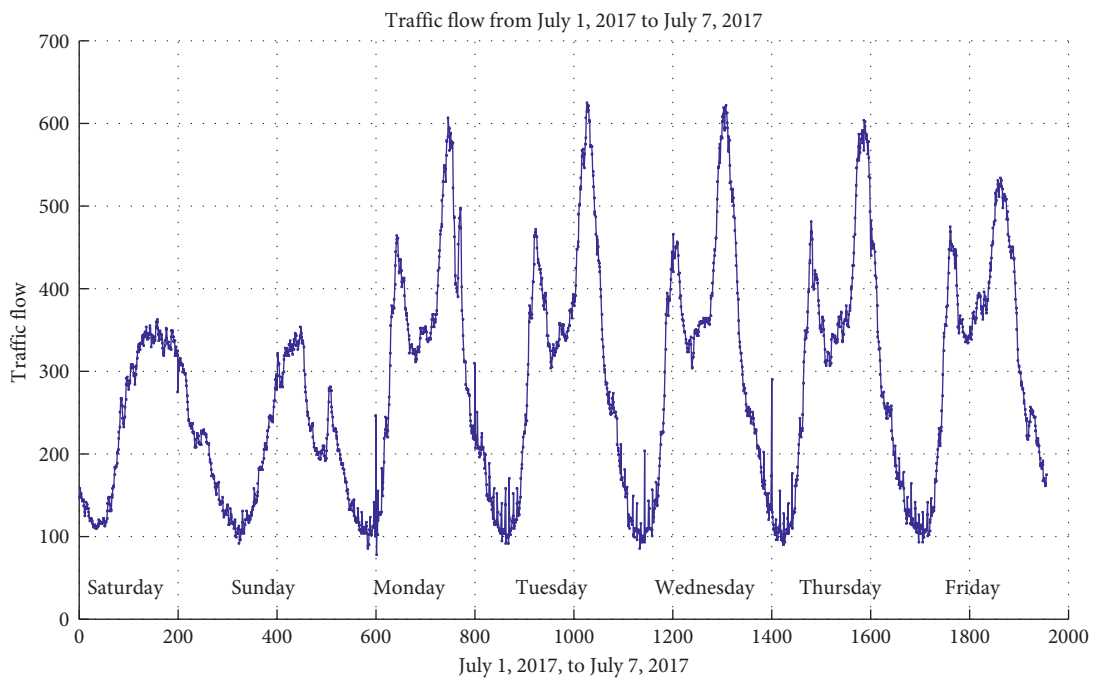


FIGURE 4: Traffic flow from July 1, 2017, to July 7, 2017.

aimed at forecasting traffic volume on the road from 621 T-junction to Thu Duc crossroad. The route is currently one of the busiest roads in Ho Chi Minh City (the largest city of Vietnam).

The total number of vehicles that pass over a given point during a given time interval is called volume. Traffic flow is the number of vehicles passing a reference point per unit of time, vehicles per hour. In the study, the numbers of vehicles crossing a fixed point of the road are counted. In mixed traffic, it is necessary to convert the whole traffic into one common standard or reference vehicle. For measuring the traffic volume, the car is

selected as the reference vehicle. Area ratio is a criterion for finding the equivalent factor of the reference vehicle and the other vehicles.

$$\text{Area ratio} = \frac{A_C}{A_Y}, \quad (21)$$

where A_C represents projected area of the reference vehicle (car) and A_Y denotes the projected area of "y" type vehicle. The area ratio for each vehicle is calculated and is presented in Table 2 [33].

A video recording of the forecasting section was done. The traffic flow datasets were then extracted from traffic

