Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada

Jan Adamowski,[1] Hiu Fung Chan,[1] Shiv O. Prasher,[1] Bogdan Ozga-Zielinski,[2] and Anna Sliusarieva[1]

Received 26 August 2010; revised 17 October 2011; accepted 21 November 2011; published 20 January 2012.

[1] Daily water demand forecasts are an important component of cost-effective and sustainable management and optimization of urban water supply systems. In this study, a method based on coupling discrete wavelet transforms (WA) and artificial neural networks (ANNs) for urban water demand forecasting applications is proposed and tested. Multiple linear regression (MLR), multiple nonlinear regression (MNLR), autoregressive integrated moving average (ARIMA), ANN and WA-ANN models for urban water demand forecasting at lead times of one day for the summer months (May to August) were developed, and their relative performance was compared using the coefficient of determination, root mean square error, relative root mean square error, and efficiency index. The key variables used to develop and validate the models were daily total precipitation, daily maximum temperature, and daily water demand data from 2001 to 2009 in the city of Montreal, Canada. The WA-ANN models were found to provide more accurate urban water demand forecasts than the MLR, MNLR, ARIMA, and ANN models. The results of this study indicate that coupled wavelet-neural network models are a potentially promising new method of urban water demand forecasting that merit further study.

Citation: Adamowski, J., H. Fung Chan, S. O. Prasher, B. Ozga-Zielinski, and A. Sliusarieva (2012), Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada, *Water Resour. Res.*, *48*, W01528, doi:10.1029/2010WR009945.

1. Introduction

[2] The city of Montreal, along with other large urban cities in Canada and elsewhere, needs to find more effective methods to optimize the operation and management of its existing water supply system in addition to exploring the implementation of water demand management programs to decrease both peak and average urban water demand. Even though it may not appear that Canada lacks fresh water, experts are seeing a growing problem involving municipal water and sewer infrastructure in Canada. Many urban water supply systems in Canada are becoming stressed due to increasing peak and total water demand, a lack of new water sources, climate change, and other environmental and socioeconomic factors. Water loss in urban water supply distribution networks is also becoming a problem in certain cities in Canada. Many water supply distribution networks currently being used in Canada were built prior to or just after World War II, and many have not been properly upgraded or maintained [*Maas*, 2003]. Moreover, in Canada the use of water per capita is one of the highest in the world (it ranks 15th out of 16 peer countries and earns a "D" grade), and a false illusion of fresh water abundance exists among Canadians. Although Canada has 7% of the world's fresh water and only 0.5% of the world's population [*Environment Canada*, 2010], most of the fresh water flows north in the opposite direction of the main areas of population.

[3] The poor infrastructure, overuse, and lack of new water sources in urban areas in Canada and around the world has increased the need for demand management as a way to ensure sufficient and sustainable water for urban use. Rather than finding additional resources, demand management strategies look at ways to decrease water requirements and to conserve water [*Dziegielewski and Baumann*, 1992]. An important component of optimizing water supply systems and implementing effective water demand management programs is the accurate forecasting of short-term water demands [*Ghiassi et al.*, 2008].

[4] Since some Canadian water supply distribution networks such as in Montreal are deteriorating, accurate short-term water demand forecasts are becoming increasingly important in helping to find solutions to Canada's urban water supply management problems. For example, water managers find accurate short-term water demand forecasts important because it allows them to: (i) develop a better understanding of the dynamics and underlying factors that affect water use; (ii) manage and optimize maintenance and operating schedules for pumps, wells, reservoirs, and

[1]Department of Bioresource Engineering, McGill University, Ste. Anne de Bellevue, QC H9X 3V9, Canada.
[2]Centre of Hydrology, National Research Institute, Institute of Meteorology and Water Management, Warsaw PL-01-673, Poland.

0043-1397/12/2010WR009945

mains; (iii) balance the needs of water supply, residential/industrial demands, and stream flows for ecosystem health; (iv) analyze the benefits and costs of water conservation, as well as the water saved by imposition of emergency water restrictions; and (v) provide information on when peak day events are likely to occur.

