Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2014, Article ID 910196, 10 pages
http://dx.doi.org/10.1155/2014/910196

Research Article

A Conjunction Method of Wavelet Transform-Particle Swarm Optimization-Support Vector Machine for Streamflow Forecasting

Fanping Zhang, Huichao Dai, and Deshan Tang

College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

Correspondence should be addressed to Huichao Dai; dai_huichao@263.net

Received 21 January 2014; Revised 27 April 2014; Accepted 27 April 2014; Published 22 May 2014

Academic Editor: Y. P. Li

Streamflow forecasting has an important role in water resource management and reservoir operation. Support vector machine (SVM) is an appropriate and suitable method for streamflow prediction due to its best versatility, robustness, and effectiveness. In this study, a wavelet transform particle swarm optimization support vector machine (WT-PSO-SVM) model is proposed and applied for streamflow time series prediction. Firstly, the streamflow time series were decomposed into various details (Ds) and an approximation (A_3) at three resolution levels (2^1-2^2-2^3) using Daubechies (db3) discrete wavelet. Correlation coefficients between each D subtime series and original monthly streamflow time series are calculated. Ds components with high correlation coefficients (D_3) are added to the approximation (A_3) as the input values of the SVM model. Secondly, the PSO is employed to select the optimal parameters, C, ε, and σ, of the SVM model. Finally, the WT-PSO-SVM models are trained and tested by the monthly streamflow time series of Tangnaihai Station located in Yellow River upper stream from January 1956 to December 2008. The test results indicate that the WT-PSO-SVM approach provide a superior alternative to the single SVM model for forecasting monthly streamflow in situations without formulating models for internal structure of the watershed.

1. Introduction

The accuracy of streamflow forecasting is a key factor for reservoir operation and water resource management. However, streamflow is one of the most complex and difficult elements of the hydrological cycle due to the complexity of the atmospheric process. The elements affecting streamflow forecasting precision include catchment, geomorphologic and climate characteristics, and so forth [1]. The process of streamflow is extremely complex due to the influence of these variables and their combinations. Therefore, there are many forecasting techniques that have been proposed for streamflow forecasting [2–4].

Among them, the most popular and widely known statistical method used in time series forecasting is autoregressive integrated moving average (ARIMA) model due to its superiority of forecasting capabilities and richness of information on time-related changes [5]. Several studies have shown that ARIMA can be trusted as a reliable model in water resources time series analysis [6]. For example, Lee and Tong [7] proposed a hybrid model for nonlinear time series forecasting by combining ARIMA and genetic programming and demonstrated the effectiveness of the proposed forecasting model. But the ARIMA models are a class of linear model and thus only suitable for capturing linear features of data time series [8]. In recent years, gray model, artificial neural network (ANN), and support vector machine (SVM) have been frequently used to predict the nonlinear time series and achieved good results [9–11]. For instance, Kişi [12] used three different ANN techniques, namely, feed forward neural networks, generalized regression neural networks and, radial basis ANN in one-month-ahead streamflow forecasting. However, there are some disadvantages of ANN due to its network structure, which is hard to determine and usually established using a trial-and-error approach [13].

Support vector machines (SVM) were suggested by Vapnik [14] as one of the soft computational techniques and are widely used for classification and regression based on

statistical learning theory (SLT). The basic idea of SVM for regression is to introduce a kernel function, map the input data into a high-dimensional feature space by a nonlinear mapping, and then perform linear regression in the feature space [5]. Currently, SVM were frequently applied in a number of different fields, such as fault diagnosis [16], pattern recognition [17], and classification [18]. In the hydrology context, SVM has been successfully applied to forecast the flood stage [19–21], to predict future water levels in Lake Erie [22], and to forecast discharges [23, 24]. Previous studies have indicated that SVM is an effective method for streamflow forecasting [5, 23–25].

More recently, the conjunction model of wavelet and SVM has drawn increasing interest and has displayed advantages over a single SVM model in terms of prediction accuracy. Wavelet analysis (WA) is an advanced method proposed by Morlet et al. [26] in signal processing and has attracted much attention due to its ability to reveal simultaneously both spectral and temporal information within one signal [27]. The application of WA in the areas of hydrology and water resource research mainly includes these aspects: identification of hydrologic series deterministic components such as trend, periods, and change points [28–31]; wavelet denoising in hydrologic series [30, 31]; and hydrologic series simulation and prediction based on wavelet [27, 32, 33]. Wavelet analysis can be used to decompose an observed time series (such as streamflow time series) into various components so that the new time series can be used as inputs for SVM models [34].

SVM implements the principle of structure risk minimization in place of experiential risk minimization, which makes it have excellent generalization ability in the situation of small sample. However, the practicability of SVM is affected by the difficulty of selecting appropriate SVM parameters [35]. At present, the most common parameters selection method for SVM is the cross validation method but it is time-consuming [36]. Recently, some intelligent algorithms have been applied for parameters selecting. Compared with cross validation, genetic algorithm (GA) is less time-consuming and can obtain the optimal solution well, but the operation of genetic algorithm is difficult with the steps of choosing, crossover, and mutation for different optimal problems [37]. As a new global optimizing algorithm, particle swarm optimization (PSO), proposed by Kennedy and Eberhart in 1995, is based on swarm intelligent by generating a random decision variable set called "particles" [35]. PSO is a versatile algorithm and can be used to solve different optimizing problems. In recent years, because of the best global searching ability and the simple implementing procedure, PSO has been successfully applied for function optimization [38], data mining [39], and other engineering optimization problems [15, 40] and achieved good results. Therefore, the PSO can be applied to optimize the parameters of SVM model for streamflow forecasting in this paper.

This paper is organized as follows. Section 2 introduces the principle theory of wavelet analysis, parameter selection method of SVM models based on PSO, and SVM regression forecasting model. The study area and streamflow time series analysis are introduced in Section 3. The forecasting results of the conjunction model with the real streamflow time series data sets from Tangnaihai hydrology station in China are analyzed in Section 4. Finally, the conclusion is presented in Section 5.