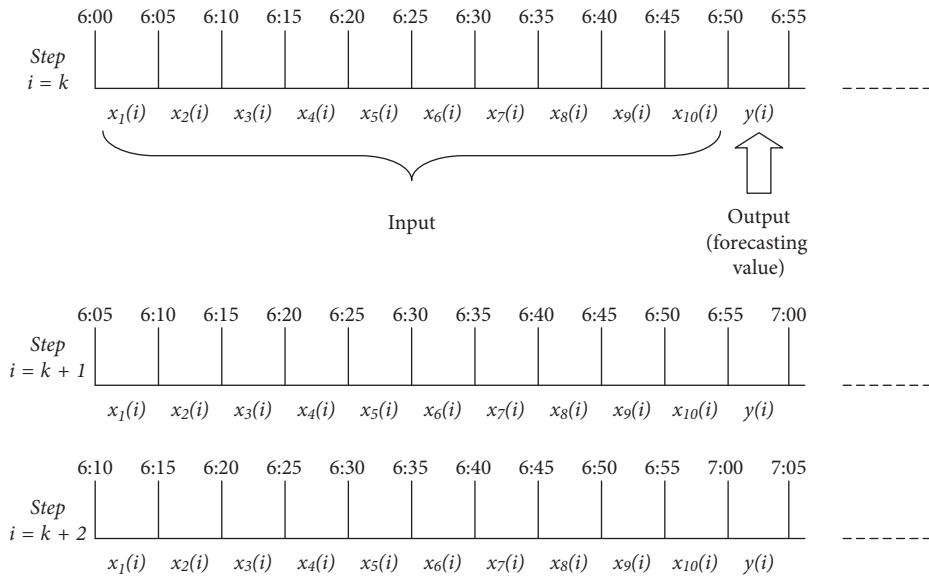


FIGURE 5: Illustration of 1-step traffic flow forecasting.

TABLE 3: Performance statistics of different forecasting models.

	RMSE	MAPE	R
ANN-based model	74.0569	0.2529	0.8880
FWNN-based model	32.5282	0.1025	0.9451
FWNN-BBO-based model	27.4678	0.0924	0.9768
FWNN-iBBO-based model	20.4034	0.0719	0.9846
WNN-based model	52.1006	0.1171	0.9617

cameras at a 15-minute interval. The data collection was conducted during the first six months of 2017 at an interval of 5 minutes. About 288 data samples were collected each day. The traffic flow in one week (from July 1 to July 7, 2017) on the monitoring site is shown in Figure 4. The traffic flow data show characteristic patterns tied to work-week activities. In workdays, the typical morning and evening peak hours are evident for urban routes. The evening peak has higher traffic volumes than the morning peak. Weekend days have lower-level peaks.

The traffic flow data were split into two parts: training data and testing data. Based on these data, several forecasting models are developed and evaluated.

For each model, ten historical data points (a total of 50 minutes) are used as inputs and the output is the forecasting value for the traffic volume in the next five minutes. For instance, in Figure 5, if at Step i , the current time is 6:50, then the inputs for each forecasting model are the 5 minutes traffic flow data from 6:00 to 6:50, and the output is the traffic flow forecasting value from 6:50 to 6:55. At each succeeding step, a newly observed traffic flow value is added as the input and the oldest value is removed, such that the input dimension is constant.

5. Results and Discussion

In this section, different forecasting models including FWNN-BBO-based model, ANN-based model, FWNN-based model,

WNN-based model, and the proposed FWNN-iBBO forecasting model are developed and investigated. For each model, we conducted 10 independent tests, and each test produced a set of performance criteria values. The average performance criteria for each model were calculated and are presented in Table 3. The scattering diagrams and traffic flow graphics are also drawn in Figure 6. A fivefold cross-validation method was used to avoid an overfitting problem.

For ANN-based model, we adopt a feedforward network (FFN) with one hidden layer to forecast traffic volume. The optimum number of neurons in the hidden layer was determined by varying their numbers, starting with a minimum of one, and then increasing in steps by adding one neuron each time. Hence, various FFN architectures were tested to achieve the optimum number of hidden neurons. The best performing architectures for ANN were found to be 10-6-1. The activation function from input layer to hidden layer is sigmoid. With no loss of generality, a commonly used activation function, $f(n) = 2/(1+e^{-2n})-1$, is utilized; while a linear function is used from the hidden layer to the output layer. The parameters for backpropagation were set as follows: the learning and momentum rates were 0.5 and 0.3, respectively. For the proposed FWNN-BBO forecasting model, the parameters for the BBO algorithm and iBBO algorithm were determined by trial and error. The parameters were set as follows: population size, $p_s=100$; maximum immigration rate, $I=1$; the maximum emigration rate, $E=1$; mutation probability, $m_{\max}=0.005$; $\varepsilon=0.1$; limit = 100; and $G=500$.