[5] Traditionally, urban water demand has been forecasted by either analyzing historical water consumption or using statistical models with a range of inputs such as population, price of water, income, and various meteorological variables. However, disadvantages of traditional urban water demand forecasting models include the requirement of a large number of input parameters, and the fact that most traditional modeling methods assume the data is linear and stationary. Statistical methods such as multiple linear regression (MLR) and autoregressive moving average (ARIMA) models have traditionally been used for short-term urban water demand forecasting.

[6] Examples of short-term water demand forecast modeling which use regression analysis include *Cassuto and Ryan* [1979], *Hughes* [1980], *Anderson et al.* [1980], and *Maidment and Parzen* [1984]. *Maidment et al.* [1985] developed daily municipal water consumption time series models as a function of rainfall and air temperature. *Maidment and Miaou* [1986] applied their model to the water consumption of nine cities in the United States. *Smith* [1988] developed a time series model to forecast daily municipal water use in Washington. *Miaou* [1990] developed a monthly time series urban water demand model. *Zhou et al.* [2000] developed a time series forecasting model for an urban sector in Australia. More recently, *Wong et al.* [2010] developed a series of seasonal autoregressive moving series models to analyze the structure of daily urban water consumption in Hong Kong. *Wong et al.* [2010] concluded that an increase in the rainfall amount generally resulted in a reduction in seasonal water use, and that the day-of-the-week effect was characterized by higher water use on weekdays as opposed to weekends.

[7] Originally developed by *Ivanchenko* [1970] for use in engineering cybernetics, multiple nonlinear regression (MNLR, or the group method of data handling or GMDH) has been shown to provide accurate forecasting results in various fields. MNLR is traditionally used in fields where seemingly random recurring events need to be accurately and quickly predicted. These areas include economics, finance, medicine, and marketing. The appeal of MNLR is that very high order multiples should be able to approximate complex multivariate functions [*Cogger*, 2010]. One of the first environmental applications of MNLR was in temperature forecasting. *Miyagashi et al.* [1999] determined that neurofuzzy MNLR models performed better than traditional radial basis functions and numerical weather predictions. More recently, *Saraycheva* [2003] developed a modified version of MNLR which was successfully applied to ecological and socioeconomical monitoring problems (including ecological forecasting) in Ukraine. In hydrology, *Emiroglu et al.* [2011] compared NLR, MLR, and ANNs in the context of discharge coefficient prediction for triangular labyrinth wires. This study found that ANNs performed the best out of all models and that NLR performed better than MLR. However, to date no studies have explored MNLR in urban water demand forecasting.

[8] In recent years, artificial neural networks (ANNs) have been introduced in urban water demand forecasting. Two of the main advantages of ANNs over other methods are that their application does not require a priori knowledge of the process, and they are effective with nonlinear data. *Jain and Ormsbee* [2002] examined regression, time series analysis, and ANN models for daily water demand forecasting. The best ANN model in their study based on the daily water demand from the previous day and the daily maximum temperature of the current day. It was concluded that this ANN model performed slightly better than time series and regression–disaggregation models. *Adamowski* [2008a] examined regression, time series analysis, and ANN models for short-term peak water demand forecasting in Ottawa, Canada. It was found that the use of artificial neural networks in forecasting peak daily water demand in the summer months in an area of high outdoor water demand was slightly better than multiple linear regression and time series analysis.

[9] Support vector machines (SVMs) are a relatively new form of machine learning that was originally developed by *Vapnik* [1995] for use in the telecommunications industry. One of the first uses of a SVM in environmental science was for air pollutant forecasting [*Lu et al.*, 2002]. *Wang et al.* [2009] showed that SVMs outperformed ANN models in forecasting monthly river flow discharges. *Khan and Coulibaly* [2006] found that an SVM performed better than MLP ANNs in 3–12 month predictions of lake water levels. *Yu et al.* [2006] successfully predicted flood stages using SVMs and *Han et al.* [2007] found that SVMs performed better than other models for flood forecasting. *Rajasekaran et al.* [2008] used SVMs for storm surge predictions and *Cimen and Kisi* [2009] used SVMs to estimate daily evaporation. SVMs have also been successfully used to forecast hourly streamflow by *Asefa et al.* [2006], and were shown to perform better than ANN [*Wang et al.*, 2009] and ARIMA [*Maity et al.*, 2010] models for monthly streamflow prediction. However, *Msiza et al.* [2008] compared the use of ANNs and SVMs in water demand forecasting in South Africa, and found that ANNs performed significantly better than SVMs.

