Engineering Applications of Artificial Intelligence 23 (2010) 586–594

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

A hybrid neural network and ARIMA model for water quality time series prediction

Durdu Ömer Faruk *

Water Resources Research Center, Adnan Menderes University, Aydın, Turkey

ARTICLE INFO

Article history:
Received 22 March 2009
Received in revised form
28 August 2009
Accepted 3 September 2009
Available online 28 October 2009

Keywords:
Neural networks
Backpropagation
ARIMA
Hybrid model
Time series

ABSTRACT

Accurate predictions of time series data have motivated the researchers to develop innovative models for water resources management. Time series data often contain both linear and nonlinear patterns. Therefore, neither ARIMA nor neural networks can be adequate in modeling and predicting time series data. The ARIMA model cannot deal with nonlinear relationships while the neural network model alone is not able to handle both linear and nonlinear patterns equally well. In the present study, a hybrid ARIMA and neural network model is proposed that is capable of exploiting the strengths of traditional time series approaches and artificial neural networks. The proposed approach consists of an ARIMA methodology and feed-forward, backpropagation network structure with an optimized conjugated training algorithm. The hybrid approach for time series prediction is tested using 108-month observations of water quality data, including water temperature, boron and dissolved oxygen, during 1996–2004 at Büyük Menderes river, Turkey. Specifically, the results from the hybrid model provide a robust modeling framework capable of capturing the nonlinear nature of the complex time series and thus producing more accurate predictions. The correlation coefficients between the hybrid model predicted values and observed data for boron, dissolved oxygen and water temperature are 0.902, 0.893, and 0.909, respectively, which are satisfactory in common model applications. Predicted water quality data from the hybrid model are compared with those from the ARIMA methodology and neural network architecture using the accuracy measures. Owing to its ability in recognizing time series patterns and nonlinear characteristics, the hybrid model provides much better accuracy over the ARIMA and neural network models for water quality predictions.

1. Introduction

The water quality is a subject of ongoing concern. Deterioration of water quality has initiated serious management efforts in many countries. Most acceptable ecological and water related decisions are difficult to make without careful modeling, prediction and analysis of river water quality for typical development scenarios. Accurate predictions of future phenomena are the lifeblood of optimal water resources management in a watershed. Computer science and statistics have improved modeling approaches for discovering patterns found in water resources time series data. Much effort has been devoted over the past several decades to the development and improvement of time series prediction models. One of the most important and widely used time series model is the autoregressive integrated moving average (ARIMA) model (Shahwan and Odening, 2007).

Over the past several years, nonlinear models have been proposed as alternative techniques, as i.e. in Pisoni et al. (2009), where nonlinear autoregressive models (NARX) and artificial neural networks (ANNs) have been applied for environmental prediction. Zhang and Hu (1998) summarized the different applications of neural networks for predictions. There are a number of studies in which neural networks are used to address water resources problems. Maier and Dandy (2000) reviewed recent papers dealing with the use of neural network models for the prediction and forecasting of water resources variables. Flood and Kartam (1994), Hassoun (1995) and Rojas (1996) have used feedforward networks with sigmoidal-type transfer functions for the prediction and forecasting of water resources variables. Chau (2006) has reviewed the development and current progress of the integration of artificial intelligence into water quality modeling. Hatzikos et al. (2005) utilized neural networks with active neurons as a modeling tool for the prediction of seawater quality indicators like water temperature, pH, dissolved oxygen (DO) and turbidity. Palani et al. (2008) demonstrated the application of ANNs to model the values of selected seawater quality variables, having the dynamic and complex processes hidden in the monitored data itself.

Most of the studies reported above were simple applications of using traditional time series approaches and ANNs. Many of the

* Tel.: +90 256 7727022.
E-mail address: odurdu@adu.edu.tr

0952-1976/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2009.09.015

real-life time series are extremely complex to be modeled using simple approaches especially when high accuracy is required. There have been several studies suggesting hybrid models, combining the ARIMA model and neural networks. Su et al. (1997) used a hybrid model to predict a time series of reliability data with growth trend. Their results showed that the hybrid model produced better predictions than either the ARIMA model or the neural network by itself. Zhang (2003) proposed a hybrid ARIMA and ANN model to take advantage of the two techniques and applied the proposed hybrid model to some real data sets. He concluded that the combined model can be an effective way to improving predictions achieved by either of the models used separately. Jain and Kumar (2006) proposed a hybrid approach for time series forecasting using monthly stream flow data at Colorado river. They indicated that the approach of combining the strengths of the conventional and ANN techniques provides a robust modeling framework capable of capturing the nonlinear nature of the complex time series and thus producing more accurate forecasts.

In the present paper, a hybrid approach, combining seasonal ARIMA model and neural network backpropagation model, is developed to predict water quality time series data. The use of combined models in water quality time series data could be complementary in capturing patterns of data sets and could improve the prediction accuracy. The motivation behind this hybrid approach is largely due to the fact that a water quality problem is often complex in nature and any individual model may not be able to capture different patterns equally well. The objectives of the present study are to: (1) develop a hybrid model, an ANN and an ARIMA model, to predict water quality time series data, (2) assess the performance of each modeling approach using observed data versus predicted data and (3) evaluate the predictive performance of hybrid model in comparison to ANN architecture and ARIMA model using accuracy measures.

2. Materials and methods

2.1. Study area and water quality data

The Büyük Menderes basin, 3.2% of the total area of the country, is located in southwest Turkey and it drains a total area of 25,000 km^2 into the Aegean Sea (Fig. 1). Annual rainfall ranges between 350 and 1000 mm and total mean annual evaporation, measured by Class A pans, is 2122 mm. Precipitation occurs

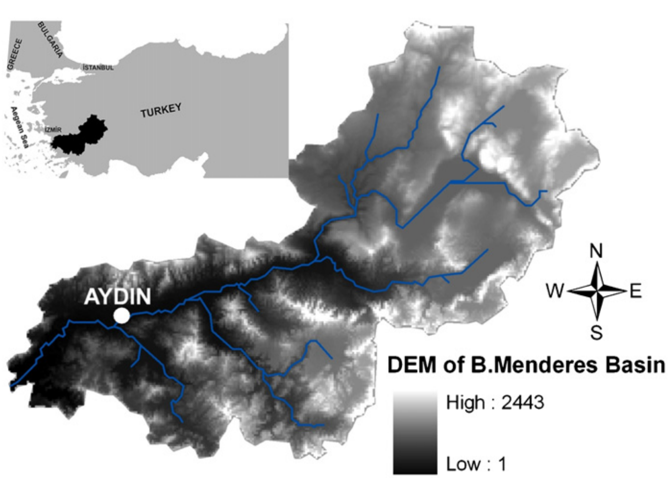

Fig. 1. Digital Elevation Model (DEM) of the Büyük Menderes basin and location of the measurement location used in this study.