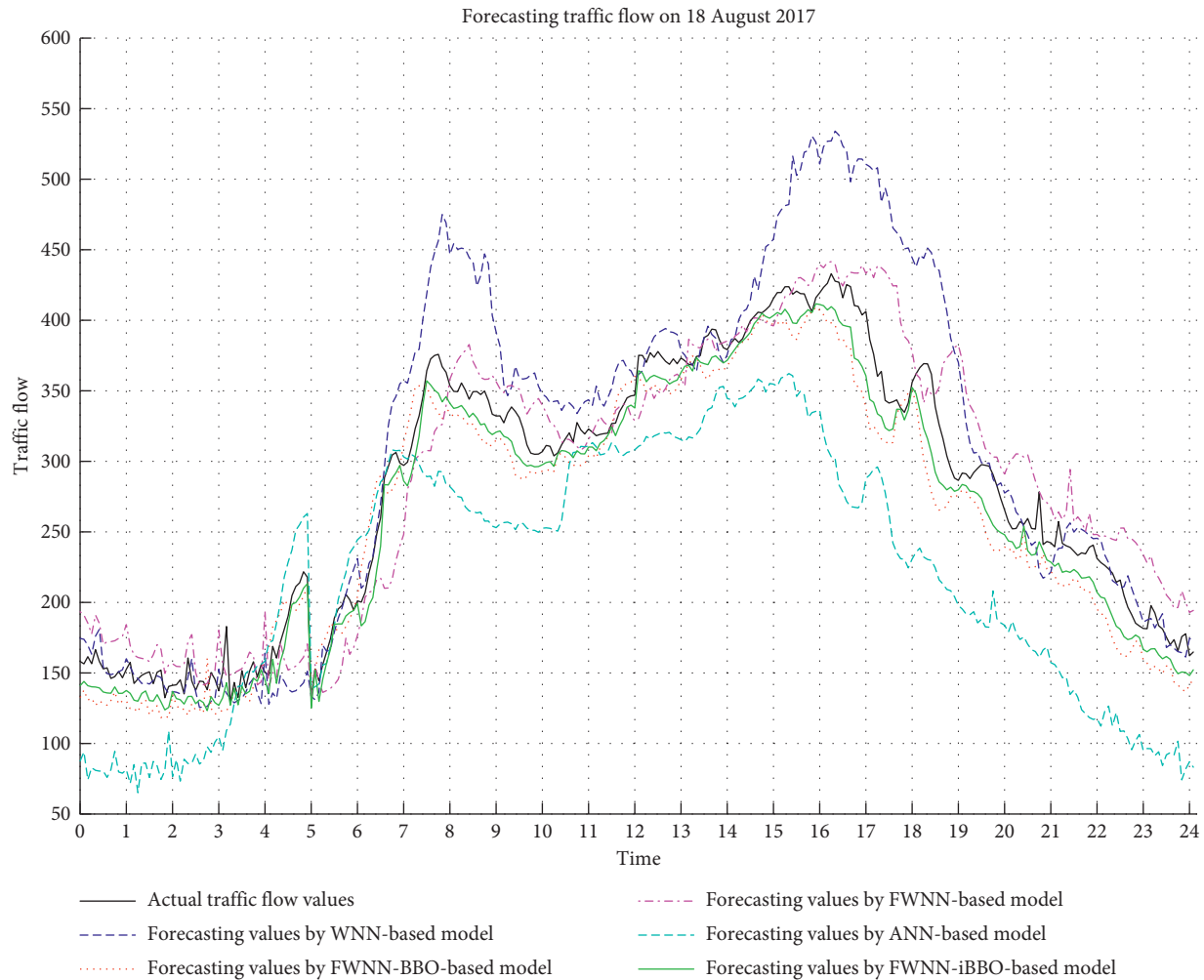


FIGURE 6: Forecasting traffic values by different forecasting models.

The models were implemented in the MATLAB 2015a environment. The simulation results were then obtained and are presented in Figures 6 and 7 and Table 3. The time series of actual and forecasting values obtained by the WNN-based model, FWNN-BBO-based model, FWNN-iBBO-based model, FWNN-based model, and ANN-based model are compared in Figure 6. The nearly perfect agreement between the trends in the plots of the actual and forecasting values indicates that the FWNN-BBO-based model is the most suitable model.