2. Methodology

2.1. Support Vector Machine (SVM). The basic idea of SVM for regression is to introduce a kernel function, map the input data into a high-dimensional feature space by a nonlinear mapping, and then perform linear regression in the feature space [5]. Supposing that there is a training dataset $D = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_n, \mathbf{y}_n)\} \in R^p \times R$, $\mathbf{x}$ is the input vector, $\mathbf{y}$ is the excepted output, n is the number of data, and p is the total number of data patterns. By nonlinear mapping function Φ, $\mathbf{x}$ is mapped into a feature space in which a linear estimate function is defined as

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\omega}) = \langle \boldsymbol{\omega}, \Phi(\mathbf{x}) \rangle + b; \quad \boldsymbol{\omega}, \mathbf{x} \in R^p, \ b \in R, \tag{1}$$

where $\Phi(\mathbf{x})$ represents the high-dimensional feature spaces, which is nonlinearly mapped from the input space $\mathbf{x}$; ω and b are coefficients that have to be estimated from the input data. By introducing the slack variables ξ and ξ^* and following the regularization theory, parameters ω and b are estimated by minimizing the cost function

$$\min J(\omega, \xi) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{2}$$

subject to the constraints:

$$\begin{aligned} y_i - \langle \omega, x_i \rangle - b &\le \varepsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i &\le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\ge 0. \end{aligned} \tag{3}$$

The first term $(1/2)\|\omega\|^2$ is weight vector norm; C is referred to as the regularized constraint determining the tradeoff between the empirical error and the regularized term; and ε is the insensitive loss function.

By using Lagrange multiplier techniques, the minimization of (2) leads to the following dual-optimization problem:

$$\begin{aligned} \max \quad & W(\alpha_i, \alpha_i^*) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(\mathbf{x}_i, \mathbf{x}_j)\right] \\ & - \varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) \\ \text{s.t.} \quad & \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0; \quad \alpha_i, \alpha_i^* \in [0, C], \end{aligned} \tag{4}$$

where (α_i^*, α_i) are coefficients determined by training and $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function which can be expressed as inner product:

$$K(\mathbf{x}_i, \mathbf{x}) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}). \tag{5}$$

The decision function takes the form

$$f(x) = \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) K(\mathbf{x}, \mathbf{x}_i) + b. \quad (6)$$

The selection of an appropriate kernel function plays an important role in SVM regression since the kernel function defines the feature space. Gaussian radial basis function kernel has received significant attention from the machine learning community. Gaussian radial basis function (RBF) kernel is defined as

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2}\right). \quad (7)$$

Here σ is the kernel parameter.

2.2. Wavelet Analysis. In wavelet analysis, the signals are analyzed in both the time and the frequency domain by decomposing the original signals in different frequency bands using wavelet functions. The wavelet transform (WT) uses the scalable windowing technique for analyzing local variation in the time series [41]. WT provides useful decompositions of original time series, so that wavelet-transformed data improve the ability of a forecasting model by capturing useful information on various resolution levels [42]. The time series data are preprocessed using wavelet transformation techniques to obtain decomposed wavelet coefficients that are used as inputs in the forecasting models.

The basic objective of WT is to achieve a complete timescale representation of localized and transient phenomena occurring at different timescales [4, 43]. The continuous wavelet transform is defined as the sum over all time of the signal multiplied by scale and shifted versions of wavelet function ψ:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi\left(\frac{x-b}{a}\right) dx, \quad (8)$$

where a is a scale parameter; b is a position parameter; and ψ corresponds to the complex conjugate. The coefficient plots of the continuous wavelet transform are precisely the timescale view of the signal. However, calculating wavelet coefficients at every possible scale is time-consuming and generates large amount of information. Thus, the use of the continuous wavelet transform for forecasting is not practically possible.

In hydrology, observed hydrologic series are often expressed as discrete series, so the discrete wavelet transform is usually employed to decompose a hydrologic series into a set of coefficients and subsignals under different scales, and then guide other time series analyses [44]. The DWT is defined as the following form:

$$f(t) = \sum C_{jo,k} \phi_{jo,k}(t) + \sum_{j>jo} \sum \omega_{j,k} 2^{j/2} \psi\left(2^j t - k\right), \quad (9)$$

where j is the dilation or level index, k is the translation or scaling index, and $\phi_{jo,k}$ is a scaling function of coarse scale coefficients. $C_{jo,k}, \omega_{j,k}$, is the scaling function of detailed (fine scale) coefficients and all functions of $\psi(2^j t - k)$ are orthonormal.

The original time series are decomposed into various details (Ds) and an approximation (As) at different resolution levels using DWT. The approximations are the high-scale, low frequency components of the signal and the details are the low-scale, high frequency components. Normally, the low frequency component of the signal (A) is the most important part which demonstrates the signal identity [45]. The choice of wavelet type is an important issue. The Daubechies wavelets are one of the widely used in wavelet family, which are written dbN, where db is the "surname" and N is the order of the wavelet [46]. Daubechies wavelets exhibit good tradeoff between parsimony and information richness [34], so in this study the Daubechies wavelets were employed as the mother wavelet to decompose the time series.

2.3. Parameters Selection of SVM Based on PSO

2.3.1. The Principle of PSO. PSO, deriving from the research for the movement of organisms in a bird flocking or fish schooling, performs searches using a population (called swarm) of individuals (called particles) that are updated from iteration to iteration [47, 48]. An equation (velocity update) controls the swarm in moving around the search space seeking the optimum state. In each iteration, the algorithm saves the local optimum and compares it with the global (best yet) optimum values. Definitely the criteria for being chosen as an optimum state depend on the fitness of the objective function. Candidate solutions (decision variables) of any particle calculate and remember its own fitness. The position of any particle accelerated towards the global best position by using (10) and (11) [49]. In any search step t, the ith particle is used to update its candidate solution's current position $x_{i,j}(t)$ by using local best $p_{i,j}(t)$ and best $p_{g,j}(t)$ position achieved yet. Consider the following:

$$v_{i,j}(t+1) = \omega v_{i,j}(t) + c_1 r_1 \left[p_{i,j} - x_{i,j}(t)\right] + c_2 r_2 \left[p_{g,j} - x_{i,j}(t)\right] \quad (10)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1), \quad j = 1, 2, \ldots, d, \quad (11)$$

where $v_{i,j}$ is velocity measures for particles; ω is inertial weight controlling velocity direction; c_1 and c_2 are acceleration coefficients; r_1 and r_2 are random numbers uniformly distributed between $[0, 1]$. $x_{i,j}$ is the position of any particle.

