Monthly Water Consumption Prediction Using Season Algorithm and Wavelet Transform–Based Models

Abdusselam Altunkaynak, A.M.ASCE[1]; and Tewodros Assefa Nigussie[2]

Abstract: Accurate prediction of water demand is essential for optimum management of water resources and sustainable growth and development. Recently, models based on artificial neural networks (ANNs), in combination with data preprocessing techniques, have been used for water demand prediction due to their ability to handle large amounts of complex nonlinear data. Discrete wavelet transform (DWT) is one of the most widely employed data preprocessing techniques, and is used in combination with ANNs to improve prediction accuracy and extend prediction lead time. However, DWT is known to have serious drawbacks, and the accuracy and prediction lead times of the models have not been satisfactory. In this study, multiplicative season algorithm (MSA) is applied for the first time as an alternative data preprocessing technique in the area of hydrology and its performance is compared with DWT. The outputs of MSA and DWT are used as inputs to a multilayer perceptron (MLP) in order to develop combined models called discrete wavelet transform–multilayer perceptron (DWT-MLP) and multiplicative season algorithm–multilayer perceptron (MSA-MLP), which are compared with the stand-alone MLP model. The results demonstrate that MSA is a better preprocessing technique than DWT and, thus, that MSA captures periodicity and converts nonstationary time series into stationary time series better than DWT. **DOI: 10.1061/(ASCE)WR.1943-5452.0000761.** *© 2017 American Society of Civil Engineers.*

Author keywords: Season algorithm; Wavelet transform; Water consumption.

Introduction

Water demand has increased in many parts of the world, particularly in urban areas, as a result of rapid population growth and industrial development (Tiwari and Adamowski 2015). Water demand is a very important variable that is affected by several factors, such as climatic conditions, population and city size, commercial and social conditions of people, cost of supply, and characteristics of the water-distribution system. Thus it is considered to be a dynamic process whose prediction is essential for optimum operation of water resources and sustainable growth and development (Sastri and Valdes 1989). Prediction of the magnitude of water demand into the future would allow land planners and managers to tailor various activities that can improve water supply and customer satisfaction because it helps managers understand the temporal patterns of future water use. According to Altunkaynak et al. (2005), water demand should be predicted because it can affect many properties of engineering applications, such as timing and sizing. Accurate prediction of water demand is required as an input in order to make sound decisions concerning supply-system development and expansion, the staging of such projects, and the establishment of water rates (Polebitski and Palmer 2010). It is also used in activities related to the optimization of system operations (Donkor et al. 2014) and is key to municipal water-supply management (Tian et al. 2016).

[1]Associate Professor, Faculty of Civil Engineering, Hydraulics and Water Resource Engineering Division, Istanbul Technical Univ., Maslak, Istanbul 34469, Turkey. E-mail: altunkay@itu.edu.tr

[2]Ph.D. Student, Faculty of Civil Engineering, Hydraulics and Water Resource Engineering Division, Istanbul Technical Univ., Maslak, Istanbul 34469, Turkey (corresponding author). E-mail: hiyawtewodros@gmail.com

Note. This manuscript was submitted on February 8, 2016; approved on November 15, 2016; published online on February 10, 2017. Discussion period open until July 10, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

A number of water-demand prediction models have been developed and applied during the last several years (Aly and Wanakule 2004) for long-term, medium-term, and short-term urban water-demand prediction. The models also attempted to represent the growth and periodic behavior of monthly, daily, or hourly water consumption (Arandia et al. 2016). According to Mohammed and Ibrahim (2012), long-term prediction of urban water demand refers to up to 25 years of prediction lead time and it is a critical and essential factor for water-supply planning, which includes the determining of type, size, location, and timing of the required improvements and developments of the water-supply systems. Medium-term prediction refers to up to two years of prediction lead time, and it is needed for planning large-scale maintenance activities. Short-term prediction of municipal water demand refers to up to two days of prediction lead time, and it is required for water utilities to proactively optimize pumping and treatment operations to minimize energy costs, water-supply costs, and treatment costs while maintaining a reliable and high-quality product for their customers. In general, accurate short-term, medium-term, and long-term demand forecasting provides water-distribution companies with information for capacity planning, maintenance activities, system improvements, pumping-operations optimization, and the development of purchasing strategies (Ghiassi et al. 2008). Short-term water-demand forecasting model results can be used for optimal overall quantity control or for optimal pump scheduling. Even though the time frames are different, long-term predictions of future water demand are necessary in order to develop new water sources and/or expand treatment plants, and short-term water predictions are needed to operate treatment plants and wells so that the required demand is met with the required quality Donkor et al. (2014). Mechanisms and systems that can be developed to predict both short-term and long-term water demands need to be performed so that water-supply entities can supply clean and safe water at the rate required by consumers (Msiza et al. 2007).

The simplest way to forecast future water demand has been to calculate the product of present per capita water consumption and

J. Water Resour. Plann. Manage., 2017, 143(6): 04017011

the predicted size of population. This technique is not a reliable means because it does not take into consideration the changes in per capita water consumption as a result of weather variability (seasonal variation), economic growth, change in lifestyle, and technological development. Classical regression models have also been widely used for predicting water demand, although they have drawbacks arising from the strict restrictive assumptions such as linearity, constant variance, and normal distribution (Altunkaynak 2010) that need to be fulfilled before developing regression models for prediction purposes. In recent years, other modeling methods that do not have these restrictive assumptions and others, such as stationarity and ergodicity, which are primary requirements for stochastic modeling, have been developed and used to predict water demand. Maier and Dandy (2000) observed that the use of models that are based on artificial neural networks (ANNs) is on the increase in the field of water management due to their ability to handle large amounts of nonlinear data and their ability to handle complex associations between inputs and outputs (Andrade et al. 2016). The information-processing characteristics of ANNs such as nonlinearity, parallelism, noise tolerance, learning, and generalization capabilities have made ANNs an attractive method (Seckin et al. 2009). In addition, ANN-based models have been found to perform better when compared with classical models such as autoregressive integrated moving average (ARIMA) (Adamowski 2008; Adebiyi et al. 2014) and autoregressive integrated moving average with explanatory variable (ARIMAX) (Akbari and Bozorg 2012; Young and Liu 2015; Young et al. 2015). Above all, according to Zhang (2003), the major advantage of neural networks is their flexible nonlinear modeling capability, as opposed to ARIMA models, where no nonlinear patterns can be captured. Thus ANNs were selected as a modeling tool in this study.

According to Maier and Dandy (2000), data preprocessing techniques have been used in combination with ANN models and have resulted in significant improvements in the models' performance. In this regard, singular spectrum analysis (SSA), moving average (MA), and principal component analysis (PCA) are recognized as efficient preprocessing algorithms for time series prediction. Other preprocessing techniques that were found to improve the water-demand prediction accuracy of ANN models are fuzzy regression (Azadeh et al. 2012) and Fourier analysis (Odan and Reis 2012). According to Dabhi and Chaudhary (2014), however, wavelet transform—discrete wavelet transform (DWT), in particular—has been used as an effective preprocessing technique in a number of studies involving nonstationary time series that used ANN as a modeling tool. Mohammed and Ibrahim (2012), for example, used a combined wavelet–neural network model to forecast municipal water demand of Tampa, Florida. Their results from both daily and monthly forecasting models for the municipal water-consumption time series demonstrated that the combined wavelet–ANN approach provided accurate daily and monthly prediction, with R^2 values of greater than 0.967. In addition, studies by Huang et al. (2014) and Tiwari and Adamowski (2015) demonstrated that the performance of ANNs in predicting water demand can be improved when they are combined with wavelet transform.