[10] A problem with presently available data-based methods, including ANNs, is that they have limitations with nonstationary data, which can result in poor forecasting. Relative urban water demand variation is daily, by day of week, by weekend and holiday patterns, by month, and season, and is modified by weather. The pattern of water demand is a nonstationary stochastic time series that may include a nonlinear trend in the mean, nonconstant variance, and discontinuities. Many methods such as ANNs may not be able to handle nonstationary data if preprocessing of the input data is not done. The methods for dealing with nonstationary data are not as advanced as those for stationary data [*Cannas et al.*, 2006]. Two very recent publications [*Solomatine and Ostfeld*, 2008; *Maier et al.*, 2010] on the current status and future directions of the use of data-based modeling in the area of water resources variable forecasting highlighted several critical issues that need to be explored in greater detail, including: (i) the development and evaluation of hybrid model architectures that

attempt to draw on the strengths of different modeling methods and (ii) the development of robust modeling procedures that are able to work with "noisy" data. This research explored these two issues in the context of short-term urban water demand forecasting.

[11] A data analysis technique that effectively deals with the above described type of multiscale and nonstationary behavior is wavelet analysis, which can be used to detect and extract signal variance both in time and scale simultaneously, without any assumptions of stationarity. Wavelet transforms are well suited for dealing with nonstationary time series in forecasting applications since they can automatically localize and filter the nonstationary component of the signal, instead of trying to de-trend or suppress quasi-periodic smooth components as, for example, in the classical nonstationary ARIMA approach. Through the use of wavelet analysis, a water demand series can be decomposed into a few selected component series. The selected component series carry most of the information which can then be selectively used in forecasting. This allows for the removal of most of the noisy data and facilitates the extraction of quasiperiodic and periodic signals in the water demand time series.

[12] In the last decade, wavelet analysis has begun to be investigated in the water resources engineering and hydrology literature. *Wang and Ding* [2003] developed a hybrid model to forecast groundwater levels in China. *Cannas et al.* [2006] developed a hybrid model for monthly rainfall-runoff forecasting in Italy. *Adamowski* [2007, 2008a, 2008c] developed a new method of wavelet and cross wavelet based forecasting for floods. *Kisi* [2008] and *Partal* [2009] developed a hybrid model for monthly flow forecasting in Turkey. *Kisi* [2009] explored the use of wavelet ANN (WA-ANN) models for daily flow forecasting of intermittent rivers, while *Adamowski and Sun* [2010] developed a WA-ANN model for flow forecasting of nonperennial rivers at lead times of one and three days in Cyprus. *Shiri and Kisi* [2010] successfully combined wavelets with ANN and fuzzy interference systems to accurately forecast short-term and long-term streamflow. Overall, these studies found that the WA-ANN models generally outperformed other methods such as MLR, ARIMA, and ANN in hydrological forecasting applications.

[13] To date, no research has been published in the literature that explores the use of the coupled wavelet-neural network method (or multiple nonlinear regression) for short-term urban water demand forecasting. In this research, a method based on coupling discrete wavelet transforms (WA) and artificial neural networks (ANNs) for urban water demand forecasting applications is proposed and tested, and compared to the MLR, MNLR, ANN, and ARIMA methods.

2. Methods

2.1. Multiple Linear Regression

[14] The objective of multiple linear regression (MLR) analysis is to study the relationship between several independent or predictor variables and a dependent or criterion variable. The assumption of the model is that the relationship between the dependent variable Y_i and the p vector of regressors X_i is linear. The following represents a MLR equation [*Pedhazur*, 1982]:

$$Y = a + \beta_1 X_1 + \cdots + \beta_k X_k, \tag{1}$$

where a is the intercept, β is the slope or coefficient, and k is the number of observations. For forecasting purposes, the linear regression equation will fit a forecasting model to an observed data set of Y and X values. The fitted model can be used to make a forecast of the value of Y with new additional observed values of X.