2.3.2. Parameters Selection of SVM Based on PSO. In the SVM regression model, three parameters, namely, C, ε, and σ, should be identified before forecasting. Therefore, PSO algorithm is used for optimizing the SVM parameters. The process of optimizing the SVM parameters with PSO is presented in Figure 1 and the steps are described as follows [47].

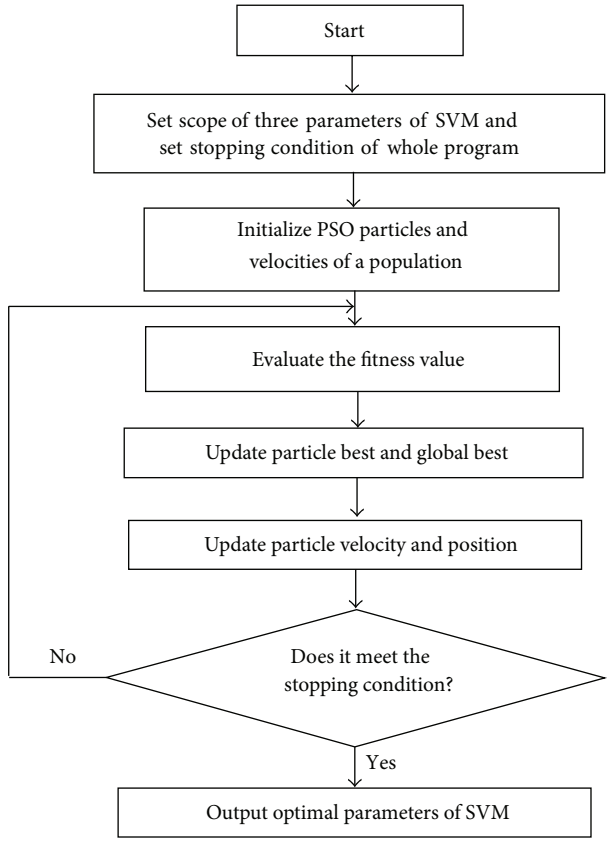

Figure 1: The process of optimizing the SVM parameters with PSO.

Step 1. Initialization: consider randomly initial particles and velocities of a population (every particle contains three variables, namely, C, ε, and σ).

Step 2. Fitness evaluation: the fitness function of PSO is shown as follows:

$$\text{fitness} = \eta_{\text{MAPE}} = \frac{1}{m}\sum_{i=1}^{m}\frac{\left|y_i - \widehat{y}_i\right|}{y_i}, \tag{12}$$

where η_{MAPE} is the mean absolute percentage error; y_i is the actual value and $\widehat{y}_i$ is the predicted value; and m is the number of subsets. The solution with a smaller η_{MAPE} has a smaller fitness value.

Step 3. Update global and personal best according to fitness evaluation results.

Step 4. Calculation of velocity: particle flies toward a new position by calculating the velocity of position change. Velocity of each particle is calculated by (10).

Step 5. Update position value: each particle moves to its next position according to (11).

Step 6. Termination: repeat the same procedures from Step 2 to Step 5 until stopping conditions are satisfied.

2.4. Model Evaluation. It is essential to evaluate the performance of the models by employing appropriate methods. In this study, the performance of the models is evaluated by the indexes of the correlation coefficients (R), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute relative error (MARE). These indexes are respectively defined as follows.

Correlation coefficients (R):

$$R = \frac{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)\left(\widehat{y}_i - \overline{\widehat{y}}_i\right)}{\sqrt{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2 \sum_{i=1}^{n}\left(\widehat{y}_i - \overline{\widehat{y}}_i\right)^2}}. \tag{13}$$

Root mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2}. \tag{14}$$

Mean absolute error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y}_i\right|. \tag{15}$$

Mean absolute relative error (MARE):

$$\text{MARE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right|; \tag{16}$$

where y_i stands for the observed data and $\widehat{y}_i$ stands for the forecasting data. n is number of the data.

3. Study Area and Data Analysis

In this study, we examined the data obtained from the monthly streamflow of the Tangnaihai Hydrological station located at the upper catchment of the Yellow River in Qinghai Province of China. Tangnaihai Hydrological station is the upstream hydrological station of Longyang Gorge Reservoir which is the largest regulating reservoir in the upper catchment of the Yellow River, so there are few human impacts disordering streamflow regular. Location of Tangnaihai Hydrological station is shown in Figure 2. The Yellow River catchment covers an area of 95,000 km^2.

The monthly streamflow time series of Tangnaihai Hydrological station, consisting of 636 monthly records (January 1956 to December 2008), are used in this study. The dataset was split up into two parts: training and testing, where the first dataset consisting of 536 monthly records (January 1956 to August 2000) was used for training, while the final dataset contains 100 monthly records (September 2000 to December 2008). Training data were used exclusively for model development and testing data were used to measure the performance of the model on untrained data. The testing set was also used to evaluate the forecasting ability of the model and to compare the proposed model with others.

4. Results Analysis

4.1. Wavelet Decomposition of Streamflow Time Series. The WT-PSO-SVM model structure is shown in Figure 3. For the SVM model inputs, the original time series are decomposed into subseries with an approximation (As) with low frequency and details ($D_1, D_2, \ldots, D$s) with high frequency by Daubechies DWT algorithm.

The optimal decomposition level of the streamflow time series in wavelet analysis plays an important role in preserving the information and reducing the distortion of the datasets [4]. The number of decomposition levels controls the streamflow approximation in the data. The general rule for the appropriate decomposition levels is that the largest levels should be shorter than the size of the testing data [50].