Despite this, however, DWT is shift sensitive because input-signal shifts generate unpredictable changes in DWT coefficients. It also suffers from poor directionality because its coefficients reveal only three spatial orientations. Furthermore, analysis with DWT lacks phase information that accurately describes nonstationary signal behavior, and decomposition of original time series data into bands by DWT involves complex procedures (Fernandes 2001). These limitations call for the development of a simple but accurate preprocessing technique that does not suffer from these drawbacks. In relation to this, Altunkaynak and Nigussie (2015) and Altunkaynak and Nigussie (2016) applied an additive season algorithm (SA) in hydrological studies in their studies involving the development of ANN-based daily-precipitation and streamflow prediction tools, respectively, and found that SA outperformed DWT. However, it was recommended that the performance of SA as an effective preprocessing method should be further investigated by taking various hydrological variables into consideration. Therefore this study was initiated to investigate the performance of multiplicative season algorithm (MSA) as an effective preprocessing method in comparison with DWT by taking monthly water-consumption data into consideration. As a result, three models—the stand-alone multilayer perceptron (MLP), the combined discrete wavelet transform–multilayer perceptron (DWT-MLP), and the combined multiplicative season algorithm–multilayer perceptron (MSA-MLP)—were developed for predicting the monthly water consumption of Istanbul for prediction lead times of 1, 2, 3, 6, 9, and 12 months. Although it is possible to develop sophisticated models by taking into consideration hydrological and meteorological variables such as precipitation, temperature, humidity, and evaporation, it is economically preferable that a model that simulates the water demand on the basis of past consumption records be available to decision makers, whether administrators, local authorities, or technical operators (Kisi 2005). As a result, in this study the ANN-based models use previous water consumption data (raw or preprocessed) as inputs to predict the water-consumption values with the indicated prediction lead times

Apart from this, it can be seen from the literature that most of the water-consumption predictions focused on short-term (24–48 h) predictions. This shows that medium-term predictions (usually up to 12 months), which are required for planning large-scale maintenance activities, have not been given due attention. Therefore this study also addresses this gap by undertaking a medium-term water-demand prediction. The performance of the three models was investigated with respect to the observed data by using root-mean-square error (RMSE) and the Nash-Sutcliffe coefficient of efficiency (CE) (Nash and Sutcliffe 1970) as evaluation criteria.

Material and Methods

Data Collection and Analysis

This study was conducted using a total of 23 years of monthly water-consumption data (January 1992–December 2014) that were obtained from Istanbul Metropolitan Municipality Water Works and Canalization Administration (ISKI). The observed monthly water-consumption data (original data) are depicted in Fig. 1. In general there is an increasing trend in the monthly water consumption of Istanbul with increasing fluctuations depending on season, making it nonstationary. However, there was a drop in the water consumption of Istanbul in the summer months of 1994 and from 2007 to 2009 as a whole. Based on the information obtained from ISKI, this was because of the low amount of water supplied to the city as a result of drought that reduced the amount of water stored in the reservoirs. The mean, variance, and skewness coefficient values of the observed data were found to be 56.07, 135.75 million m^3/month, and 0.156, respectively.

As part of the analysis, DWT and MSA preprocessing techniques were used to decompose the observed time series into components to remove the trend and seasonal components of the original data, allowing accurate and reliable prediction by employing MLP. The original monthly water-consumption data were decomposed into three wavelets (bands) by DWT as depicted in Fig. 2. Bands 2 and 3 in Figs. 2(b and c) exhibit some level of

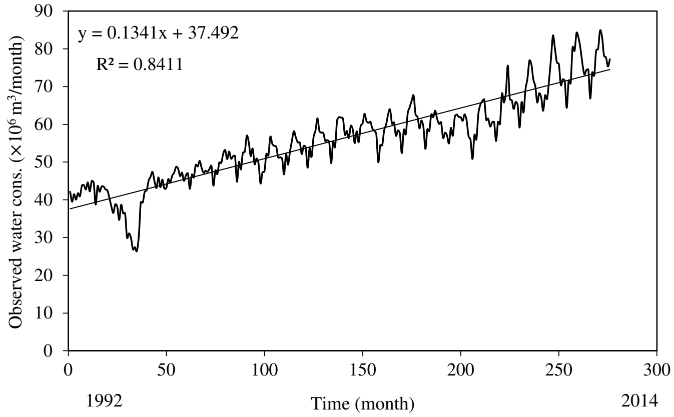

Fig. 1. Monthly water-consumption time series of Istanbul and a linear line showing trend

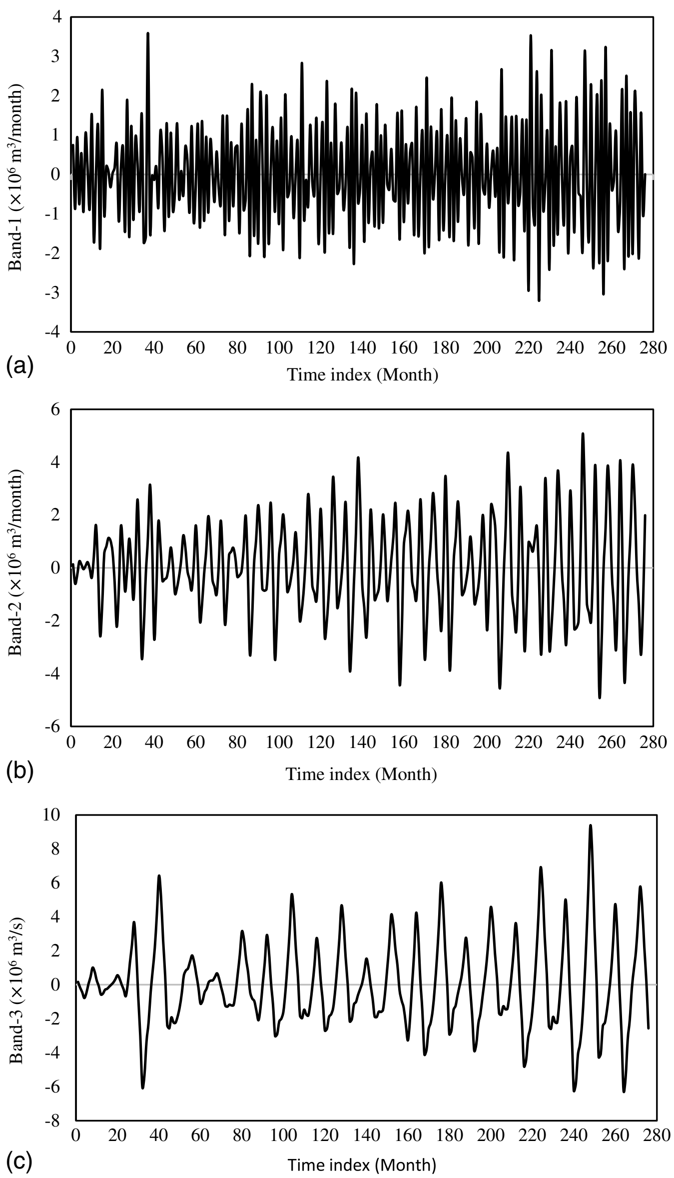

Fig. 2. Plots of bands 1(a), 2(b), and 3(c) of the original monthly water-consumption time series data for Istanbul decomposed using DWT

nonstationarity as the amplitudes increases along the x-axes of the figures. In addition, the observed monthly water consumption data were decomposed into trend-cycle, seasonal-index, and irregular (error-term) components by MSA [Figs. 3(a–c), respectively]. Fig. 3(c) depicts the irregular (error-term) component of the observed time series as decomposed by MSA. It can be seen from this figure that the new time series is stationary because it does not exhibit any trend or change in variability. The new time series obtained after decomposing the original (raw) data with DWT and MSA were then used as inputs into the MLP modeling tool to form the combined DWT-MLP and MSA-MLP models, respectively. The new model (MSA-MLP) was compared with DWT-MLP and the stand-alone MLP models in terms of its capability in predicting monthly water consumption of Istanbul for prediction lead times of 1, 2, 3, 6, 9, and 12 months by taking the observed value as a reference. In short, three MLP-based models—the stand-alone MLP model, the combined DWT-MP model, and the combined MSA-MLP model—were developed and their performance in predicting monthly water consumption for prediction lead times of up to 12 months was assessed.