mainly in the winters while during the summer irrigation period, there is very little rain. The main river of the basin is the Büyük Menderes river. The land use in the Büyük Menderes river basin is as follows: 40% agriculture, 45% forest and scrubland, 10% meadow and pasture, 3% empty, 1% settlement, 1% surface water. The agricultural economy of the basin depends on the irrigated cotton cultivation, corn, fig and olives. Total population of the basin is 2.5 million. The river during its course of about 584 km receives pollution load both from the point and non-point sources. It receives agricultural run-off from its vast catchments area directly or through its tributaries and wastewater drains. The declining water quality in the Büyük Menderes river has resulted in negative impacts to fish, wildlife, aquatic life. The coastal waters of the Aegean Sea have been plagued by algal blooms, leading to fish kills and unpleasant conditions for tourism (Durdu, 2007). In the present study, the Aydın measurement location was selected for the analysis (Fig. 1). The dataset of one measurement location, comprising three water quality parameters monitored over 9 years (1996–2004), was measured by the State Hydraulic Works (DSI), the General Directorate of Electrical Power Resources, Survey and Development Administration, General Directorate of Rural Services and Municipalities. The selected water quality parameters include water temperature, Boron (Br) and dissolved oxygen (DO). The data for these parameters available for the analysis are on a monthly basis for the period of nine years from 1996 to 2004. In this study, the first 72 months of water quality data for boron, dissolved oxygen and water temperature were used for model training. The remaining 36 months of water quality data were used for verification of the model prediction results. The statistical properties of water quality time series data are demonstrated in Table 1. The minimum, maximum, mean, standard deviation (SD), skewness and kurtosis can describe variability of a water quality parameter. As described in Table 1, water temperature and pH are the water quality variables that have low skewness coefficients. Dissolved oxygen has a large skewness coefficient and this indicates that the mean and median values of dissolved oxygen have large differences. Probably the mean of dissolved oxygen dataset is heavily influenced by the presence of a few extreme values.

2.2. ARIMA modeling approach

In an autoregressive integrated moving average model (ARIMA), the future value of a variable is assumed to be a linear function of several past observations and random errors. An ARIMA model can be explained as ARIMA$(p,d,q)(P,D,Q)_s$, where (p,d,q) is the non-seasonal part of the model and $(P,D,Q)_s$ is the seasonal part of the model (Box et al., 1991), which is mentioned below

$$\varphi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t=\theta_q(B)\Theta_Q(B^s)a_t \tag{1}$$

where p is the order of non-seasonal autoregression, d is the number of regular differencing, q is the order of non-seasonal MA, P is the order of seasonal autoregression, D is the number of seasonal differencing, Q is the order of seasonal MA, s is the length of season (periodicity), φ is the AR operator of order p, Φ is the seasonal AR parameter of order P, ∇^d is the differencing operator,

Table 1
Statistical properties of the water quality parameters.

Parameters	Max	Min	Mean	Std. Dev.	Skewness	Kurtosis
Water temperature (°C)	32.2	7.2	18.81	8.01	0.10	−1.401
Boron (mg l^{-1})	0.6	0	0.207	0.139	0.10	−0.074
DO (mg l^{-1})	9.3	6.1	8.2	2.40	−1.82	−0.169

∇_s^D is the seasonal differencing operator, z_t is the observed value at time point t, θ is the MA operator of order q, Θ is the seasonal MA parameter of order Q and a_t is the noise component of the stochastic model assumed to be NID(0, σ^2).

The ARIMA modeling approach involves the following three steps: model identification, parameter estimation, diagnostic checking. Identification of the general form of a model includes two stages: (1) if it is necessary, appropriate differencing of the series is performed to achieve stationary and normality; (2) the temporal correlation structure of the transformed data is identified by examining its autocorrelation (ACF) and partial autocorrelation (PACF) functions (Mishra and Desai, 2005). The ACF is a useful statistical tool that measures if earlier values in the series have some relation to later values. PACF is the amount of correlation between a variable and a lag of itself that is not explained by correlations at all low order lags. Considering the ACF and PACF graphs of water quality concentration series, different ARIMA models are identified to model selection. The model that gives the minimum Akaike Information Criterion (AIC) is selected as the best fit model. The mathematical formulation for the AIC is developed as

$$AIC = n(\ln((2\pi RSS)/n)+1)+2m \quad (2)$$

where m=(p+q+P+Q) is the number of terms estimated in the model and RSS denotes the sum of squared residuals.

After the functions of the ARIMA model have been specified, the parameters of these functions must be estimated. Once an appropriate model is chosen and its parameters are estimated, the Box–Jenkins methodology requires examining the residuals of the model to verify that the model is an adequate one for the series. Several tests are employed for diagnostic check to determine whether the residuals of the selected ARIMA models from the ACF and PACF graphs are independent, homoscedastic and normally distributed. If the homoscedasticity and normality assumptions are not provided, the observations are transformed by a Box–Cox transformation (Wei, 1990). For a good forecasting model, the residuals, left over after fitting the model, must satisfy the requirements of a white noise process (uncorrelated and normally distributed around a zero mean). In order to determine whether water quality time series are independent, the residual autocorrelation (RACF) function of the series is studied. There are several useful tests related to RACF for the independence of residuals. The first one is the correlograms drawn by plotting the residual ACF function against lag number. If the ARIMA model is correct, the estimated autocorrelations of the residuals are uncorrelated and distributed approximately normally about zero. The second one is Ljung–Box–Pierce statistics. In order to test the null hypothesis that a current set of autocorrelations is white noise, test statistics are calculated for different total numbers of successive lagged autocorrelations using the Ljung–Box–Pierce statistics ($Q(r)$ test) to test the adequacy of the model. $Q(r)$ values are compared to a critical test value (χ^2) distribution with respective degree of freedom at a 5% significant level. The third one is the cumulative periodogram, employed to diagnose the residuals for a white noise sequence. When modeling seasonal time series line the one in the present study, the periodic characteristics of water quality concentration time series might not be taken into account, therefore, the periodicities in the residuals should be investigated (El-Din and Smith, 2002).