The performance criteria RMSE, MAPE, and R obtained by FWNN-iBBO-based model were calculated as 20.4034, 0.0719, and 0.9846, respectively. Theoretically, a forecasting model is accepted as ideal when RMSE and MAPE are small, and R is close to 1. It is very clear from Table 3 that the FWNN-BBO-based model has a smaller RMSE and MAPE as well as a bigger R than those of the ANN-based model, FWNN-based model, and WNN-based model. These performance criteria indicate that the assessed results obtained by the FWNN-iBBO-based model are highly correlated and more precise.

The comparison between actual values and forecasting values obtained by FWNN-BBO-based model and FWNN-iBBO-based model are also shown in Figure 7. The figure

presents the scatter diagrams that illustrate the degree of correlation between forecasting values and actual values. An identity line was drawn as a reference. In this figure, the identity line represents that the two sets of data are identical. The more the two datasets agree, the more the points tend to concentrate in the vicinity of the identity line. It may be observed that most forecasting values are very close to the actual values. This indicates a sound agreement between the forecasts by FWNN-BBO-based model and the actual values.

In order to evaluate the performance of the proposed approach, several popular and recent optimization algorithms including genetic algorithm (GA), particle swarm optimization (PSO), and cuckoo search (CS) algorithm were also applied to training FWNN (abbreviated as FWNN-GA, FWNN-PSO, and FWNN-CS). For each training algorithm, different sets of parameters were tried to obtain the best performance. For FWNN-GA, the population size was set at 30 and p_c and p_m were set at 0.6 and 0.4, respectively; and the number of iterations was set at 500. For FWNN-PSO, the number of initial population was set at 30 with c_1 and c_2 set to be 2, w decreased linearly from 0.9 to 0.4, and the initial velocities of particles were randomly generated from $[0,1]$. For FWNN-CS, the step size (α) was set at 0.01, the number

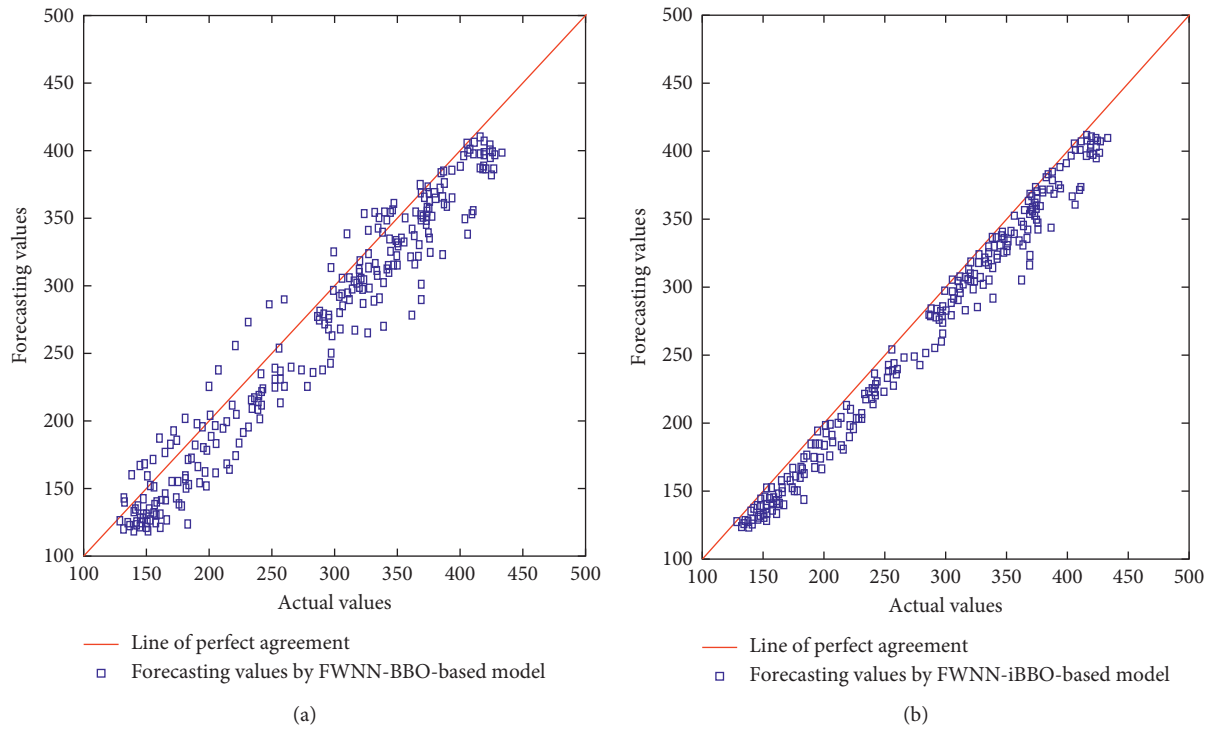


FIGURE 7: Forecasting values by FWNN-BBO-based model and FWNN-iBBO-based model.