In this case, the largest scales were chosen as three for the Tangnaihai station streamflow time series. Therefore, the flow data sets are decomposed into various details (Ds) and an approximation (A_3) at three resolution levels (2^1-2^2-2^3) using db3 DWT shown in Figure 4. The new decomposed subseries present variations of the original time series on different periods. MATLAB codes were developed using its library functions to perform wavelet decomposition of the time series data. The correlation coefficients between each D subtime series and original monthly streamflow time series are given in Table 1 for the Tangnaihai station. In the table, the D_{t-1} and Q_t denote the D subtime series at time $t-1$ and measured streamflow at time t, respectively. These correlation values provide information for the determination of effective wavelet components on streamflow. It can be seen from Table 1 that the D_3 has the highest correlation ($R = 0.259$) among Ds. The average correlation between A and Q_t is 0.363. According to the correlation analysis between Ds and the original current streamflow data, the effective component (D_3) is selected. Then, the new series obtained by adding the effective D_3 and approximation component are used as an input combination to the SVM model.

Table 1: The correlation coefficients between subtime series and original streamflow data.

Discrete wavelet components	Correlations			
	D_{t-1} and Q_t	D_{t-2} and Q_t	D_{t-3} and Q_t	Average
D_1	0.103	−0.003	0.105	0.070
D_2	−0.081	0.084	0.328	0.164
D_3	−0.023	0.295	0.459	0.259
A	0.211	0.390	0.488	0.363

4.2. Parameters Selection of SVM Based on PSO. In this study, RBF is employed as kernel function of SVM forecasting model, so three parameters, namely, balance parameter C, insensitive parameter ε, and kernel function, parameter σ should be selected. Some researchers have shown that different kernel functions have little impact on performance, but kernel function parameter σ is a key factor affecting performance of SVM. Among three parameters, σ precisely defines structure of highly dimensional space, so it controls complexity of ultimate solution; C determines complexity of model and punishment level of fitting deviation; ε indicates forecasting model's expectation on estimating functions' error of sample data, and the larger ε, the less support vector number and more sparse solution expression. But large ε can also reduce accuracy of SVM forecasting model.

For monthly streamflow time series $X = \{x_1, x_2, \ldots, x_N\}$, the flow data at time $i + p$ is predicted based on the previous flow data. The general expression is shown as follows:

$$x_{i+p} = f\left(x_{i+1}, x_{i+2}, \ldots, x_{i+p-1}\right), \tag{17}$$

where f is a nonlinear function indicating relationship of monthly streamflow time series; x_i is streamflow data at time i, $i = 1, 2, \ldots, N$; and p is the forecasting step (month), which is set as 3 in this paper. $N - p$ monthly streamflow time series data sets are used for training and testing SVM forecasting model. The particle swarm optimization is employed to optimize the best parameters set (σ, C, ε) of SVM model [51].

For the Tangnaihai station, three input combinations based on preceding monthly streamflows are evaluated to estimate current streamflow value. The input combinations evaluated in the study are as follows: (i) Q_{t-1}, Q_{t-2}, and Q_{t-3}; (ii) $Q_{t-1}(A)$, $Q_{t-2}(A)$, and $Q_{t-3}(A)$; (iii) $Q_{t-1}(A+D_3)$, $Q_{t-2}(A+D_3)$, and $Q_{t-3}(A+D_3)$. In all cases, the output is the discharge Q_t for the current month.

In the training stage, firstly the parameters σ, C, and ε of SVM model are optimized by PSO, the validation error is measured by (12), and the adjusted parameters with minimum validation error are selected as the most appropriate parameters which are provided in Table 2. Then, the optimal parameters are utilized to train SVM and WSVM models.