2.3. Structure of neural network model

Among many neural network architectures, the three-layer-feedforward back propagation network is the most commonly used (Haykin, 1999). This network architecture consists of one hidden layer of neurons with nonlinear transfer functions and an output layer of linear neurons with linear transfer functions. A schematic diagram of backpropagation network is given in Fig. 2, where x_j (j=1, , N) represents the input variables; n_i (i=1, …, S) represents the outputs of neurons in the hidden layer and y_t (t=1, …, L) represents the outputs of the neural network. A neural network must be trained to determine the values of the weights that will produce the correct outputs. In a training step, a set of input data is used for training and presented to the network many times. The performance of the network is tested after the training step is stopped. The backpropagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient). This is the direction in which the performance function is decreasing most rapidly. It turns out that although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. Therefore, the basic gradient descent training algorithm is inefficient owing to its slow convergent speed and at times the poor accuracy in model predictions (Huang et al., 2004). From an optimization point of view, training a neural network can be considered as equivalent to minimizing a multivariable global error function of the network weights. There are several optimized training algorithms, as described by Haykin (1999), such as resilient backpropagation, Levenberg–Marquardt and conjugated gradient backpropagation. On of the optimized methods developed by Moller (1993) is the scaled conjugate gradient (SCG) algorithm. The SCG training algorithm was developed to avoid the time-consuming line search. In the conjugate gradient algorithm a search is performed along conjugate directions, which produces faster convergence than steepest descent directions. The standard backpropagation algorithm, traditionally employed in neural network learning, evaluates the gradient of the global error function with respect to the weights, $f(W^k)$, at each iteration k, and updates the weights according to

$$W^{k+1} = W^k - \alpha^k \nabla f(W^k) \quad (3)$$

The step size $\alpha^k > 0$ is a user-selected learning rate parameter, which affects the performance of the learning algorithm to a great extent. In all cases, the backpropagation algorithm may follow a

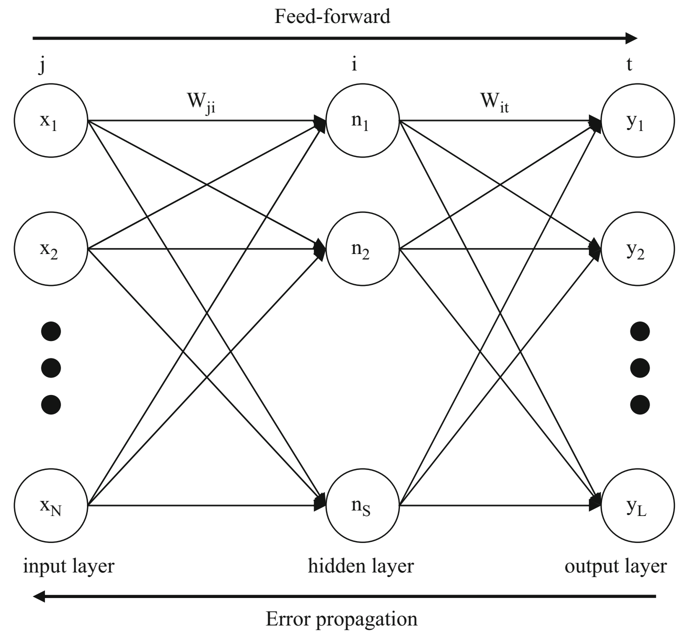

Fig. 2. A three-layer feed-forward back propagation neural network.

zigzag path to the minimum, typical for a steepest gradient descent method (Falas and Stafylopatis, 2005). A conjugate gradient algorithm avoids the zigzag approach to the minimum point by incorporating a special relationship between the direction and gradient vector at each iteration. If D^k represents the direction vector at iteration k of the algorithm, then the weight vector is updated according to the rule

$$W^{k+1} = W^k + \alpha^k D^k \tag{4}$$

Given values of W^k and D^k, a particular values of α^k that reduce the objective function as much as possible needs to be found. After a small number of iterations, the search along the line direction to find the optimum step size for the actual minimum should stop. Estimating the optimum step size with scaled conjugate gradient (SCG) training algorithm increases the learning speed and eliminates the dependence on critical user-selected parameters. The main idea behind the algorithm is the use of a factor ρ which is raised or lowered within each iteration during the execution of the algorithm, looking at the sign of the quantity δ^k, which reveals if the Hessian matrix is not positive definite. A brief algorithm of SCG in neural network is given as follows (Falas and Stafylopatis, 2005).

1. Initialization: At k=0, choose an initial weight vector W^o, and set the initial direction vector to the negative gradient vector $D^o = G^o = -\nabla f(W^o)$. Set the scalars $0 < \sigma < 10^{-4}$, $0 < \rho^o \leq 10^{-6}$, $\overline{\rho}^o$, set the boolean *success=true*.
2. If success=true, then calculate second order information: $\sigma^k = \sigma/|D^k|$, $S_k = (\nabla f(W^k + \sigma^k D^k) - \nabla f(W^k))/\sigma^k$, $\delta^k = transpose$ $(D^k) + S^k$
3. Scale δ^k: $\delta^k = \delta^k + (\rho^k - \overline{\rho}^k)|D^k|^2$, look at the sign of δ^k for each iteration adjusting ρ^k. If $\delta^k \leq 0$ then ρ^k and S_k is estimated again.
4. If $\delta^k \leq 0$ then make the Hessian positive definite
$$\overline{\rho}^k = 2(\rho^k - \delta^k/\left|D^k\right|^2), \quad \delta^k = -\delta^k + \rho^k\left|D^k\right|^2, \quad \rho^k = \overline{\rho}^k$$
5. Calculate the step size: $\xi^k = transpose(D^k)G^k$, $\alpha^k = \xi^k/\delta^k$, the values of ρ^k directly scale the step size in the way, that the bigger ρ^k, the smaller the step size.
6. Calculate the comparison parameter c^k: $c^k = 2\delta^k \lfloor f(W^k) - f(W^k + \alpha^k D^k) \rfloor/(\xi^k)^2$
7. Weight and direction update: If $c^k \geq 0$, then a successful update can be made: $W^{k+1} = W^k + \alpha^k D^k$, $G^{k+1} = -\nabla f(W^{k+1})$, $\overline{\rho}^k = 0$, $success = true$. If k mod N=0 then restart algorithm with $D^{k+1} = G^{k+1}$ *else* $\beta^k = (|G^{k+1}|^2 - G^{k+1\mathrm{T}}G^k)/\xi^k$, $D^{k+1} = G^{k+1} + \beta^k D^k$. If $c^k \geq 0.75$ then reduce the scale parameter to $\rho^k = \frac{1}{4}\rho^k$ else $\overline{\rho}^k = \rho^k$, $success = false$.
8. If $c^k \geq 0.25$ then increase the scale parameter to $\rho^k = \rho^k + \delta^k(1 - c^k)/|D^k|^2$.
9. Repetition: If the steepest descent direction $G^k \neq 0$, *set k* $= k+1$ and go back to step 2 else terminate and return W^{k+1} as the desired minimum.