The monthly water-consumption data were divided into two groups: observed consecutive monthly water-consumption data for the first 14 years (60%) (January 1992–December 2005) were used for training (calibrating) the models and the remaining nine years (40%) of observed consecutive monthly water-consumption data (January 2006–December 2014) were used for testing (validating) the performance of the three models. The RMSE and CE were used as performance-evaluation criteria of the models during both the calibration and testing steps. On top of these, visual investigation was implemented in evaluating the performance of the

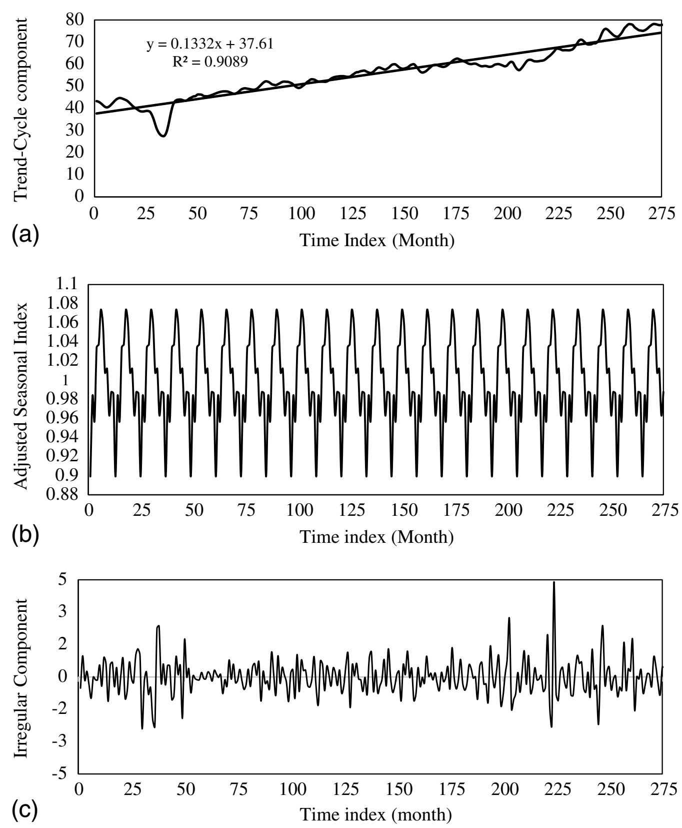

Fig. 3. Time series of the (a) trend-cycle component; (b) adjusted seasonal index; (c) irregular component of the monthly water consumption of Istanbul (10^6 m^3/month)

models with the help of scatter plots of observed data versus results of model predictions. Brief descriptions of the preprocessing tools, the developed models, and the performance-evaluation criteria are given in the following section.

Description of the Preprocessing Methods

Wavelet Transform

Wavelet transforms are mathematical functions that allow one to unfold a time series into both space and scale, and possibly directions based on group theory and square integrable representations (Farge 1992). Wavelet transforms use analyzing functions, called wavelets, which are mathematical functions that give a time-frequency representation of a time series in the time domain and their relationships in order to analyze time series that contain nonstationarity Partal and Kisi (2007). Wavelet analyses allow the decomposition of complex information into simple forms at different positions and scales and subsequent reconstruction with high precision (Sifuzzaman et al. 2009). In other words, wavelet analyses decompose signals (both stationary and nonstationary) into shifted and scaled versions of the basic (original) wavelet, called a mother wavelet. Each wavelet obtained after decomposition carries information of the original time series but with a distinct contribution to the original time series (Wang and Ding 2003). According to Ramana et al. (2013), wavelet analysis is used to explore, denoise, and smoothen a certain time series (signal), and it is used as a preprocessing tool in predictions and other empirical analyses because it eliminates the trend and periodicity of the time series (Misiti et al. 2000). Wavelet analysis is also known to be capable of revealing aspects of the original data such as trends, breakdown points, and discontinuities that other signal-analysis techniques might miss (Adamowski and Sun 2010).

The basic (mother) wavelet transform function, $\psi(t)$, is expressed as

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{1}$$

This wavelet is dilated and contracted to calculate scale decomposition, $\psi_{a,b}(t)$, as

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right) \tag{2}$$

where s = real number that represents a scale or frequency parameter and is taken as dilation when $s > 1$ or contraction when $s < 1$; and τ = real number that denotes a translation parameter, and is interpreted as shift of the wavelet transform function.

Wavelet transform has been successfully used as a preprocessing tool in many studies involving prediction of hydrological components. The prediction accuracy of modeling tools has been found to be enhanced when wavelet analysis is applied as a preprocessing tool (Altunkaynak 2014). The literature mainly uses two types of wavelet transforms, continuous wavelet transform (CWT) and DWT.

According to Polikar et al. (2001), CWT is used to decompose a continuous time function (signal) into wavelets, and it is designed to work with functions defined over the whole real axis. The CWT function of a signal $x(t)$ is expressed mathematically, according to Cannas et al. (2006), as

$$\mathrm{CWT}_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}}\int_{-\infty}^{+\infty} x(t)\psi * \left(\frac{t-\tau}{s}\right)dt \tag{3}$$

where s = scale parameter; τ = translation parameter; and $*$ = complex conjugate.

Because of the computation required over the whole real axis in CWT, the calculation of CWT coefficients requires a large amount of work and a long computation time. In order to avoid this limitation of CWT, another transform—DWT—was developed. DWT deals with functions that are defined over a range of integers for which the wavelets are discretely sampled and it results in wavelet coefficients only at selected scales and periods. This provides many advantages because fewer but still representative coefficients that give the variable's frequency-scale information over time are achieved. According to Sharma et al. (2014), DWT is described as an implementation of wavelet transform using a discrete set of wavelet scales and translations obeying some defined rules. DWT decomposes the signal into a mutually orthogonal set of wavelets (Adamowski and Sun 2010). This is achieved by modifying the wavelet representation into the following form (Cannas et al. 2006):

$$\psi_{j,k}(t) = \frac{1}{\sqrt{S_o^j}}\psi\left(\frac{t-k\tau_o S_o^j}{S_o^j}\right) \tag{4}$$

where j and k = integers and S_o = fixed dilation step and is >1.

DWT has been found to be an effective preprocessing tool that could improve prediction accuracy of modeling tools in hydrological studies. However, according to Fernandes (2001), it is also known to have three serious limitations in addition to the complex procedures involved in its computation:

- DWT is shift sensitive because input-signal shifts generate unpredictable changes in its coefficients,
- It suffers from poor directionality because its coefficients reveal only three spatial orientations, and
- Its analysis lacks the phase information that accurately describes nonstationary signal behavior.