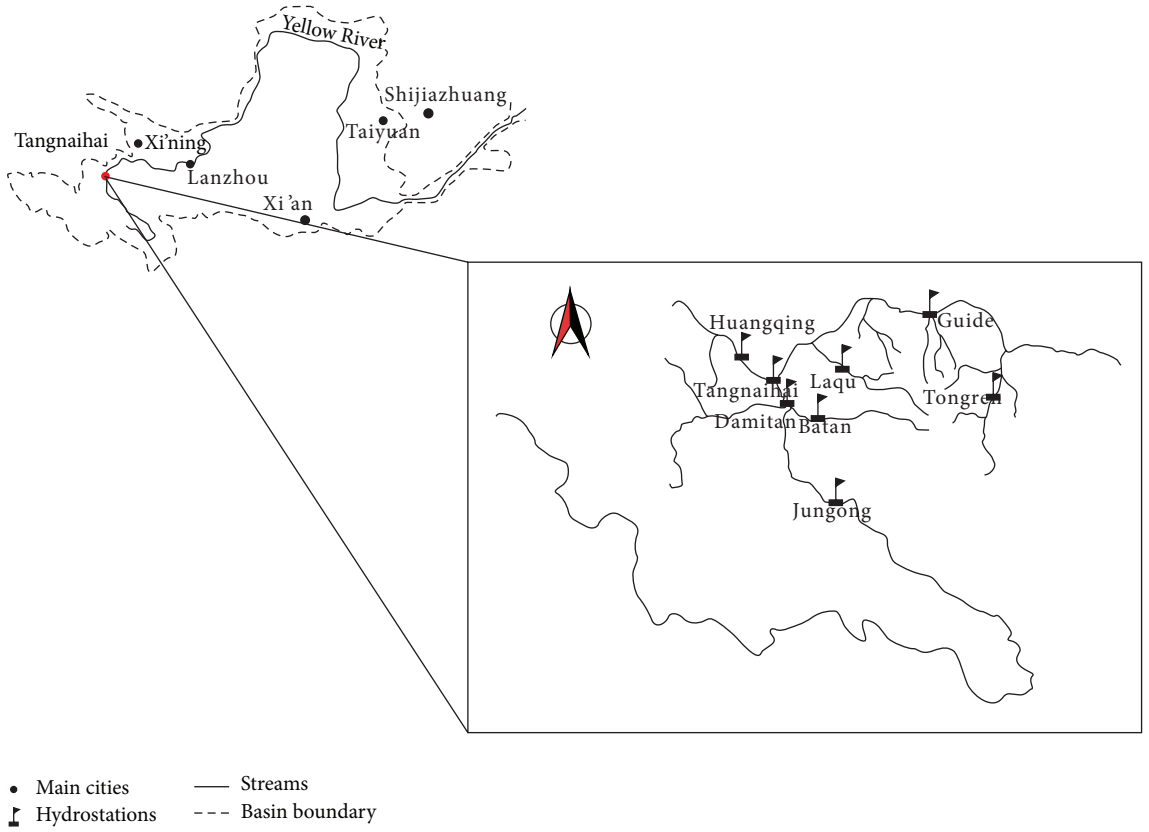

Figure 2: Location map of Yellow River and Tangnaihai station.

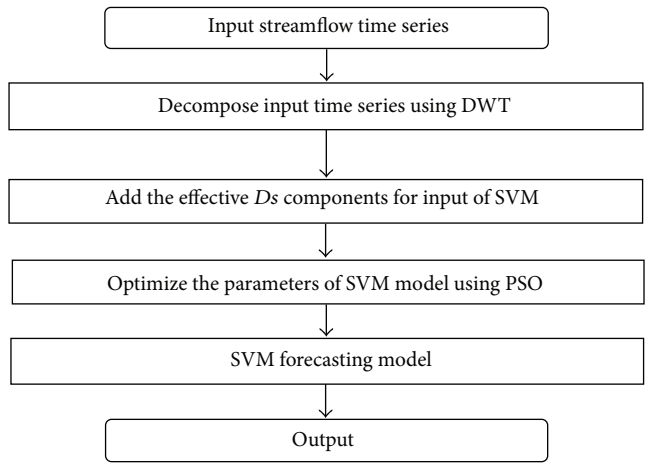

Figure 3: The WA-LSSVM-PSO model structure.

Table 2: Optimal parameters of SVM and WSVM models.

Model	Input variables	Output	Parameters		
			σ	C	ε
(i) SVM	(i) Q_{t-1}, Q_{t-2}, and Q_{t-3}	Q_t	0.90156	2.3149	0.01
(ii) $WSVM_1$	(ii) $Q_{t-1}(A)$, $Q_{t-2}(A)$, and $Q_{t-3}(A)$		0.01	46.8765	0.01
(iii) $WSVM_2$	(iii) $Q_{t-1}(A + D_3)$, $Q_{t-2}(A + D_3)$, and $Q_{t-3}(A + D_3)$		28.4876	6.0929	0.01

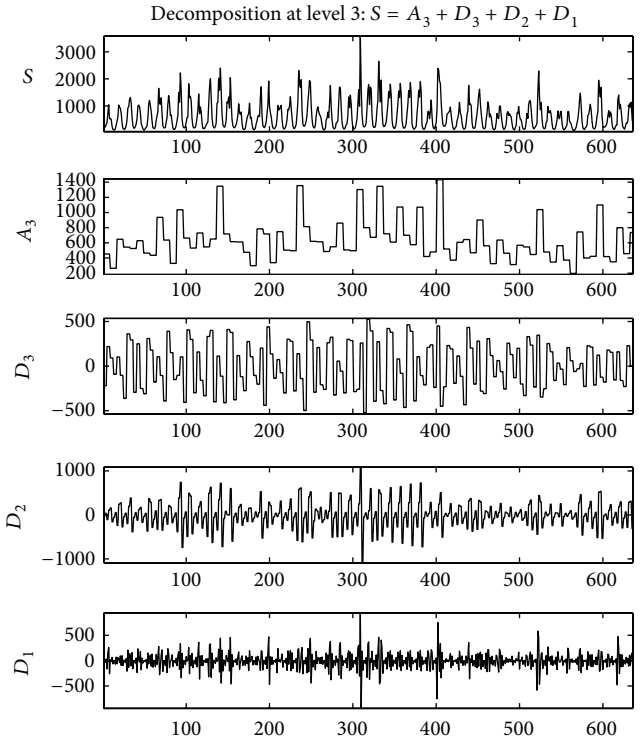

Figure 4: Decomposed wavelet subtime series components (Ds) of runoff data of Tangnaihai station.

Before the training process begins, data normalization is often performed. Runoff time series data was normalized in the range [0, 1] by the following equation:

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \tag{18}$$

where y_i represents the normalized data, while x_i is the actual observation value and $x_{\max}$, $x_{\min}$, respectively, represent the maximum and minimum value among the actual observation values.

4.3. Streamflow Forecasting Based on WT-PSO-SVM. As shown in Table 2, the optimal parameters for SVM (input model (i)) and WSVM (input models (ii) and (iii)) models are $\sigma_1 = 0.90156$, $C_1 = 2.3149$, and $\varepsilon_1 = 0.01$; $\sigma_2 = 0.01$, $C_2 = 46.8765$, and $\varepsilon_2 = 0.01$; and $\sigma_3 = 28.4876$, $C_3 = 6.0929$, and $\varepsilon_3 = 0.01$. The optimal parameters are used to examine the accuracy of the SVM and WSVM forecasting models with the testing data sets. Table 3 shows the performance results obtained in the training and testing periods of the SVM and WSVM models for Tangnaihai station.