2.4. Hybrid model

The behavior of water quality data may not easily be captured by stand-alone models because water quality time series data could include a variety of characteristics such as seasonality, heteroskedasticity or a non-Gaussian error. The approximation of ARIMA models to complex nonlinear problems may not be adequate. On the other hand, using ANNs to model linear problems have yielded unsatisfactory results. Zhang (2003) indicated that it is not wise to apply ANNs blindly to any type of data. Therefore, a hybrid model having both linear and nonlinear modeling abilities could be a good alternative for predicting water quality data. By combining different models, different aspects of the underlying patterns may be captured. Following the hybrid model structure of Zhang (2003), a water quality time series could be composed of a linear autocorrelation structure and a nonlinear component. That is,

$$y_t = L_t + S_t \tag{5}$$

where L_t represents the linear component and S_t represents the nonlinear component. Both of these two parameters have to be estimated from the time series data. First ARIMA model is used to capture the linear component, then the residuals from the linear model will contain only the nonlinear relationship. The residuals e_t at time t from the linear model is defined by

$$e_t = y_t - \hat{L}_t \tag{6}$$

where $\hat{L}_t$ is the predicted value of the ARIMA model at time t. The diagnostic check of the residuals is important to determine the adequacy of the ARIMA models. An ARIMA model is not sufficient is not adequate if there are still linear correlation structures left in the residuals. However, diagnostic check of the residuals is not able to detect any nonlinear patterns in the times series data. For this reason, even if the residuals pass the diagnostic check and the model is an adequate one, the model may still not be sufficient in that nonlinear relationships have not been appropriately modeled. Therefore, the residuals can be modeled by using ANNs to discover nonlinear relationships. With N input nodes, the ANN model for the residuals will be

$$e_t = f(e_{t-1}, e_{t-2}, \ldots., e_{t-N}) + \varepsilon_t \tag{7}$$

Where f is a nonlinear function determined by the neural network and ε_t is the random error. Finally the combined prediction will be

$$\hat{y}_t = \hat{L}_t + \hat{S}_t \tag{8}$$

where $\hat{S}_t$ represents the prediction from Eq. (7).

2.5. Model verification and comparison methods

Three different forecast consistency measures are used in order to compare the performances of obtained ARIMA and artificial neural network (ANN): root mean square error (RMSE), the mean absolute percentage error (MAPE) and the Nash–Sutcliffe coefficient of efficiency (NSC).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{9}$$

where P_i and O_i are the predicted and observed weekly water temperatures, respectively, and n is the number of data. The second criterion is the mean absolute percentage error (MAPE), which is defined as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_i - O_i}{P_i}\right|100 \tag{10}$$

The third criterion is the Nash–Sutcliffe coefficient of efficiency (NSC) (Nash and Sutcliffe, 1970), which is <1 (equal to1 when P=0), and is defined as (Benyahya et al., 2007)

$$NSC = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(\overline{O} - O_i)^2} \tag{11}$$

where $\overline{O}$ is the average of observed monthly water quality parameters. Scatter plots and time series plots are used for visual comparison of the observed and predicted values. NSC values of zero, one and negative indicate that the observed mean as good a

Table 2
Summary of the statistical parameters of the best fitted multiplicative ARIMA models fitted to water quality parameters.

Parameters	Model	ϕ_1	θ_1	Φ_1	Θ_1	Q	χ^2(95%)	AIC
Water temperature (°C)	$(1.1.1)(0.0.1)_{12}$	0.639	−0.062	–	−0.584	27.01	32.67	478
Boron (mg l^{-1})	$(1.0.1)(0.1.1)_{12}$	0.677	0.458	–	0.843	29.64	32.67	−126.3
DO (mg l^{-1})	$(1.1.0)(1.0.0)_{12}$	−0.017	–	0.428	–	31.28	32.67	192.1

predictor as the model, a perfect fit and a better predictor than the model, respectively (Palani et al., 2008). Decision makers can decide whether the predictability of the ANN model is accurate enough to make important decisions regarding to data usage depending on the mismatch between the forecasted water quality parameters and that measured.

3. Results and discussion

3.1. ARIMA modeling

In the present study, several trails were made to choose the optimal ARIMA model parameters. The model parameters that satisfy the statistical residual diagnostic checking were chosen in the ARIMA forecasting model. The ARIMA models were used to predict monthly water quality time series over the period between 1996 and 2004. The water quality data for the period between 1996 and 2001 were used for model calibration and to obtain the best model fit for each water quality parameter. The data for the period between 2002 and 2004 were used for model verification and comparisons for prediction purposes. In the ARIMA modeling process, the input and output water quality data sets for each parameter were normalized to the range of $[0, 1]$.

To fit ARIMA model to the available water quality time series data, three-stage procedure of model identification, estimation of model parameters and diagnostic checking of the estimated parameters was employed. In the identification stage, to determine the possible persistence structure in the time series data, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) were used. Using Akaike Information Criteria (AIC), the best fitted model has been identified out of the various competing models. As demonstrated in Table 2, the seasonal components (P, D, Q) of best fit ARIMA models are $(0, 0, 1)$ for water temperature, $(0, 1, 1)$ for boron and $(1, 0, 0)$ for dissolved oxygen. The nonseasonal components (p, d, q) are $(1, 1, 1)$, $(1, 0, 1)$ and $(1, 1, 0)$ for water temperature, boron and dissolved oxygen, respectively. The AICs for the best fit ARIMA models are 478, −126.3 and 192.1 for water temperature, boron and dissolved oxygen, respectively (Table 2). In the estimation of model parameters stage, the best fit ARIMA model statistical parameters were estimated. The computational method outlined by Box and Jenkins (1976) was employed to estimate model parameters. In the diagnostic checking of the estimated parameters stage, diagnostic checks were done to insure that the best fit model was selected by checking that assumptions of ARIMA modeling such as independence, homoscedastic (constant variance) and normality of the residual a_t were satisfied. In order to check the independence of residuals, the residual autocorrelation function (RACF), Ljung–Box–Pierce statistics and cumulative periodograms were used. The values of residual autocorrelation functions (RACF) were well settled within confidence limits except very few individual correlations appear large compared to the confidence limits, which were acceptable

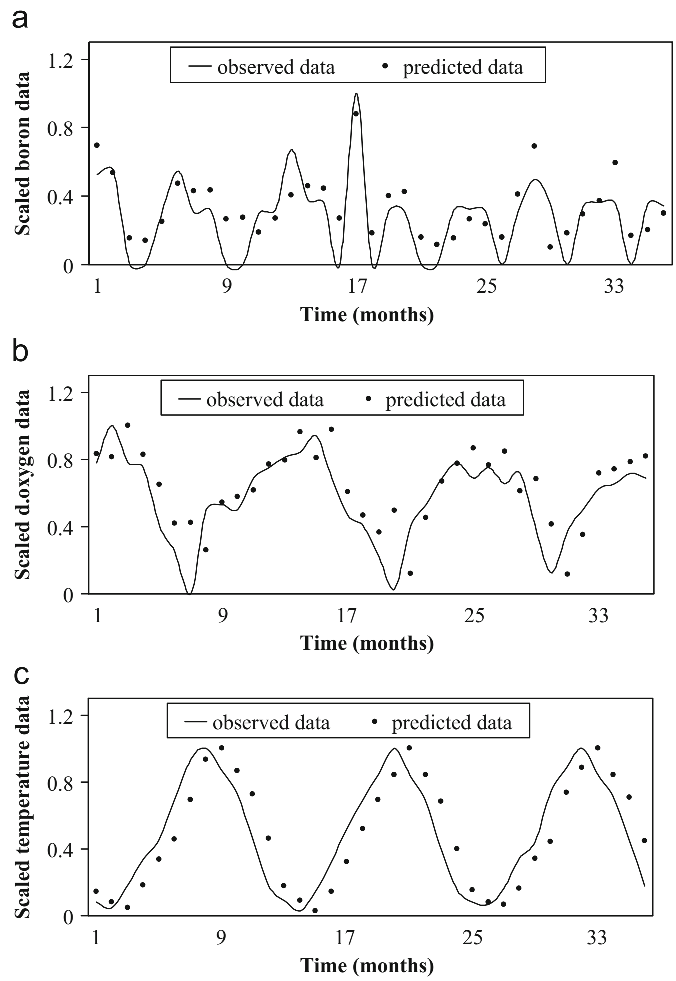

Fig. 3. ARIMA model verification for each water quality constituent.