This study used the Morlet wavelet because it is the most widely employed wavelet function among the many mother wavelets available in the literature and it provides a good balance between time and frequency localizations (Altunkaynak 2014; Altunkaynak and Ozger 2016).

Multiplicative SA

An SA is a technique that is used to decompose the original time series into its trend-cycle, seasonal, and irregular components. Two types of SA are presented by Shiskin et al. (1967): MSA and additive season algorithm (ASA).

The MSA and ASA are defined as, respectively

$$X_t = T_t \,.\, S_t .\, I_t \tag{5}$$

$$X_t = T_t + S_t + I_t \tag{6}$$

where X_t = original variable; T_t = trend-cycle component; S_t = seasonal component; and I_t = irregular component, or what is literally taken as the error term.

An SA is mainly employed to eliminate the seasonal patterns from the time series of the original variable so that reliable and accurate prediction results can be obtained. The following three procedures need to be followed in order to determine the seasonal component (IBM SPSS 2011):

1. Apply the centered moving average method to the original time series to smoothen it; the new series generated after the application of the centered moving average reflects the trend-cycle component.

J. Water Resour. Plann. Manage., 2017, 143(6): 04017011

2. Divide the original time series by the smoothed series to obtain the seasonal-irregular component if the algorithm used is multiplicative, or subtract the smoothed values from the original series if the algorithm is additive.
3. Compute the medial average (centered average, defined as defined as the mean value of a certain variable determined for the central location) of the specific seasonal relatives for each unit of periods if the model is multiplicative (additive) in order to isolate the seasonal component from the seasonal-irregular term.

An MSA is recommended when there is a visible trend in the time series of the original variable and ASA is recommended otherwise. Since there is a clear, increasing trend in the water-consumption time series of Istanbul, MSA is employed in this study. Details of the procedures mentioned earlier for MSA are given as follows:

1. Determine the periodicity (P) of the time series. Periodicity refers to the time span after which the value of an original variable is expected to repeat itself in normal conditions. For example, the monthly water consumption time series used in this study has a P of 12 months.
2. Apply the medial (centered) moving average technique to smoothen the original data. The smoothed time series is usually represented by Z_t.

   If P is odd

$$Z_t = \frac{1}{P}\sum_{j=t-[(P-1)/2]}^{t+[(P-1)/2]} X_j, \quad \text{where } t = \frac{P+1}{2}, \frac{P+3}{2}, \frac{P+5}{2}, \ldots \tag{7}$$

   If P is even

$$Z_{t+0.5} = \frac{1}{P}\sum_{j=t-[(P-1)/2]}^{t+[(P-1)/2]} X_j, \quad \text{where } t = \frac{P}{2}, \frac{P}{2}+1, \frac{P}{2}+2, \frac{P}{2}+3, \ldots \tag{8}$$

   Then Z_t is determined as

$$Z_t = \frac{Z_{t+0.5} + Z_{t+1.5}}{2} \tag{9}$$

3. Calculate the seasonal irregular component (SI_t) as the ratio of the original time series to the smoothed series as

$$\mathrm{SI}_t = \frac{X_t}{Z_t} \tag{10}$$

4. Determine the unadjusted seasonal index (USI) as a mean of the seasonal irregular component values by taking into consideration P in Step 1. This is carried out to determine the relative effect of each seasonal irregular value. For instance, if SI_{13}, SI_{25}, and SI_{37} represent the SI values for January, then the USI value for January is taken as the arithmetic mean of these three SI values.
5. Determine new values of indices, called the adjusted seasonal index (ASI), by dividing each USI determined in Step 4 by the mean USI.
6. Determine the seasonally adjusted series (SAS) by dividing the original time series by the corresponding ASI values determined in Step 5 as

$$\mathrm{SAS} = \frac{X_t}{\mathrm{ASI}} \tag{11}$$

7. Determine the smoothed trend-cycle (STC_t) component by using the following mathematical expressions:

$$\mathrm{STC}_t = \frac{1}{9}[(\mathrm{SAS})_{t-2} + 2(\mathrm{SAS})_{t-1} + 3(\mathrm{SAS})_t + 2(\mathrm{SAS})_{t+1} + (\mathrm{SAS})_{t+2}], \qquad t = 3,4,5,\ldots,n-2 \tag{12}$$

   For the first two end points

$$\mathrm{STC}_2 = \frac{1}{3}[(\mathrm{SAS})_1 + (\mathrm{SAS})_2 + (\mathrm{SAS})_3] \quad \text{and}$$

$$\mathrm{STC}_{n-1} = \frac{1}{3}[(\mathrm{SAS})_{n-2} + (\mathrm{SAS})_{n-1} + (\mathrm{SAS})_n]$$

   and for the last two end points

$$\mathrm{STC}_1 = \left[(\mathrm{STC})_2 + \frac{1}{2}[(\mathrm{STC})_2 - (\mathrm{STC})_3]\right] \quad \text{and}$$

$$\mathrm{STC}_n = \left[(\mathrm{STC})_{n-1} + \frac{1}{2}[(\mathrm{STC})_{n-1} - (\mathrm{STC})_{n-2}]\right]$$

8. Calculate the irregular (error) component as the difference between the SAS and the corresponding STC values.

As can be understood from these steps, SA involves simple arithmetical expressions and therefore it can be employed for data decomposition even with a hand calculator or with ordinary software like MS *Excel*. In general, SA involves clear and straightforward procedures and its application does not consume much time because it does not involve curve fitting and determination of coefficients.

Description of the Models

Stand-Alone MLP Model

In recent years, ANNs have become a widely used tool for predicting various hydrological variables. An ANN is a network that is inspired by the way neuron systems in the brain are organized and process information for decision making (Madala and Ivakhnenko 1994). It can be literally taken as a mathematical expression developed based on the nervous system of a brain. Hsu et al. (1995) defined an ANN as a flexible mathematical structure that is capable of identifying complex nonlinear relationships between input and output data sets.

An ANN involves a number of interconnected processing elements (neurons or nodes) working in harmony to solve complicated problems. The organization of these nodes according to a particular arrangement determines the architecture of an ANN. Basically, it consists of three distinct layers: input layer, hidden layer, and output layer. The input data are introduced at the input layer. The hidden layer is a layer or layers where the input data are processed, and the results of the process are obtained at the output layer. According to Altunkaynak (2007, 2014), the nodes are generally arranged in layers to provide information flux from the input layer to the output layer, passing through the hidden layer or layers.

The topology of an ANN includes the number of layers in the network, the number of nodes in each layer, the type of connection between nodes, and the overall structure (Pereira Filho and dos Santos 2006). Three components are considered during the design and characterization of ANNs. These components are nodes, weights, and activation (transfer) functions. Weights that are optimized through an iterative process during the model calibration phase are put between nodes and the activation function controls the generation of outputs in the nodes (Diamantopoulou et al. 2005).

Radial basis function (RBF) and MLP are the two types of predictive processes in ANNs. The RBF type of neural networks have an input layer, a hidden layer, and an output layer; this makes them simpler than the MLP networks, which are known to consist of a set of source nodes forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes (Pereira Filho and dos Santos 2006). Despite their more complex nature, however, MLP networks are the most widely used type of ANN (Cherkassky et al. 1993) because their complex nature enables them to handle complex nonlinear interactions (Altunkaynak 2013). Moreover, MLP-based modeling tools have been found to perform better than RBF-based models in a number of hydrological studies (Seckin et al. 2013; Santos et al. 2013). As a result, the MLP type of ANN was selected for this study.