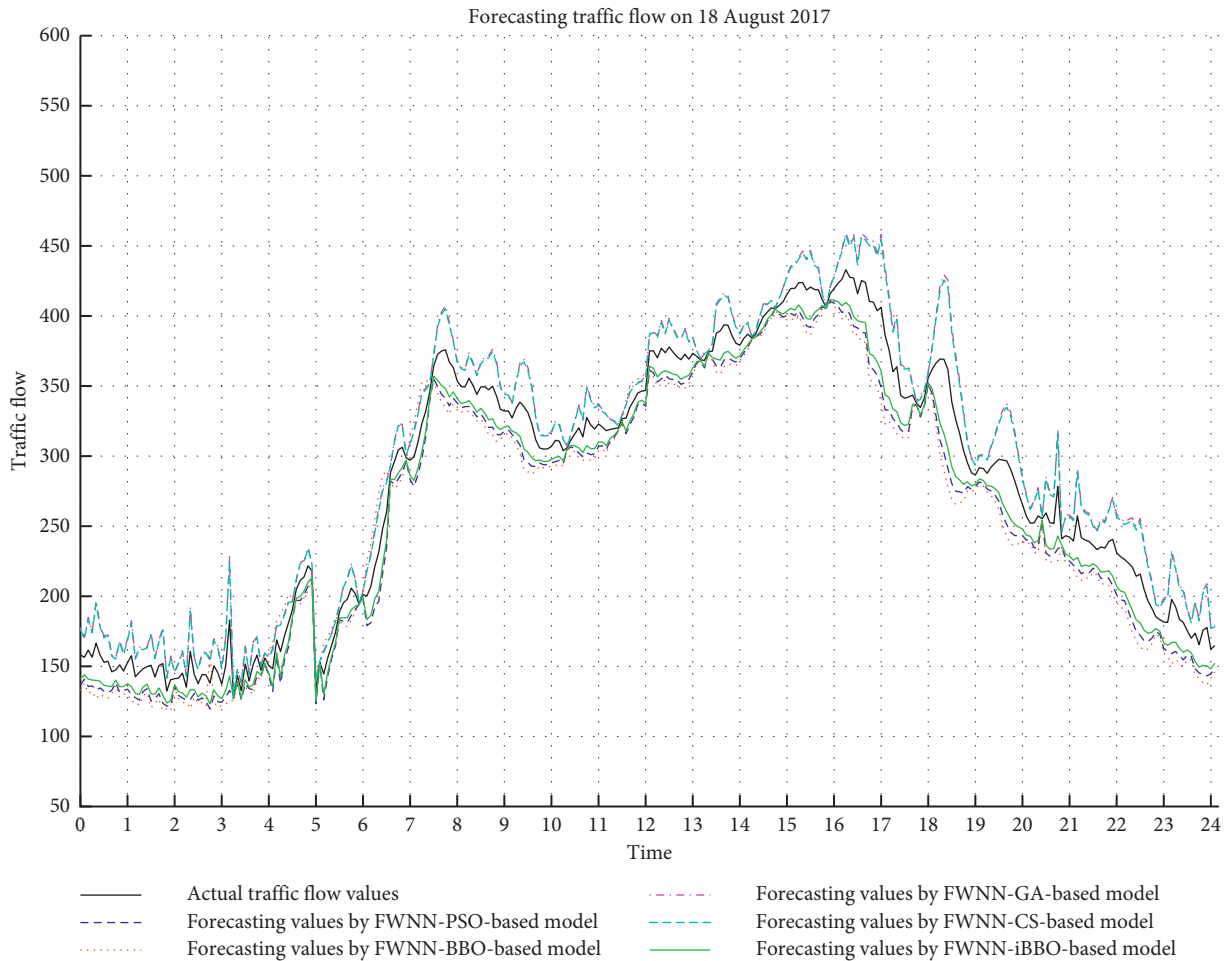


FIGURE 8: The forecasting values of FWNN trained by different optimization algorithms.

of nests was set at 30, and the net discovery rate (p_a) was set at 0.1. The results were recorded and depicted in Figure 8. As shown in the figure, the performances of FWNN-iBBO and the FWNN-BBO surpassed those of the FWNN-GA, FWNN-PSO, and FWNN-CS-based models. It can be concluded that the iBBO outperforms the GA, PSO, and CS algorithms in this study.