2.2. Multiple Nonlinear Regression

[15] The assumption of multiple nonlinear regression (MNLR) models is that the relationship between the dependent variable Y_i and the p vector of regressors X_i is nonlinear. The following represents a MNLR equation [*Ivakhnenko*, 1970]:

$$Y = a + \beta_1 X_i + \beta_2 X_j + \beta_3 X_i^2 + \beta_4 X_j^2 + \cdots + \beta_k X_i X_j, \tag{2}$$

where a is the intercept, β is the slope or coefficient, and k is the number of observations. For forecasting purposes, the multiple nonlinear regression equation will fit a forecasting model to an observed data set of Y and X values. The fitted model can be used to make a forecast of the value of Y with new additional observed values of X.

2.3. Autoregressive Integrated Moving Average

[16] The autoregressive integrated moving average (ARIMA) methodology has the ability to identify complex patterns in data and generate forecasts [*Box and Jenkins*, 1976]. ARIMA models can be used to analyze and forecast univariate time series data. The ARIMA model function is represented by (p,d,q), with p representing the number of autoregressive terms, d the number of nonseasonal differences, and q the number of lagged forecast errors in the prediction equation. The three steps to develop ARIMA models are identification, estimation, and forecasting. ARIMA (p,d,q) models are defined as follows [*Box and Jenkins*, 1976]:

$$\Phi_p(B)(1-B)^d Y_t = \delta + \theta_q(B)\varepsilon_t, \tag{3}$$

where Y_t is the original value of the time series, ε_t is a random perturbation or white noise (zero mean, constant variant, and covariance zero), B is the backshift operator, δ is a constant value, Φ_P is the autoregressive parameter of order p, θ_q is the moving average parameter of order q, and d is the differentiation order used for the regular or not seasonal part of the series. A model described as (0, 1, 3) signifies that it contains 0 autoregressive (p) parameters, and 3 moving average (q) parameters, which were computed for the series after it was differenced (d) once (1).

2.4. Artificial Neural Networks

[17] An artificial neural network (ANN) is a data-driven process with a flexible mathematical algorithm capable of solving complex nonlinear relationships between input and output data sets. A neural network can be described as a network of simple processing nodes or neurons, interconnected to each other in a specific order, performing simple

numerical manipulations. A neural network can be used to predict future values of possibly noisy multivariate time series based on past histories. ANNs have become popular in the last decade for hydrological forecasting such as rainfall-runoff modeling, streamflow forecasting, groundwater and precipitation forecasting, and water quality issues [*Kisi*, 2004; *Sahoo and Ray*, 2006; *Adamowski*, 2007, 2008a, 2008b; *Banerjee et al.*, 2009; *Pramanik and Panda*, 2009; *Sreekanth et al.*, 2009; *Sethi et al.*, 2010; *Adamowski and Chan*, 2011].

[18] The most widely used neural network is the multilayer perceptron (MLP). In the MLP, the neurons are organized in layers, and each neuron is connected only with neurons in contiguous layers. A typical three-layer feedforward ANN is shown in Figure 1. The input signal propagates through the network in a forward direction, layer by layer. The mathematical form of a three-layer feedforward ANN is given as [*Ozbek and Fidan*, 2009]

$$O_k = g_2\left[\sum_j V_j w_{kj} g_1\left(\sum_i w_{ji} I_i + w_{j0}\right) + w_{ko}\right], \tag{4}$$

where I_i is the input value to node i of the input layer, V_j is the hidden value to node j of the hidden layer, and O_k is the output at node k of the output layer. An input layer bias term $I_0 = 1$ with bias weights w_{j0} and an output layer bias term $V_0 = 1$ with bias weights w_{k0} are included for the adjustments of the mean value at each layer. Two sets of adjustable weights are presented in this equation: w_{ji} controls the strength of the connection between the input node i and the hidden node j, while w_{kj} controls the strength of the connection between the hidden node j and the output node k. g_1 and g_2 represent the activation function for the hidden layer and the output layer, respectively. The activation function is typically a continuous and bounded nonlinear transfer function, for which the sigmoid and hyperbolic tangent functions are usually selected [*Ozbek and Fidan*, 2009]. A detailed overview of ANNs is provided by *Haykin* [1998].