Altunkaynak and Strom (2009) gave the following equation to define an MLP by assuming a model having m number of input neurons $(x_1, x_2, \ldots, x_m)$, r number of neurons in the hidden layer $(h_1, h_2, \ldots, h_r)$, and p number of output neurons $(z_1, z_2, \ldots, z_p)$:

$$h_j = f\left(\sum_{i=1}^{m} x_i w_{ij} + \alpha_j\right) \tag{13}$$

where

$$z_k = g\left(\sum_{j=1}^{r} h_j \theta_{jk} + \beta_k\right) \tag{14}$$

Here, $f(.)$ and $g(.)$ = activation functions; α_j = bias for neuron h_j; β_k = bias for neuron z_k; w_{ij} = weight of the connection from neuron x_i to neuron h_j; θ_{jk} = weight of the connection from neuron h_j to z_k; and i, j, and k = indices that represent the input, hidden, and output layers, respectively.

Dawson and Wilby (2001) stated that the following steps need to be followed when developing ANN models: (1) selection of the input variables, (2) selection of ANN type, (3) selection of data and preprocessing, (4) ANN training (calibration), and (5) ANN testing (prediction) and performance assessment. However, determining the optimum architecture of the ANN model that captures the relationship between the input and the output parameters needs to be the primary objective during the development of the model (Adamowski 2008). Determining the number of neurons in the input and output layers is simple because it depends on the input and output parameters used to model the physical process (Adamowski 2008).

The MLP model consists of a group of weighting coefficients between the input layer(s) and the hidden layer(s) and between the hidden layer(s) and the output layer(s). According to Altunkaynak (2013), the optimal values of these coefficients need to be identified and then used as model parameters so that the model's prediction error is minimized. The MLP model can be trained (calibrated) based on observed data so that the optimum values of these weights of connection can be determined. There are several training algorithms in the literature. In this study, however, the backpropagation training algorithm (Rumelhart et al. 1986) was employed to obtain the optimal weighting coefficients of the MLP architecture.

The number of neurons in the hidden layer was determined through a trial-and-error procedure by making use of the observed data. While doing so, various numbers of hidden layers and neurons were taken into considerations and were tested with the objective of determining the best architecture of the MLP model for this particular study. After passing these procedures, the best architecture of the MLP model determined for modeling the monthly water consumption of Istanbul resulting in the minimum prediction error was found to consist of one input layer [with two input nodes, $D(t-1)$ and $D(t-2)$], one hidden layer with three neurons, and an output layer with one node $[D(t)]$, as depicted in Fig. 4. Moreover, after testing the tansig, logsig, and pureline activation functions by taking into consideration the lowest prediction error, tansig and pureline were selected as activation functions for the hidden layer and output layer, respectively. This same architecture was used for all prediction lead times when combined with the preprocessing techniques. In this study, the stand-alone MLP model uses unprocessed data as input.

Combined DWT-MLP Model

As indicated in the Data Collection and Analysis section, the original monthly water-consumption time series of Istanbul was decomposed by DWT into three subseries (bands), each having distinct behavior and playing a different role in the original time series. These three subseries were used as inputs individually into the MLP model that was used as a modeling tool; this was named the combined wavelet–MLP model (W-MLP). The MLP was used as a modeling tool that establishes the mathematical form of the model relating inputs and outputs. In this study, based on the best architecture of the MLP model, the subseries at time $t-1$ and $t-2$ were used as inputs of MLP and the original time series at time t was determined as the sum of the three outputs. Prediction of the

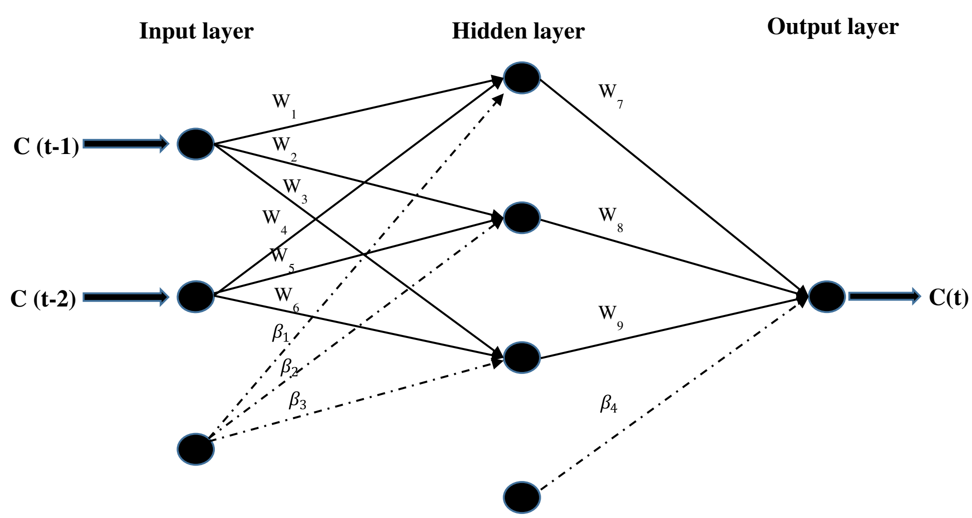

Fig. 4. Architecture of the best MLP

monthly water-consumption value was performed for lead times of 1, 2, 3, 6, 9, and 12 months from two previous monthly water-consumption values.

The following three steps were followed in the combined DWT-MLP approach for predicting the monthly water demand of Istanbul:

1. The bands of the original data that were used for modeling the objective were selected after assessing significant spectral bands from average wavelet spectra;
2. Each band was modelled individually using MLP; and
3. Relevant monthly water-consumption time series were obtained by reconstructing the modelled bands.

Combined MSA-MLP Model

The combined MSA-MLP model was developed by using the irregular component of the observed data determined by applying MSA as an input into the MLP model; this was named the combined MSA-MLP model. This means that, first, the trend-cycle and the seasonal components of the observed time series data were eliminated by MSA in order to obtain a new time series representing the irregular component. Then the time series of the irregular component of times $t-1$ and $t-2$ were used as inputs of the MLP, which was used to establish the mathematical relationship between the input and the output data (the irregular component at time t, in this case). The original time series at time t was then determined as the product of the adjusted seasonal component and the sum of the predicted irregular component and the trend-cycle component. The combined MSA-MLP model is proposed as a new alternative to the DWT-MLP model for predicting the monthly water consumption of Istanbul. The development of this model is simpler than the development of the DWT-MLP model because it does not involve complex mathematical procedures (as opposed to DWT). In addition, it was proposed as a way of improving the prediction accuracy and lead times better than the stand-alone MLP and the combined DWT-MLP models.

Performance-valuation Criteria

The performance of models is evaluated by using various efficiency-evaluation criteria during both the calibration and validation phases of modeling. According to Solomatine and Shrestha (2009), among the available evaluation criteria in the literature, RMSE and CE are the most widely used. As a result, these evaluation criteria were selected as techniques for assessing the performance of the models developed in this study.

According to Bowden et al. (2012), RMSE is a good indicator of the prediction accuracy of a model. The following equation is used to determine RMSE:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(D_{pi} - D_{oi})^2} \tag{15}$$

where n = total number of observations; D_{pi} = predicted monthly water demand; D_{oi} = observed monthly water demand; $\bar{D}_o$ = mean observed monthly water demand; and $\bar{D}_p$ = mean predicted monthly water demand. The value of RMSE is 0 for perfect predictions and increases to large positive values as the inconsistency between the predicted results and observed data increases.