Results obtained from the three models for 3-month-advance flow forecasting at Tangnaihai station are presented in Table 3. Performance of the three models was compared by evaluating indexes of R, RMSE, MAE, and MARE. It is observed that the forecasting accuracy of the $WSVM_2$ model was much better than that of corresponding SVM and $WSVM_1$ models. Comparing the forecasting results of SVM and $WSVM_1$, the R, RMSE, MAE, and MARE in the testing period were 0.768–0.613, 317.035–331.420 (m^3/s), 247.480–256.368 (m^3/s), and 17.81–23.49 (%), respectively. It is observed that SVM is more superior than WSVM1 because some useful details (Ds) of original streamflow series were eliminated in the model $WSVM_1$. Comparing the forecasting results of SVM and $WSVM_2$, the values of R, RMSE, MAE, and MARE in the testing period were 0.768–0.806, 317.035–243.268, 247.480–173.20, and 17.81–11.52, respectively. It is obvious from Table 3 and Figure 5 that the WSVM2 performs better than the SVM model. These results indicated that D_3 was an effective component for the runoff series and the D_1 and D_2 were the noise that should be eliminated before streamflow forecasting. Wavelet transform is a necessary process of data preprocessing for improving predicting accuracy.

5. Conclusion

This study developed a WT-PSO-SVM hybrid model to forecast monthly streamflow. The WT-PSO-SVM model was obtained by combining three methods, discrete wavelet transform-particle swarm optimization, and support vector machine regression. The combined model integrated the advantages of best versatility, robustness and effectiveness of SVM, the best global searching ability and the simple implementing procedure of PSO for parameter selection, and the ability of WT to reveal simultaneously both spectral and temporal information within one signal. This hybrid approach was successfully applied to simulate streamflow time series of Tangnaihai Hydrology station in the Yellow River.

The streamflow time series were decomposed into various details (Ds) and an approximation (A_3) at three resolution levels (2^1-2^2-2^3) by using db3 DWT of the wavelet function of Daubechies 3 (db3). The correlation coefficients between each D subtime series and original monthly streamflow time series were calculated. Ds(D_3) components with high correlation coefficients ($R = 0.259$) were added to the approximation (A) as the input values of SVM model. The input combinations evaluated in the study are as follows: (i) Q_{t-1}, Q_{t-2}, and Q_{t-3}; (ii) $Q_{t-1}(A)$, $Q_{t-2}(A)$, and $Q_{t-3}(A)$; (iii) $Q_{t-1}(A + D_3)$, $Q_{t-2}(A + D_3)$, and $Q_{t-3}(A + D_3)$. The PSO was employed to select the optimal parameters, C, ε, and σ, of the three input models which were used to test the accuracy of the SVM model. Three different input combinations of SVM predicting results indicated that the discrete wavelet transform can significantly increase the accuracy of the SVM model in forecasting monthly streamflow. In addition, particle swarm optimization can determine suitable parameters to forecast streamflow as well. Predicting accuracy was evaluated by indexes of R, RMSE, MAE, and MARE. At the Tangnahai station, the best predictions belong to $WSVM_2$ model. $WSVM_2$ model increased the prediction R by 0.038 and 0.193 with respect to the SVM and $WSVM_1$ models and reduced MARE by 6.29% and 11.97%, respectively, in the testing period. These results indicated that D_3 was an effective component for the runoff series and the D_1 and D_2 were the noise that should be eliminated before streamflow forecasting. Wavelet transform

Table 3: Forecasting performance indexes of SVM and WSVM models in Tangnaihai station.

Model	Training				Testing			
	R	RMSE (m^3/s)	MAE (m^3/s)	MARE (%)	R	RMSE (m^3/s)	MAE (m^3/s)	MARE (%)
(i) SVM	0.853	287.315	196.524	14.33	0.768	317.035	247.480	17.81
(ii) $WSVM_1$	0.548	306.554	217.883	18.24	0.613	331.420	256.368	23.49
(iii) $WSVM_2$	0.808	201.784	141.730	7.46	0.806	243.268	173.20	11.52