among 30 lags. The results exhibited no significant correlation between the residuals of the each water quality parameter. The values of Ljung–Box–Pierce *Q(r)* statistic is shown in Table 2 and has a value of 27.01, 29.64 and 31.28 for water temperature, boron and dissolved oxygen, respectively. The values of *Q(r)* were compared to a critical test value (χ^2) distribution with respective degree of freedom at a 5% significant level. It was obvious that the computed values were less than the actual (χ^2) values, which indicated that the residuals from the best models were white noise (Table 2). Cumulative periodograms confirmed that no significant periodicity was available in the residual series at 95% confidence level and indicated that the points were clustering closely about the theoretical line and there was no evidence of periodic characteristics buried in the residual series. In order to check the normality of residuals, the histograms and

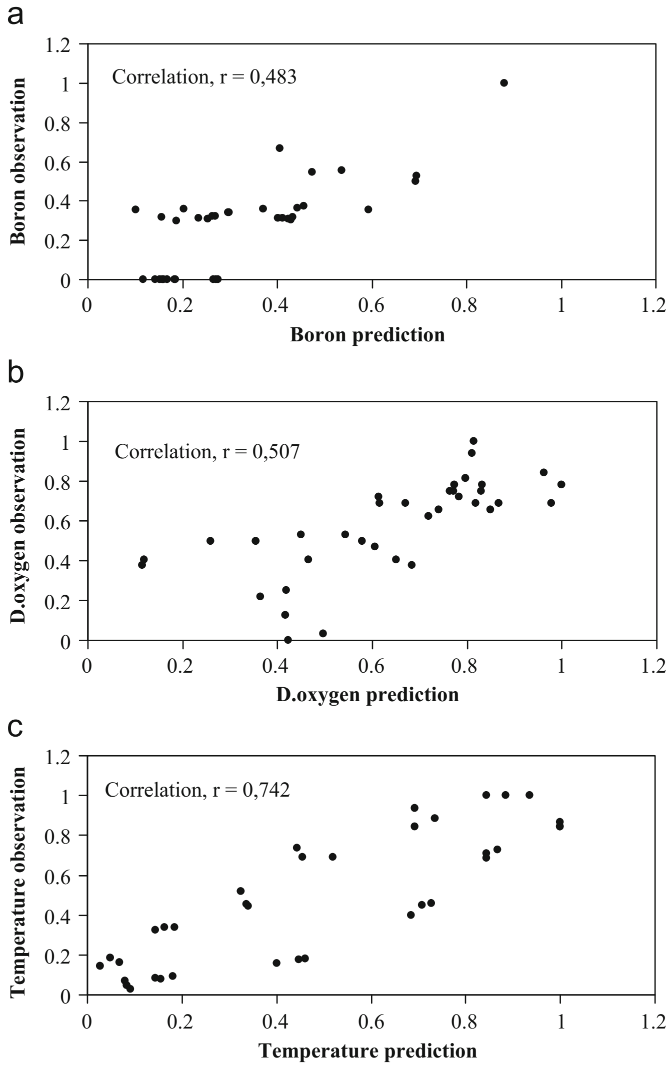

Fig. 4. Observed versus ARIMA predicted data for each water quality constituent.

normal probability plot of residuals were investigated and they clearly supported the assumption of normality. In order to investigate homoscedasticity of the residuals, a plot of residuals versus fitted values were examined and the plots showed a random scatter around zero. In other words, the residuals were evenly distributed around mean, which explains the models were adequate.

The trained ARIMA model was then tested using water quality data set for the period of 36 months. As shown in Fig. 3, although the ARIMA models generally vary with the range of most of the water quality data, the model predictions are not quite satisfied. The correlation coefficient values between models predicted values and observed data for boron, dissolved oxygen and water temperature are 0.483, 0.507 and 0.742, respectively, which are not satisfactory in common model applications (Fig. 4). Although the ARIMA models were able to show the cycles of the high and low water quality values in a year, they were not able to provide good predictions of the water quality value magnitudes, which changed from month to month. This limitation was due mainly to the limitations of the linear modeling algorithm in the ARIMA model, the performance of which was generally not quite satisfactory in recognizing and reproducing the nonlinear time series of water quality data.

3.2. Neural network modeling

A three-layer feedforward neural network model was developed for the prediction of water quality constituents (boron, dissolved oxygen and water temperature) using an optimized back-propagation training algorithm. In the present study, the scaled conjugated gradient algorithm was selected as the optimized training method. In the following part, artificial neural network model performances were validated for flow prediction under monthly time-step condition. The data for the period between 1996 and 2004 were available for the modeling purposes. Boron, dissolved oxygen and water temperature time series data were divided into two independent data sets. The first data set was used for model training, and the second data set was used for model verification purposes. In the ANN modeling process, the input and output water quality data sets for each parameter were normalized to the range of $[0, 1]$. Fig. 5 demonstrates the ANN model training performance for each water quality parameter. Fig. 5a indicates that the ANN model for boron time series data reaches the preset training goal of 0.001 after 100 epochs. For dissolved oxygen time series data, the ANN model run through 40 epochs to reach the required training goal

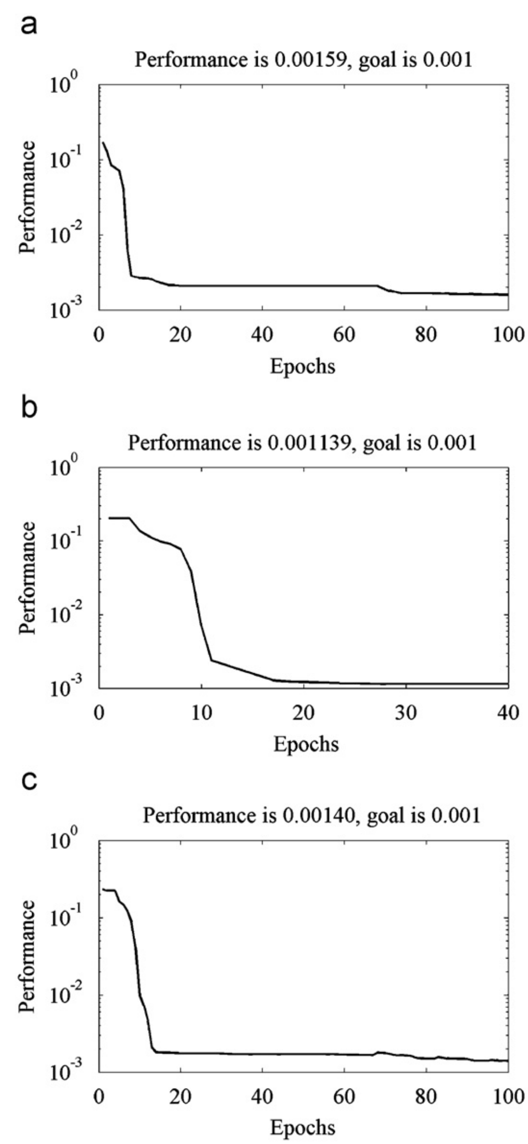

Fig. 5. Training performance of ANN model for each water quality constituent.