Another effective way of comparing the goodness of fit that exists between observed data and predicted values of a variable is CE, which is calculated as

$$\mathrm{CE} = \left[1 - \frac{\sum_{i=1}^{n}(D_{pi} - D_{oi})^2}{\sum_{i=1}^{n}(D_{oi} - \bar{D}_o)^2}\right] \tag{16}$$

where the definitions of n, D_{pi}, D_{oi}, and $\bar{D}_o$ are the same as for Eq. (15).

In general, CE values greater than 0.5 are taken as acceptable levels of performance (Moriasi et al. 2007). In addition, Donigan and Love (2003) categorized the performance of a model based on CE values as fair if the CE value is in range 0.65–0.75, good if the CE value is in range 0.75–0.85, and very good if the CE value is greater than 0.85 model. Obviously, a model with large values of R^2 and CE (1 or close to 1) and a small value of RMSE is considered to be a model with high efficiency. Finally, the model with the largest CE and the smallest RMSE was selected as the best model for predicting monthly water consumption of Istanbul and extending prediction lead time.

Results and Discussion

In this study, DWT and MSA were used as preprocessing tools with the objective of eliminating trend and seasonality from the observed data so that prediction accuracy and lead time could be improved. The best configuration of the MLP model used as a modeling tool was found to use two input layers with two neurons (two previous values, $t-1$ and $t-2$), one hidden layer with three neurons, and one neuron for current monthly water consumption at the output layer (t). The preprocessed data were used as inputs to the MLP model to develop the combined DWT-MLP and MSA-MLP models, which were compared with the stand-alone MLP model. The performance of these models was evaluated using RMSE and CE determined by taking into consideration the model prediction results and observed values as performance indicators.

Table 1 depicts the values of RMSE and CE criteria determined in the testing (validation) phase of the models' development for prediction lead times of 1, 2, 3, 6, 9, and 12 months. The RMSE and CE values of the models in the calibration phase are not presented in this study because the results have a similar trend with the corresponding values in the validation phase. The RMSE and CE values of the stand-alone MLP model were found to be the largest and the smallest, respectively, among the RMSE and CE values of the models at all prediction lead times, showing that this model performed least well.

For a prediction lead time of one month, the DWT-MLP and MSA-MLP models resulted in CE values of 0.959 and 0.986, respectively, and RMSE values of 1.64 and 0.94, respectively. These results show that the DWT-MLP and MSA-MLP models performed much better than the stand-alone MLP model, which has CE and RMSE values of 0.779 and 3.78, respectively. This implies that, based on the classification of Donigan and Love

Table 1. Values of the Performance-Evaluation Criteria of the Stand-Alone-MLP, DWT-MLP, and MSA-MLP Models

Lead time (months)	Stand-alone MLP		DWT-MLP		MSA-MLP	
	RMSE	CE	RMSE	CE	RMSE	CE
1	3.78	0.779	1.64	0.959	0.940	0.986
2	6.22	0.404	2.98	0.863	0.965	0.986
3	8.26	−0.059	3.25	0.837	0.962	0.986
6	11.61	−1.084	5.89	0.464	1.034	0.984
9	9.13	−0.287	6.03	0.439	0.981	0.985
12	6.85	0.275	5.75	0.489	1.023	0.984

J. Water Resour. Plann. Manage., 2017, 143(6): 04017011

(2003), the stand-alone MLP model's performance is good but the DWT-MLP and MSA-MLP models performed very well in predicting the monthly water consumption of Istanbul for a prediction lead time of one month. However, the new proposed MSA-MLP model performed the best because it resulted in the highest CE and the lowest RMSE values compared to the MLP and DWT-MLP models.

The difference between the prediction performance of the stand-alone MLP model and the other models becomes evident beginning with a prediction lead time of two months, when the CE value of the stand-alone MLP model was found to be less than 0.5, which shows that the model cannot be used for predicting monthly water consumption of Istanbul based on the criteria set by Moriasi et al. (2007). Obviously, the poor performance of the stand-alone MLP model is the result of the direct use of observed data as inputs to the MLP without any preprocessing. By contrast, the MSA-MLP performed very well, with CE and RMSE values of 0.996 and 0.965, respectively. The performance of the DWT-MLP model was good, with CE and RMSE values of 0.863 and 2.98, respectively.

Similar to the stand-alone MLP model, the performance of the DWT-MLP model becomes unacceptable starting with a prediction lead time of six months, when the CE value becomes less than 0.5. The poor performance of the model at this time range could be directly attributed to the failure of the DWT technique to remove the trend and seasonal components of the original data to the desired level. The performance of the MSA-MLP model was found to be very good, with more or less the highest values of CE and the lowest values RMSE at all prediction lead times. It is evident from from Table 1 that the performance of the models tends to improve beginning with a prediction lead time of nine months and approaching 12 months, where the effects of seasons are expected to repeat themselves.

In addition to the performance-evaluation criteria discussed earlier, 1:1 line scatter diagrams were used to analyze the performance of the three models in predicting monthly water-consumption values. Only scatter plots of prediction lead times of 1, 6, and 12 months are presented in this study as samples (Fig. 5). It is evident from Figs. 5(c, f, and i) that the monthly water-consumption values predicted by MSA-MLP for prediction lead times of 1, 6, and 12 months are very close to the corresponding observed values. The DWT-MLP model predictions for a lead time of 1 month [Fig. 5(b)] seem to be closer to the corresponding observed values. From Fig. 5(d) it can be seen that the stand-alone MLP model underestimates observed monthly water-consumption values for a prediction lead time of six months. Similar trends (underestimating the monthly water consumption of Istanbul for values greater than 63 million $m^3/month$) can be seen in the performance of the DWT-MP6 [Fig. 5(e)], MP12 [Fig. 5(g)], and DWT-MP12 [Fig. 5(h)] models, even though the DWT-MP12 model seems to perform better. It can be understood from the above analysis that, even though decomposing the original data into wavelets by DWT can remove trends and periodicities (seasonalities), it may not be good enough to handle this task to the desired level, because the performance the DWT-MLP model was found to be unacceptable,