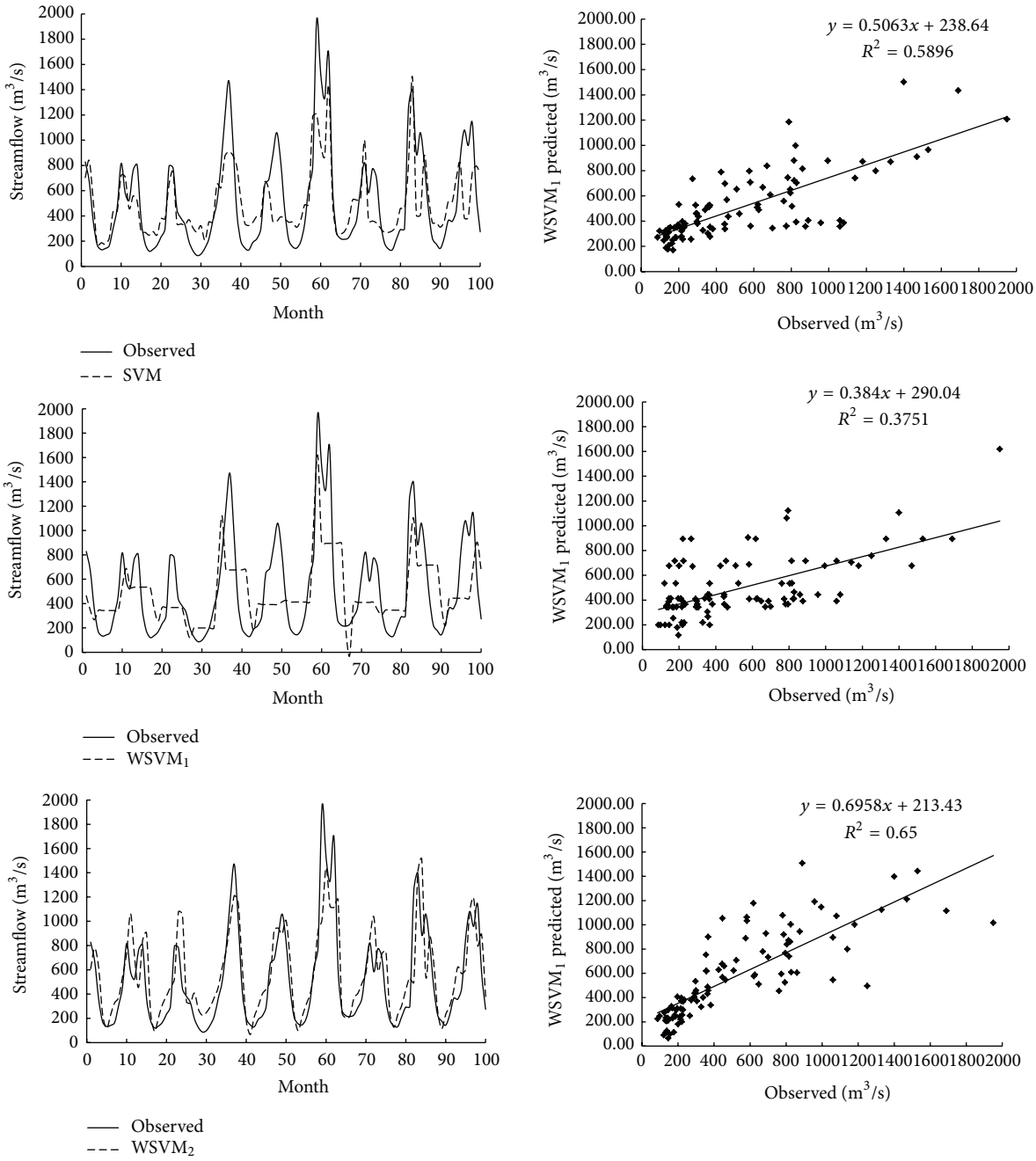

Figure 5: Predicted and observed streamflow in the testing period by SVM, $WSVM_1$, and $WSVM_2$.

is a necessary process of data preprocessing for improving predicting accuracy. The test results indicated that PSO-WT-SVM approach provides a superior alternative to the single SVM model for forecasting monthly streamflow in situations, without formulating models for the internal structure of the watershed.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Key Basic Research Program of China (973 Program) (2012CB417006), the National Science Fund for Distinguished Young Scholars (Grant no. 50925932), National Natural Science Foundation of China (Grant no. 50979050), and the College Graduate Research and Innovation Projects of Jiangsu Province of China (Grant no. 2013B20414). The authors wish to thank the Yellow River Conservancy Commission for providing the requisite meteorological data.

References

[1] A. Jain and A. M. Kumar, "Hybrid neural network models for hydrologic time series forecasting," *Applied Soft Computing Journal*, vol. 7, no. 2, pp. 585–592, 2007.

[2] O. Kisi and M. Cimen, "A wavelet-support vector machine conjunction model for monthly streamflow forecasting," *Journal of Hydrology*, vol. 399, no. 1-2, pp. 132–140, 2011.

[3] J. Adamowski and K. Sun, "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds," *Journal of Hydrology*, vol. 390, no. 1-2, pp. 85–91, 2010.

[4] N. Pramanik, R. K. Panda, and A. Singh, "Daily river flow forecasting using wavelet ANN hybrid models," *Journal of Hydroinformatics*, vol. 13, no. 1, pp. 49–63, 2011.

[5] A. Shabri and Suhartono, "Streamflow forecasting using least-squares support vector machines," *Hydrological Sciences Journal*, vol. 57, no. 7, pp. 1275–1293, 2012.

[6] J. Adamowski and C. Karapataki, "Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms," *Journal of Hydrologic Engineering*, vol. 15, no. 10, pp. 729–743, 2010.

[7] Y.-S. Lee and L.-I. Tong, "Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 66–72, 2011.

[8] P. G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[9] H. V. Trivedi and J. K. Singh, "Application of grey system theory in the development of a runoff prediction model," *Biosystems Engineering*, vol. 92, no. 4, pp. 521–526, 2005.

[10] N. Vivekanandan, "Prediction of annual runoff using artificial neural network and regression approaches," *Mausam*, vol. 62, no. 1, pp. 11–20, 2011.

[11] G. Yu, J. Yang, and Z. Xia, "The prediction model of chaotic series based on support vector machine and its application to runoff," *Advanced Materials Research*, vol. 255–260, pp. 3594–3599, 2011.

[12] Ö. Kişi, "River flow forecasting and estimation using different artificial neural network techniques," *Hydrology Research*, vol. 39, no. 1, pp. 27–40, 2008.

[13] Ö. Kişi, "River flow modeling using artificial neural networks," *Journal of Hydrologic Engineering*, vol. 9, no. 1, pp. 60–63, 2004.

[14] V. Vapnik, "The nature of statistical learning theory," *Data Mining and Knowledge Discovery*, vol. 6, pp. 1–47, 1995.

[15] M. H. Afshar, "Large scale reservoir operation by constrained particle swarm optimization algorithms," *Journal of Hydro-Environment Research*, vol. 6, no. 1, pp. 75–87, 2012.

[16] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2560–2574, 2007.

[17] M. Zhu, K. Fu, X. Huang, S. Wu, and W. Jin, "A novel recognition approach for radar emitter signals based on on-line independent support vector machines," *Advances in Computer Science and Its Applications*, vol. 2, no. 3, pp. 390–396, 2013.

[18] E. Carrizosa and D. Romero Morales, "Supervised classification and mathematical optimization," *Computers & Operations Research*, vol. 40, no. 1, pp. 150–165, 2013.

[19] D. Han, L. Chan, and N. Zhu, "Flood forecasting using support vector machines," *Journal of Hydroinformatics*, vol. 9, no. 4, pp. 267–276, 2007.

[20] P.-S. Yu, S.-T. Chen, and I.-F. Chang, "Support vector regression for real-time flood stage forecasting," *Journal of Hydrology*, vol. 328, no. 3-4, pp. 704–716, 2006.

[21] C. Sivapragasam and S.-Y. Liong, "Flow categorization model for improving forecasting," *Nordic Hydrology*, vol. 36, no. 1, pp. 37–48, 2005.

[22] M. S. Khan and P. Coulibaly, "Application of support vector machine in lake water level prediction," *Journal of Hydrologic Engineering*, vol. 11, no. 3, pp. 199–205, 2006.

[23] J.-Y. Lin, C.-T. Cheng, and K.-W. Chau, "Using support vector machines for long-term discharge prediction," *Hydrological Sciences Journal*, vol. 51, no. 4, pp. 599–612, 2006.