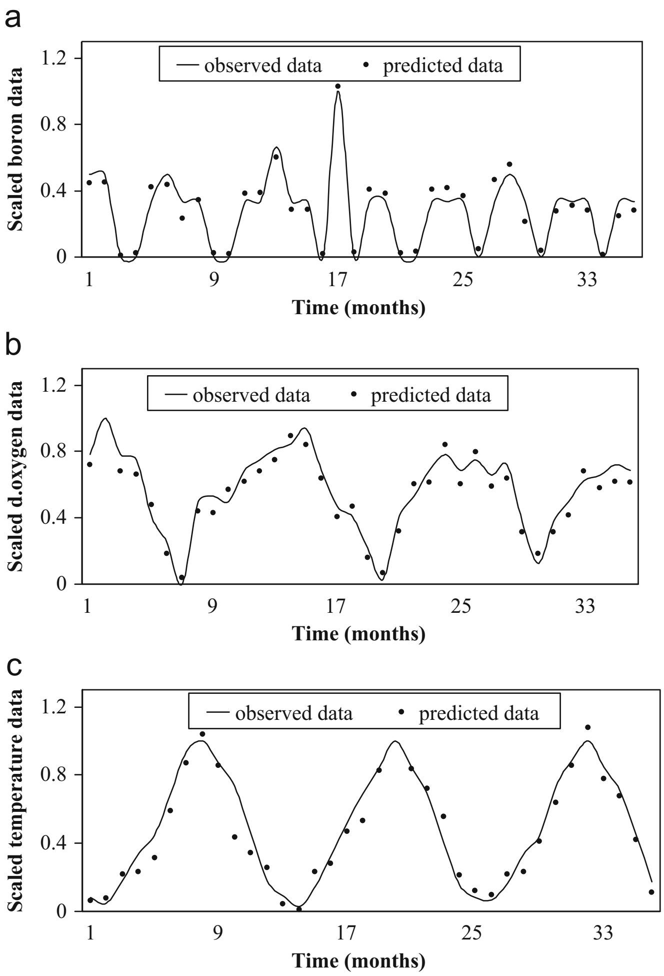

Fig. 6. ANN model verification for each water quality constituent.

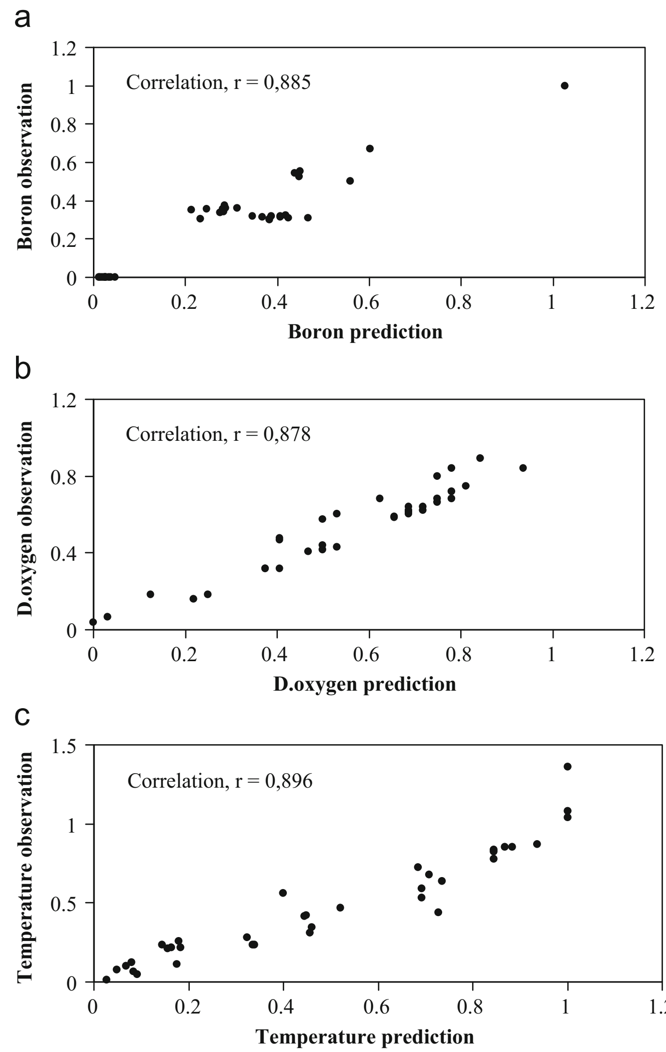

Fig. 7. Observed versus ANN predicted data for each water quality constituent.

of 0.001 (Fig. 5b). By comparing model predictions with observations, model parameters (weights) were calibrated. It takes 100 epochs for the neural network model to reach the required training goal of 0.001 for water temperature data (Fig. 5c). These results indicate that the ANN model can be trained to adjust model weights so that the model predicted water quality parameters match well with observed data.

In the model verification phase, the trained network was used to predict the monthly water quality parameters. Fig. 6 compares the model predictions for water quality parameters with the observations. The verifications stage indicate that the model prediction results reasonably match the observed water quality parameters. The correlation coefficient between the ANN model predicted values and observed data for boron, dissolved oxygen and water temperature are 0.885, 0.878 and 0.896, respectively, which are satisfactory in common model applications (Fig. 7). These results indicate that the neural network model is able to recognize the pattern of the water quality parameters to provide good predictions of the monthly variations of water quality data of the Büyük Menderes river.

3.3. Hybrid modeling

The proposed algorithm of the hybrid system consisted of two steps. In the first step, to analyze the linear part of the problem, an ARIMA model was employed. In the second step, the residuals from the ARIMA model were modeled by using a neural network model. Since the ARIMA model cannot detect the nonlinear structure of the water quality time series data, the residuals of linear model will contain information about the nonlinearity. The outputs from the neural network can be used as predictions of the error terms of the ARIMA model. The hybrid model utilizes the unique feature and strength of ARIMA model as well as ANN model in determining different patterns. Therefore, it may be favorable to model linear and nonlinear patterns separately by using different models and then combine the predictions to improve the overall modeling and predicting performance. In the hybrid modeling algorithm, the input and output water quality data sets for each parameter were normalized to the range of [0,1]. In the modeling process, the hybrid model was trained to adjust the model so that the model predicted water quality parameters match well with observed data. Fig. 8 compares the model predictions for water quality parameters with the observations. The verifications results indicate that the model prediction results reasonably match the observed water quality parameters. The correlation coefficient between the hybrid model predicted values and observed data for boron, dissolved oxygen and water temperature are 0.902, 0.893 and 0.909, respectively, which are satisfactory in common model applications (Fig. 9).