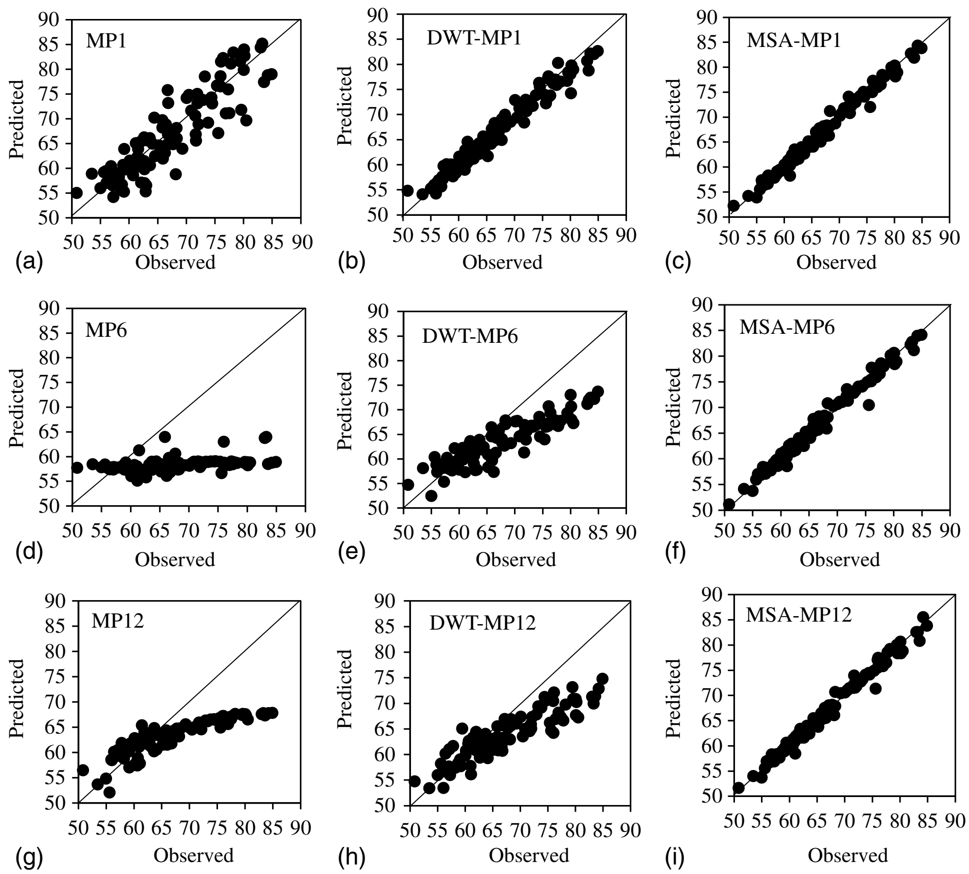

Fig. 5. Scatter plots of the observed (a) stand-alone MLP; (b) DWT-MLP; and (c) MSA-MLP models for a prediction lead time of 1 month, the (d) stand-alone MLP; (e) DWT-MLP; and (f) MSA-MLP models for a prediction lead time of 6 months; and the (g) stand-alone MLP; (h) DWT-MLP; and (i) MSA-MLP models for a prediction lead time of 12 months (values on the axis are 10^6 $m^3/month$)

especially for predictions with longer lead times. By contrast, the performance of the MSA-MLP model was found to be high throughout the entire prediction lead time.

According to Kim et al. (2004), stationarity of a time series is the basic underlying assumption in the practical application of ANN-based models for prediction purposes. In relation to this, the high prediction accuracy obtained by applying the MSA-MLP model could be related to the ability of SA to convert nonstationary time series into a stationary time series by removing seasonal-index and trend-cycle components better than wavelet transform, although it had been identified as an effective preprocessing tool in studies involving nonstationary time series. The application of the MA technique clearly removes the irregular component and periodicity. In addition, the division of the original time series by the new time series in the case of MSA removes the trend component and therefore converts a nonstationary time series into stationary time series. The performance demonstrated by SA in this study agrees with the result obtained by Altunkaynak and Nigussie (2015, 2016) despite the differences in the parameters considered in the studies. From this, it can be suggested that SA may be a superior alternative to wavelet transform for preprocessing data used in developing models for predicting hydrological parameters.

Conclusion

In this study, three MLP-based models (the stand-alone MLP, the combined DWT-MLP that combines DWT with the MLP model, and the combined MSA-MLP that combines MSA with MLP) were developed for predicting the monthly water consumption of Istanbul. The RMSE and CE values were used as evaluation criteria for comparing the performance of the three models in terms of prediction accuracy and their ability to extend prediction lead time. It can be concluded from the results that the stand-alone MLP model cannot be used to the predict monthly water consumption of Istanbul beyond a prediction lead time of one month, and thus there is a need to apply preprocessing tools to the original data before using them as inputs for prediction purposes. The combined MSA-MLP was found to perform the best throughout the prediction lead times, showing that MSA is a simpler but better preprocessing technique than DWT in removing nonstationarity from the original data, increasing the accuracy of the MLP model, and extending prediction lead time. However, it should be noted that the study was carried out by using monthly water consumption time series from one urban area; more studies need to be undertaken to determine how MSA will perform in various geographical and climatic regions in order to generalize this result. In addition, considering the dynamic nature of water demand, further study needs to be undertaken to determine how MSA will perform in comparison with other preprocessing tools for short-term water-demand prediction, such as at hourly and daily time scales. Moreover, the performance of MSA in terms of improving the prediction accuracy and extending the prediction lead time of MLP when weather data and other socioeconomic factors are used as inputs needs to be investigated.

References

Adamowski, J., and Sun, K. (2010). “Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds.” *J. Hydrol.*, 390(1), 85–91.

Adamowski, J. F. (2008). “Peak daily water demand forecast modelling using artificial neural networks.” *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2008)134:2(119), 119–128.

Adebiyi, A. A., Adewumi, A. O., and Ayo, C. K. (2014). “Comparison of ARIMA and artificial neural networks models for stock price prediction.” *J. Appl. Math.*, 2014(1) (2014), 1–7.

Akbari, H., and Bozorg, H. O. (2012). “Comparison of artificial intelligence based models and methods in monthly flow forecasting.” *Water Eng.*, 5(13), 35–53.

Altunkaynak, A. (2007). “Forecasting surface water level fluctuations of Lake Van by artificial neural networks.” *Water Resour. Manage.*, 21(2), 399–408.

Altunkaynak, A. (2010). “A predictive model for well loss using fuzzy logic approach.” *Hydrol. Process.*, 24(17), 2400–2404.

Altunkaynak, A. (2013). “Prediction of significant wave height using geno-multilayer perceptron.” *Ocean Eng.*, 58, 144–153.

Altunkaynak, A. (2014). “Predicting water level fluctuations in Lake Michigan-Huron using wavelet-expert system methods.” *Water Resour. Manag.*, 28(8), 2293–2314.

Altunkaynak, A., and Nigussie, T. A. (2015). “Prediction of daily rainfall by a hybrid wavelet-season-neuro technique.” *J. Hydrol.*, 529, 287–301.

Altunkaynak, A., and Nigussie, T. A. (2016). “Performance comparison of SAS-multilayer perceptron and wavelet-multilayer perceptron models in terms of daily streamflow prediction.” *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0001263, 04015051.

Altunkaynak, A., and Ozger, M. (2016). “Comparison of discrete and continuous wavelet-multilayer perceptron methods for daily precipitation prediction.” *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0001376, 04016014.

Altunkaynak, A., Ozger, M., and Cakmakci, M. (2005). “Water consumption prediction of Istanbul city by using fuzzy logic approach.” *Water Resou. Manag.*, 19(5), 641–654.

Altunkaynak, A., and Strom, K. B. (2009). “A predictive model for reach morphology classification in mountain streams using multilayer perceptron methods.” *Water Resour. Res.*, 45(12), W12501.

Aly, A., and Wanakule, N. (2004). “Short-term forecasting for urban water consumption.” *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2004)130:5(405), 405–410.

Andrade, M. A., Choi, C. Y., Lansey, K., and Jung, D. (2016). “Enhanced artificial neural networks estimating water quality constraints for the optimal water distribution systems design.” *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000663, 04016024.

Arandia, E., Ba, A., Eck, B., and McKenna, S. (2016). “Tailoring seasonal time series models to forecast short-term water demand.” *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000591, 04015067.

Azadeh, A., Neshat, N., and Hamidipour, H. (2012). “Hybrid fuzzy regression-artificial neural network for improvement of short-term water consumption estimation and forecasting in uncertain and complex environments: Case of a large metropolitan city.” *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000152, 71–75.

Bowden, G. J., Maier, H. R., and Dandy, G. C. (2012). “Real-time deployment of artificial neural network forecasting models: Understanding the range of applicability.” *Water Resour. Res.*, 48(10), W10549.

Cannas, B., Fanni, A., Sias, G., Tronei, S., and Zedda, M. K. (2006). “River flow forecasting using neural networks and wavelet analysis.” *Proc., European Geosciences Union. EUG, 2005*, Copernicus Publications, Gottingen, Germany.