Based on the obtained results, it can be inferred that the proposed FWNN-iBBO-based model can be used to forecast the short-term traffic flow. The FWNN-iBBO-based model outperforms the FWNN-BBO-based model, ANN-based model, FWNN-based model, and WNN-based model, and the results show that its forecasting outcome is more accurate and reliable. Hence, the FWNN-iBBO-based model is acceptable and good enough to serve as a predictor of traffic flow.

6. Conclusions

In this study, different traffic flow forecasting models have developed and applied to forecast traffic flow on the road from 621 T-junction to Thu Duc crossroad which is characterized by mixed traffic flow. This study proposed a model based on fuzzy logic, wavelet transform, neural network, and the heuristic algorithm to forecast traffic flow. The results clearly demonstrated the superior forecasting performance of FWNN-iBBO model. It is concluded that FWNN can be utilized for short-term traffic flow prediction with mixed traffic conditions in Vietnam. The numerical experiments indicate the potential of the proposed method for large-scale network-wide traffic forecasting applications. As for the future research, it may be desirable to apply the proposed model to evaluate more traffic flow data from different locations. In addition, the proposed model can also consider additional factors, such as social events and weather to forecast the traffic flow.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Ministry of Science and Technology under grant no. MOST 106-2221-E-035-069.

References

- [1] G. E. P. Box and G. M. Jenkins, *Time series analysis, Forecasting and Control*, John Wiley & Sons, Inc, Hoboken, NJ, USA, 1970.
- [2] C. Moorthy and B. Ratcliffe, "Short term traffic forecasting using time series methods," *Transportation Planning and Technology*, vol. 12, no. 1, pp. 45–56, 1988.
- [3] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Journal of the Transportation Research Board*, vol. 1678, pp. 179–188, 1999.
- [4] S.V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *European Transport Research Review*, vol. 7, no. 3, p. 21, 2015.
- [5] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.
- [6] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: overview of objectives and methods," *Transport Rev*, vol. 24, no. 5, pp. 533–557, 2004.
- [7] J. Z. Zhu, J. X. Cao, and Y. Zhu, "Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections," *Transportation Research Part C: Emerging Technologies*, vol. 47, pp. 139–154, 2014.
- [8] H. Su and S. Yu, "Hybrid GA based online support vector machine model for short-term traffic flow forecasting," in *APPT 2007, LNCS 4847*, M. Xu, Y. Zhan, J. Cao, and Y. Liu, Eds., pp. 743–752, Springer-Verlag Berlin Heidelberg, Heidelberg, Germany, 2007.
- [9] M. Lippi, M. Bertini, and P. Frasconi, "Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 871–882, 2013.
- [10] M. G. Karlaftis and E. I. Vlahogianni, "Statistical methods versus neural networks in transportation research: differences, similarities and some insights," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 3, pp. 387–399, 2011.
- [11] V. Topuz, *Hourly Traffic Flow Predictions by Different ANN Models, Urban Transport and Hybrid Vehicles*, IntechOpen Limited, London UK, 2010.
- [12] Y. Xie and Y. Zhang, "A wavelet network model for short-term traffic volume forecasting," *Journal of Intelligent Transportation Systems*, vol. 10, no. 3, pp. 141–150, 2006.
- [13] W. Yu, T. Du, and W. Zhang, "Short-time traffic flow prediction using fuzzy wavelet neural network based on master-slave PSO," in *Proceedings of 2008 Fourth International Conference on Natural Computation (ICNC)*, pp. 321–325, Jinan, China, October 2008.
- [14] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.
- [15] H. Huang, Q. Tang, and Z. Liu, "Adaptive correction forecasting approach for urban traffic flow based on fuzzy c-mean clustering and advanced neural network," *Journal of Applied Mathematics*, vol. 2013, Article ID 195824, 7 pages, 2013.
- [16] E. Dogan and A. Akgungor, "Forecasting highway casualties under the effect of railway development policy in Turkey using artificial neural networks," *Neural Computing and Applications*, vol. 22, no. 5, pp. 869–877, 2013.
- [17] Y. Erzin and T. Oktay Gul, "The use of neural networks for the prediction of the settlement of one-way footings on cohesionless soils based on standard penetration test," *Neural Computing and Applications*, vol. 24, no. 4, pp. 891–900, 2014.
- [18] A. L. Santana, G. B. Conde, L. P. Rego et al., "PREDICT-Decision support system for load forecasting and inference:

- a new undertaking for Brazilian power suppliers,” *International Journal of Electrical Power and Electricity Systems*, vol. 38, no. 1, pp. 33–45, 2012.
- [19] M. Buragohain and C. Mahanta, “A novel approach for ANFIS modeling based on full factorial design,” *Applied Soft Computing*, vol. 8, no. 1, pp. 609–625, 2008.
- [20] D. Simon, “Biogeography-based optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008.
- [21] O. Bozorg-Haddad, S.-M. Hosseini-Moghari, and H. Loaiciga, “Biogeography-based optimization algorithm for optimal operation of reservoir systems,” *Journal of Water Resources Planning and Management*, vol. 142, article 04015034, 2015.
- [22] M. Tamjidy, S. Paslar, B. T. Hang Tuah Baharudin, T. S. Hong, and M. K. A. Ariffin, “Biogeography based optimization (BBO) algorithm to minimise nonproductive time during hole-making process,” *International Journal of Production Research*, vol. 53, no. 6, pp. 1880–1894, 2015.
- [23] R. Cheng and Y. Bai, “A novel approach to fuzzy wavelet neural network modeling and optimization,” *Electrical Power and Energy Systems*, vol. 64, pp. 671–678, 2015.
- [24] N. Amjady and F. Keynia, “Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm,” *Energy*, vol. 34, pp. 46–57, 2009.
- [25] K. Antonios, A. Alexandridis, and D. Zaprani, “Wavelet neural networks: a practical guide,” *Neural Networks*, vol. 42, pp. 1–27, 2013.
- [26] S. Yilmaz and Y. Oysal, “Fuzzy wavelet neural network models for prediction and identification of dynamical systems,” in *IEEE Transactions on Neural Networks*, vol. 21, no. 10, pp. 1599–1609, 2010.
- [27] L. S. L. Linhares, J. M. Araújo Jr., F. M. U. Araújo, and T. Yoneyama, “A nonlinear system identification approach based on Fuzzy Wavelet Neural Network,” *Journal of Intelligent and Fuzzy Systems*, vol. 28, no. 1, pp. 225–235, 2015.
- [28] D. Simon, *A Probabilistic Analysis of a Simplified Biogeography Based Optimization Algorithm*, Massachusetts Institute of Technology Press, Cambridge, MA, USA, 2009, <http://academic.csuohio.edu/simond/bbo/simplified/bbosimplified.pdf>.
- [29] Q. Feng, S. Liu, J. Zhang, G. Yang, and L. Yong, “Improved biogeography-based optimization with random ring topology and Powell’s method,” *Applied Mathematical Modelling*, vol. 41, pp. 630–649, 2016.
- [30] Y. Zheng, X. Wu, H. Ling, and S. Chen, “A simplified biogeography-based optimization using a ring topology,” in *Advances in Swarm Intelligence. ICSI 2013. Lecture Notes in Computer Science*, Y. Tan, Y. Shi, and H. Mo, Eds., Springer, Berlin, Heidelberg, 2013.
- [31] M. Lohokare, S. S. Pattnaik, B. Panigrahi, and S. Das, “Accelerated biogeography-based optimization with neighborhood search for optimization,” *Applied Soft Computing*, vol. 13, no. 5, pp. 2318–2342, 2013.
- [32] S. Chandra and U. Kumar, “Effect of lane width on capacity under mixed traffic conditions in India,” *Journal of Transportation Engineering*, vol. 129, no. 2, pp. 155–160, 2003.
- [33] C. Patel and G. Joshi, “Equivalency factor using optimization for Indian mixed traffic condition,” *International Journal for Traffic and Transport Engineering*, vol. 5, no. 2, pp. 210–224, 2015.