[19] The Levenberg-Marquardt algorithm (LM) is a modification of the classic Newton algorithm. The goal of the LM algorithm is to search for the minimum point of a nonlinear function [*Karul et al.*, 2000]. The LM algorithm is represented by the following equation [*Karul et al.*, 2000]:

$$X_{k+1} = X_k - (J^T J + uI)^{-1} J^T e, \tag{5}$$

where X is the weight of the neural network, J is the Jacobian matrix of the performance criteria to be minimized, u is the learning rate that controls the learning process, and e is the vector of the case error. The LM algorithm performs a curve fitting on the data. The LM algorithm is designed to approach second-order training speed and accuracy without having to compute the Hessian matrix. Second-order nonlinear optimization techniques are usually faster and more reliable compared to most other optimization algorithm

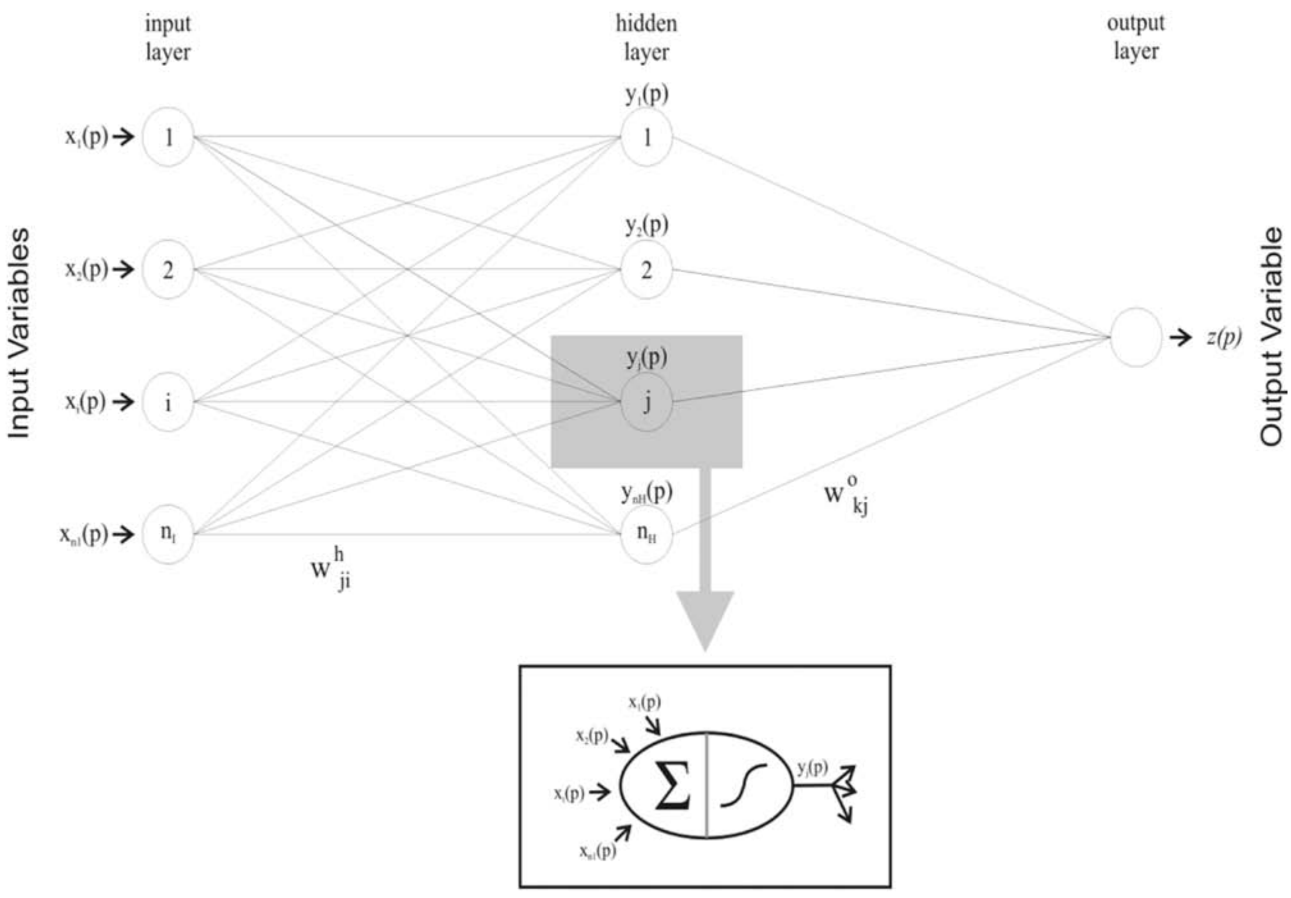

Figure 1. ANN architecture with one hidden layer [*Adamowski and Chan*, 2011].

methods. Several previous research studies have found the performance of the LM algorithm to be superior to other training algorithms such as the conjugate gradient (CG) and gradient descent with momentum (GD) algorithms [*Pramanik and Panda*, 2009; *Adamowski and Karapataki*, 2008; *Adamowski and Chan*, 2011]. Therefore, the LM algorithm was used to train the ANN models in this study. A more detailed description of the application of the LM algorithm to ANN training is provided by *Hagan and Menhaj* [1994].

2.5. Wavelet Analysis

[20] Wavelets are mathematical functions that give a time-scale representation of the time series and their relationships to analyze time series that may contain nonstationarities. The data series is broken down by the transformation into its wavelets, that is a scaled and shifted version of the mother wavelet. Wavelet transform analysis, developed during the last three decades, is often a more effective tool than the Fourier transform (FT) in studying nonstationary time series [*Partal and Kisi*, 2007]. The signal to be analyzed is multiplied with a wavelet function and the transform