Cherkassky, V., Friedman, J. H., and Wechsler, H. (1993). *From statistics to neural networks*, Springer, Berlin.

Dabhi, V. K., and Chaudhary, S. (2014). “Hybrid wavelet-postfix-GP model for rainfall prediction of Anand region of India.” *Adv. Artif. Intell.*, 2014, 1–11.

Dawson, C. W., and Wilby, R. L. (2001). “Hydrological modelling using artificial neural networks.” *Prog. Phys. Geogr.*, 25(1), 80–108.

Diamantopoulou, M. J., Antonopoulos, V. Z., and Papamichail, D. M. (2005). “The use of a neural network technique for the prediction of water quality parameters of Axios River in northern Greece.” *Eur. Water*, 11/12, 55–62.

Donigan, A. S., and Love, J. T. (2003). "Sediment calibration procedures and guidelines for watershed modelling." ⟨http://www.hspf.com/pdf/TMDL2003PaperDonigian.pdf⟩ (Oct. 10, 2015).

Donkor, E. A., Mazzuchi, T. A., Soyer, R., and Roberson, J. A. (2014). "Urban water demand forecasting: Review of methods and models." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000314, 146–159.

Farge, M. (1992). "Wavelet transforms and their applications to turbulence." *Annu. Rev. Fluid Mech.*, 24(1), 395–458.

Fernandes, F. (2001). "Directional, shift-insensitive, complex wavelet transforms with controllable redundancy." ⟨http://dsp.rice.edu/publications/directional-shift-insensitive-complex-wavelet-transforms-controllable-redundancy⟩ (Oct. 9, 2015).

Ghiassi, M., Zimbara, D., and Saidane, H. (2008). "Urban water demand forecasting with a dynamic artificial neural network model." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2008)134:2(138), 138–146.

Hsu, K. L., Gupta, H. V., and Sorooshian, S. (1995). "Artificial neural network modelling of the rainfall-runoff process." *Water Resour. Res.*, 31(10), 2517–2530.

Huang, L., Zhang, C., Peng, Y., and Zhou, H. (2014). "Application of a combination model based on wavelet transform and KPLS-ARMA for urban annual water demand forecasting." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000397, 04014013.

IBM SPSS. (2011). "IBM SPSS statistics 20 Algorithms." ⟨ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf⟩ (Dec. 12, 2015).

Kim, T. Y., Oh, K. J., Kim, C., and Do, J. D. (2004). "Artificial neural networks for non-stationary time series." *Neurocomputing*, 61(1), 439–447.

Kisi, O. (2005). "Daily river flow forecasting using artificial neural networks and auto-regressive models." *Turk. J. Eng. Environ. Sci.*, 29(1), 9–20.

Madala, H. R., and Ivakhnenko, A. G. (1994). *Inductive learning algorithms for complex systems modelling*, CRC Press, Boca Raton, FL.

Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications." *Environ. Modell. Software*, 15(1), 101–124.

Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J. M. (2000). "Wavelet toolbox user's guide, the math works." ⟨http://mathworks.com/help/pdf_doc/wavelet/wavelet_ug.pdf⟩ (Sep. 19, 2015).

Mohammed, J. R., and Ibrahim, H. M. (2012). "Hybrid wavelet artificial neural network model for municipal water demand forecasting." *ARPN J. Engin. and Appl. Sci.*, 7(8), 1047–1065.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D., and Veith, T. L. (2007). "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Trans. ASABE.*, 50(3), 885–900.

Msiza, I. S., Nelwamondo, F. V., and Marwala, T. (2007). "Artificial neural networks and support vector machines for water demand time series forecasting." *IEEE Int. Conf. on Systems, Man and Cybernetics*, IEEE, New York.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models. Part 1: A discussion of principle." *J. Hydrol.*, 10(3), 282–290.

Odan, F., and Reis, L. (2012). "Hybrid water demand forecasting model associating artificial neural network with Fourier series." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000177, 245–256.

Partal, T., and Kisi, O. (2007). "Wavelet and neuro-fuzzy conjunction model for precipitation forecasting." *J. Hydrol.*, 342(1), 199–212.

Pereira-Filho, A. J., and dos Santos, C. C. (2006). "Modelling a densely urbanized watershed with an artificial neural network, weather radar and telemetric data." *J. Hydrol.*, 317(1–2), 31–48.

Polebitski, A. S., and Palmer, R. N. (2010). "Seasonal residential water demand forecasting for census tracts." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000003, 27–36.

Polikar, R., Upda, L., Upda, S. S., and Honavar, V. (2001). "Learn++: An incremental learning algorithm for supervised neural networks." *IEEE Trans. Syst. Man Cybern. Part C. Appl. Rev.*, 31(4), 497–508.

Ramana, R. V., Krishna, B., Kumar, S. R., and Pandey, N. G. (2013). "Monthly rainfall prediction using wavelet-neural network analysis." *Water Resour. Manage.*, 27(10), 3697–3711.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning representations by back-propagating errors." *Nature*, 323(6088), 533–536.

Santos, R., Rupp, M., Bonzi, S., and Filetia, A. M. (2013). "Comparison between multilayer feedforward neural networks and a radial basis function network to detect and locate leaks in pipelines transporting gas." *Chem. Eng. Trans.*, 32, 1375–1380.

Sastri, T., and Valdes, J. B. (1989). "Rainfall intervention for on-line applications." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(1989)115:4(397), 397–415.

Seckin, G., Akoz, M. S., Cobaner, M., and Haktanir, T. (2009). "Application of ANN techniques for estimating backwater through bridge constrictions in Mississippi River basin." *Adv. Eng. Software*, 40(10), 1039–1046.

Seckin, N., Cobaner, M., Yurtal, R., and Haktanir, T. (2013). "Flood flow at ungauged sites: The case of East Mediterranean river basin, Turkey." *Water Resour. Manag.*, 27(7), 2103–2124.

Sharma, A., Sharma, R., and Toshniwal, S. (2014). "Efficient use of bi-orthogonal wavelet transform for cardiac signals." *Int. J. Comput. Appl.*, 89(8), 19–23.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967). "The X-11 variant of the census method II seasonal adjustment program." ⟨https://www.census.gov/ts/papers/ShiskinYoungMusgrave1967.pdf⟩ (Dec. 6, 2015).

Sifuzzaman, M., Islam, M. R., and Ali, M. Z. (2009). "Application of wavelet transform and its advantages compared to Fourier transform." *J. Phys. Sci.*, 13, 121–134.

Solomatine, D. P., and Shrestha, D. L. (2009). "A novel method to estimate model uncertainty using machine learning techniques." *Water Resour. Res.*, 45(12), W00B11.

Tian, D., Martinez, C. J., and Asefa, T. (2016). "Improving short-term urban water demand forecasts with reforecast analog ensembles." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000632, 04016008.

Tiwari, M. K., and Adamowski, J. F. (2015). "Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000454, 04014053.

Wang, W., and Ding, J. (2003). "Wavelet network model and its application to the prediction of hydrology." *Nat. Sci.*, 1(1), 67–71.

Young, C. C., and Liu, W. C. (2015). "Prediction and modelling of rainfall-runoff during typhoon events using a physically-based and artificial neural network hybrid model." *Hydrol. Sci. J.*, 60(12), 2102–2116.

Young, C. C., Liu, W. C., and Hsieh, W. L. (2015). "Predicting the water level fluctuation in an alpine lake using physically based, artificial neural network, and time series forecasting models." *Math. Prob. Eng.*, 2015, 1–11.

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