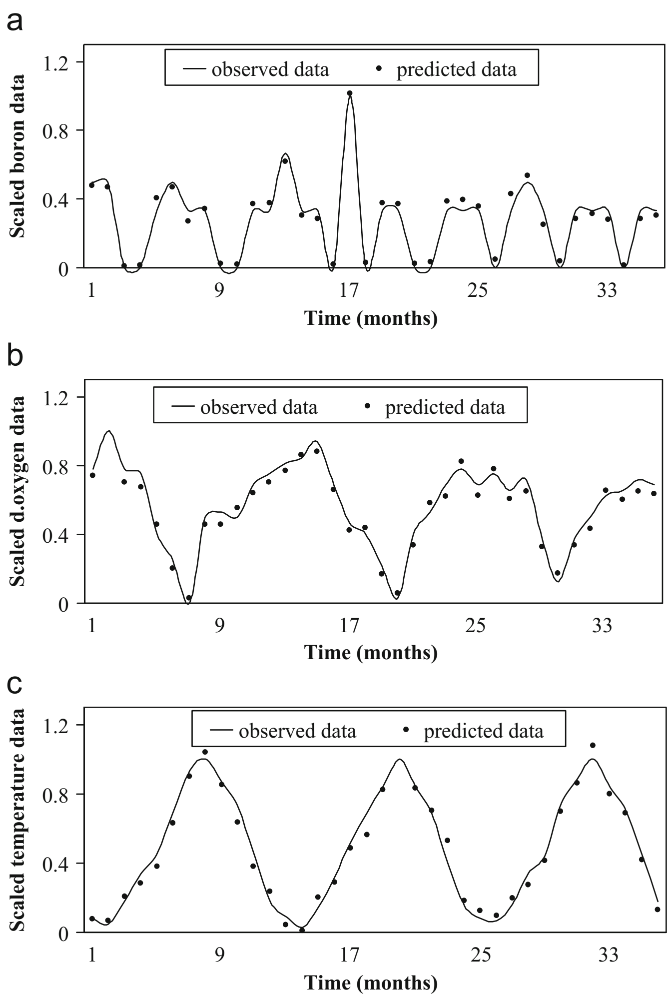

Fig. 8. Hybrid model verification for each water quality constituent.

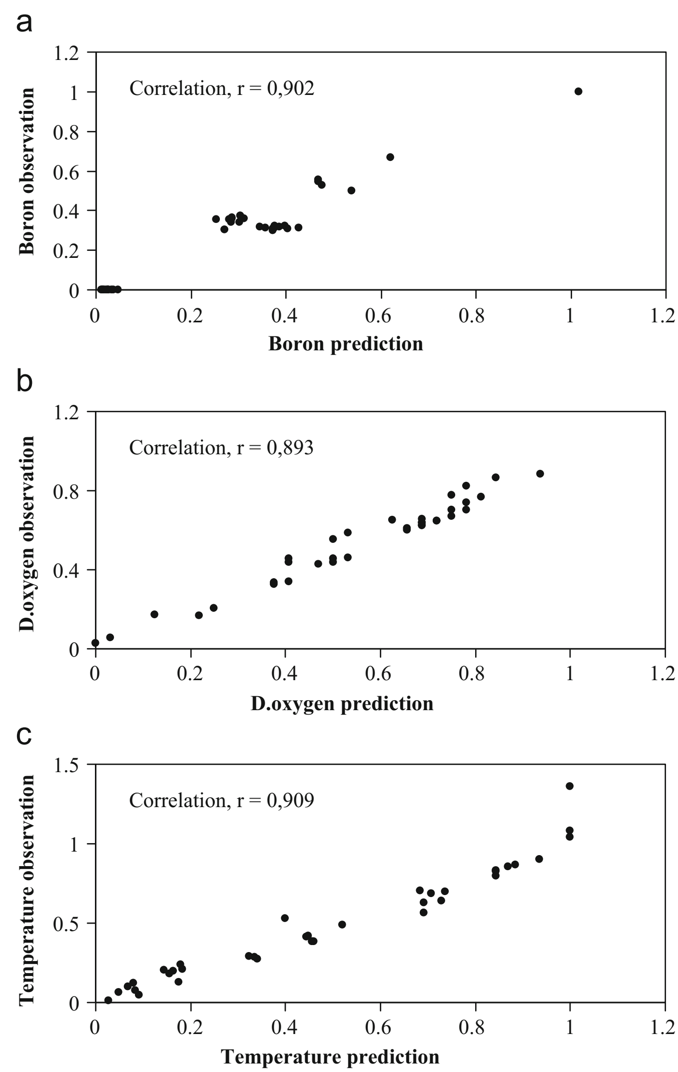

Fig. 9. Observed versus hybrid model predicted data for each water quality constituent.

3.4. Comparison of model performances

Employing accuracy measures (RMSE, MAPE and NSC), the predicted data and observed data for the period of 36-months from the hybrid, ANN and ARIMA models were compared to determine the best performed model. The predicted water quality parameters using the ARIMA models were not found to be in reasonable agreement with the observed data. However, the hybrid and ANN approaches provided reasonable precision for all water quality parameters. Table 3 gives the error estimates of the three different approaches used in the study for predicting water quality parameters. The RMSEs between observed and predicted data were calculated in ARIMA models as 0.102 °C, 0.165 and 0.113 mg l^{-1} for water temperature, boron and dissolved oxygen, respectively. In the case of ANN modeling approach, the RMSEs between observed and predicted data were computed as 0.048 °C, 0.074 and 0.061 mg l^{-1} for water temperature, boron and dissolved oxygen, respectively. Applying the hybrid method, there were a decrease of 18.75%, 14.86% and 16.39% in the RMSE values of ANN for water temperature, boron and dissolved oxygen, respectively. Furthermore, the MAPEs between observed and predicted data for water temperature, boron and dissolved oxygen were appeared to be slightly lower for the ANN modeling approach. Prediction error statistics for the ANN approach produced MAPEs of 21%, 36% and 29% for water temperature, boron and dissolved oxygen, respectively. In case of the MAPE values, the improvement of the hybrid model over the ANN model were 13.41%, 8.69% and 8.87% for water temperature, boron and dissolved oxygen, respectively. The ARIMA model produced lower NSC values in comparison to the ANN model. The NSC values for the ARIMA modeling approach were 0.716, 0.534 and 0.682 for water temperature, boron and dissolved oxygen, respectively. However, the produced NSC values for the ANN model were 0.931, 0.893 and 0.912 for water temperature, boron and dissolved oxygen, respectively. With the application of the hybrid model, the improvement over the ANN's NSC values were 15.03%, 13.43% and 12.31% for water temperature, boron and dissolved oxygen, respectively. These results indicated that the hybrid model performed well for adequate predicting of water temperature, boron and dissolved oxygen.