[24] W.-C. Wang, K.-W. Chau, C.-T. Cheng, and L. Qiu, "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," *Journal of Hydrology*, vol. 374, no. 3-4, pp. 294–306, 2009.

[25] T. Asefa, M. Kemblowski, M. McKee, and A. Khalil, "Multi-time scale stream flow predictions: the support vector machines approach," *Journal of Hydrology*, vol. 318, no. 1-4, pp. 7–16, 2006.

[26] J. Morlet, G. Arens, E. Fourgeau, and D. Glard, "Wave propagation and sampling theory—part I: complex signal and scattering in multilayered media," *Geophysics*, vol. 47, no. 2, pp. 203–221, 1982.

[27] V. Nourani, M. T. Alami, and M. H. Aminfar, "A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 3, pp. 466–472, 2009.

[28] L. C. Smith, D. L. Turcotte, and B. L. Isacks, "Stream flow characterization and feature detection using a discrete wavelet transform," *Hydrological Processes*, vol. 12, no. 2, pp. 233–249, 1998.

[29] T. Partal and M. Küçük, "Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey)," *Physics and Chemistry of the Earth*, vol. 31, no. 18, pp. 1189–1200, 2006.

[30] Y.-F. Sang, D. Wang, J.-C. Wu, Q.-P. Zhu, and L. Wang, "The relation between periods' identification and noises in hydrologic series data," *Journal of Hydrology*, vol. 368, no. 1-4, pp. 165–177, 2009.

[31] Y.-F. Sang, D. Wang, J.-C. Wu, Q.-P. Zhu, and L. Wang, "Entropy-based wavelet de-noising method for time series analysis," *Entropy*, vol. 11, no. 4, pp. 1123–1147, 2009.

[32] T. Partal and Ö. Kişi, "Wavelet and neuro-fuzzy conjunction model for precipitation forecasting," *Journal of Hydrology*, vol. 342, no. 1-2, pp. 199–212, 2007.

[33] L. Karthikeyan and D. Nagesh Kumar, "Predictability of nonstationary time series using wavelet and EMD based ARMA models," *Journal of Hydrology*, vol. 502, pp. 103–119, 2013.

[34] D. Benaouda, F. Murtagh, J.-L. Starck, and O. Renaud, "Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting," *Neurocomputing*, vol. 70, no. 1–3, pp. 139–154, 2006.

[35] S.-W. Lin, K.-C. Ying, S.-C. Chen, and Z.-J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Systems with Applications*, vol. 35, no. 4, pp. 1817–1824, 2008.

[36] X. Li, S.-D. Yang, and J.-X. Qi, "New support vector machine optimized by improved particle swarm optimization and its application," *Journal of Central South University of Technology*, vol. 13, no. 5, pp. 568–572, 2006.

[37] P.-F. Pai and W.-C. Hong, "Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms," *Electric Power Systems Research*, vol. 74, no. 3, pp. 417–425, 2005.

[38] R. Brits, A. P. Engelbrecht, and F. van den Bergh, "Locating multiple optima using particle swarm optimization," *Applied Mathematics and Computation*, vol. 189, no. 2, pp. 1859–1883, 2007.

[39] K. Sarath and V. Ravi, "Association rule mining using binary particle swarm optimization," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 8, pp. 1832–1840, 2013.

[40] L. Ostadrahimi, M. A. Mariño, and A. Afshar, "Multi-reservoir operation rules: multi-swarm PSO-based optimization approach," *Water Resources Management*, vol. 26, no. 2, pp. 407–427, 2012.

[41] H. Abghari, H. Ahmadi, S. Besharat, and V. Rezaverdinejad, "Prediction of daily pan evaporation using wavelet neural networks," *Water Resources Management*, vol. 26, no. 12, pp. 3639–3652, 2012.

[42] Y. Chen, B. Yang, and J. Dong, "Time-series prediction using a local linear wavelet neural network," *Neurocomputing*, vol. 69, no. 4-6, pp. 449–465, 2006.

[43] D. Labat, R. Ababou, and A. Mangin, "Rainfall-runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses," *Journal of Hydrology*, vol. 238, no. 3-4, pp. 149–178, 2000.

[44] Y. F. Sang, "A practical guide to discrete wavelet decomposition of hydrologic time series," *Water Resources Management*, vol. 26, no. 11, pp. 3345–3365, 2012.

[45] R. V. Ramana, B. Krishna, S. Kumar, and N. Pandey, "Monthly rainfall prediction using wavelet neural network analysis," *Water Resources Management*, vol. 27, no. 10, pp. 3697–3711, 2013.

[46] S. Wei, H. Yang, J. Song, K. Abbaspour, and Z. Xu, "A wavelet-neural network hybrid modelling approach for estimating and predicting river monthly flows," *Hydrological Sciences Journal*, vol. 58, no. 2, pp. 374–389, 2013.

[47] S.-W. Fei, M.-J. Wang, Y.-B. Miao, J. Tu, and C.-L. Liu, "Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil," *Energy Conversion and Management*, vol. 50, no. 6, pp. 1604–1609, 2009.

[48] W. Pedrycz, B. J. Park, and N. J. Pizzi, "Identifying core sets of discriminatory features using particle swarm optimization," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4610–4616, 2009.

[49] Q. Wu, "A hybrid-forecasting model based on Gaussian support vector machine and chaotic particle swarm optimization," *Expert Systems with Applications*, vol. 37, no. 3, pp. 2388–2394, 2010.

[50] C. L. Wu, K. W. Chau, and Y. S. Li, "Methods to improve neural network performance in daily flows prediction," *Journal of Hydrology*, vol. 372, no. 1–4, pp. 80–93, 2009.

[51] Y.-B. Li, N. Zhang, and C.-B. Li, "Support vector machine forecasting method improved by chaotic particle swarm optimization and its application," *Journal of Central South University of Technology*, vol. 16, no. 3, pp. 478–481, 2009.