It is obvious that the hybrid model is able to simulate the water temperature with an accuracy of a degree in comparison to the ANN and ARIMA models. In the Büyük Menderes river at Aydın location, the hybrid model approach was successful in simulating the magnitude and patterns in observed dissolved oxygen concentrations that resulted from seasonal temperature variations, periodic blooms of phytoplankton and point source

Table 3
Statistical comparison of observed and predicted data from the hybrid, ARIMA and artificial neural network (ANN) modeling approaches.

Parameters	RMSE			MAPE			NSC		
	ARIMA	ANN	Hybrid	ARIMA	ANN	Hybrid	ARIMA	ANN	Hybrid
Water temperature (°C)	0.102	0.048	0.039	42.076	21.114	18.282	0.716	0.931	0.945
Boron (mg l^{-1})	0.165	0.074	0.063	57.512	36.813	33.611	0.534	0.893	0.904
DO (mg l^{-1})	0.113	0.061	0.051	47.103	29.612	26.985	0.682	0.912	0.924

discharges of oxygen consuming substances like ammonia. The developed hybrid algorithm for boron constitute was able to recognize the pattern of the input data to provide good predictions of the monthly variations of boron content. Therefore, it can be concluded that the hybrid modeling approach can give more reliable predictions of water temperature, boron and dissolved oxygen time series of a river than the ANN and ARIMA modeling approach.

4. Conclusions

A new approach of modeling water quality time series, capable of exploiting the advantages of both the conventional methods and the ANNs, was proposed. An empirical comparative evaluation of the performance of hybrid model to the ANN and ARIMA modeling approach was presented for river water quality predictions. The proposed modeling framework gradually receives the data filtered using the ARIMA models and then the residuals from the ARIMA approach were analyzed by ANNs to capture the nonlinearity in the time series involved. Investigations were conducted to examine the hybrid model performance for predicting river water quality in monthly time steps. The results from the ARIMA models poorly represented the pattern of water quality data for boron and dissolved oxygen, but the model produced acceptable results for water temperature. The results from the ANN model were capable of providing accurate predictions of water quality parameters at the proposed time step. In the proposed hybrid model, an ARIMA model was used to analyze the linear part of the problem and then the residuals from the ARIMA model were modeled by using a neural network model. The results from the hybrid model indicated that the modeling approach gave more reliable predictions of water temperature, boron and dissolved oxygen time series data.

The predictions from hybrid model were compared with those obtained from the ANN and ARIMA traditional time series approaches. Owing to its ability in recognizing time series patterns and nonlinear characteristics, the accuracy measures RMSE, MAPE and NSC demonstrated that the hybrid model provided much better accuracy over the ANNs and ARIMA methods for water quality predictions. Therefore, the proposed hybrid algorithm can be used for the Büyük Menderes river and other hydrometerologically similar rivers for predicting water quality data of monthly time step to detect water quality severity with respect to water temperature, boron and dissolved oxygen in future. The hybrid model developed for the Büyük Menderes river can be employed for the development of a water quality emergency management plan so as to ensure sustainable water resources management in the basin.

References

Benyahya, L., Hilaire, A.S., Quarda, B.M.J.T., Bobee, B., Nedushan, B.A., 2007. Modeling of water temperatures based on stochastic approaches: case study of the Deschutes River. Journal of Environmental Engineering Science 6, 437–448.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis Forecasting And Control. Holden-Day, San Francisco, USA.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 1991. Time Series Analysis, Forecasting and Control. Prentice Hall, Englewood Cliffs, NJ. USA.

Chau, K.W., 2006. A review on integration of artificial intelligence into water quality modeling. Marine Pollution Bulletin 52, 726–733.

Durdu, O.F., 2007. GIS based PCraster-polflow approach to modeling of coupled surface water and groundwater hydrology in the Büyük Menderes catchment. In: Proceedings of International Congress on River Basin Management, Antalya, Turkey, 258-270.

El-Din, A.G., Smith, D.W., 2002. A combined transfer-function noise model to predict dynamic behavior of a full-scale primary sedimentation tank. Water Research 36, 3747–3764.

Falas, T., Stafylopatis, A., 2005. Implementing temporal difference learning with the scaled conjugated gradient algorithm. Neural Processing Letters 22, 361–375.

Flood, I., Kartam, N., 1994. Neural networks in civil engineering I: principles and understanding. Journal of Computing in Civil Engineering 8, 131–148.

Hassoun, M.H., 1995. Fundamentals of Artificial Neural Networks. MIT Press, Cambridge.

Hatzikos, E., Anastasakis, L., Bassiliades, N., Vlahavas, I., 2005. Simultaneous prediction of multiple chemical parameters of river water quality with tide. In: Proceedings of the Second International Scientific Conference on Computer Science, IEEE Computer Society, Bulgarian Section.

Haykin, S., 1999. Neural Networks, a Comprehensive Foundation. Prentice Hall, Upper Saddle River, NJ, USA.

Huang, W., Bing, X., Chan-Hilton, A., 2004. Forecasting flows in Apalachicola River using neural neworks. Hydrological Process 18, 2545–2564.

Jain, A., Kumar, A., 2006. An evaluation of artificial neural network technique for the determination of infiltration model parameters. Applied Soft Computing 6, 272–282.

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environmental Modelling & Software 15, 101–124.

Mishra, A.K., Desai, V.R., 2005. Drought forecasting using stochastic models. Stochastic Environmental Research Risk Assessment 19, 326–339.

Moller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 6, 525–533.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. Journal of Hydrology 10, 282–290.

Palani, S., Liong, S.Y., Tkalich, P., 2008. An ANN application for water quality forecasting. Marine Pollution Bulletin 56, 1586–1597.

Pisoni, E., Farina, M., Carnevale, C., Piroddi, L., 2009. Forecasting peak air pollution levels using NARX models. Engineering Applications of Artificial Intelligence 22, 593–602.

Rojas, R., 1996. Neural Networks: A Systematic Introduction. Springer-Verlag, Berlin.

Shahwan, T., Odening, M., 2007. Computational Intelligence in Economics and Finance. Springer, Berlin Heidelberg, New York, pp. 63–74.

Su, C.T., Tong, L.I., Leou, C.M., 1997. Combination of time series and neural network for reliability forecasting modeling. Journal of Chinese Industrial Engineering 14, 419–429.

Wei, W.W.S., 1990. Time series analysis. Addison-Wesley Publishing Company, Inc., New York.

Zhang, G., Hu, M.Y., 1998. Neural network forecasting of the British pound US dollar exchange rate, Omega. International Journal of Management Science 26, 495–506.

Zhang, P.G., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50, 159–175.