is computed for each segment generated. However, unlike with the Fourier transform, the width of the wavelet function changes with each spectral component in the wavelet transform. The wavelet transform, at high frequencies, gives good time resolution and poor frequency resolution, while at low frequencies the wavelet transform gives good frequency resolution and poor time resolution.

[21] The continuous wavelet transform (CWT) of a signal $x(t)$ is defined as follows [*Kim and Valdes*, 2003]:

$$CWT_x^{\psi}(\tau, s) = |s|^{1/2}\psi^* \int_{-\infty}^{+\infty} \left(\frac{t-\tau}{s}\right) dt, \tag{6}$$

where s is the scale parameter, τ is the translation parameter, and * denotes the complex conjugate [*Cannas et al.*, 2006]. The translation parameter τ relates to the location of the wavelet function as it is shifted through the signal, which corresponds to the time information in the wavelet transform. The scale parameter s is defined as $|1/\text{frequency}|$ and corresponds to frequency information. Scaling either dilates (expands) or compresses a signal. The mother wavelet $\psi(t)$ is the transforming function. Large scales (low frequencies) dilate the signal and provide detailed information hidden in the signal, while small scales (high frequencies) compress the signal and provide global information about the signal [*Cannas et al.*, 2006].

[22] The wavelet transform performs the convolution operation of the signal and the basis function. The above analysis becomes very useful as in most practical applications, high frequencies (low scales) do not last for a long duration, but instead, appear as short bursts, while low frequencies (high scales) usually last for the entire duration of the signal [*Cannas et al.*, 2006].

[23] The continuous wavelet transform (CWT) calculations require a significant amount of computation time. In contrast, the discrete wavelet transform (DWT) requires less computation time and is simpler to implement than the CWT. DWT scales and positions are usually based on powers of two (dyadic scales and positions) [*Partal and Kisi*, 2007]. This is achieved by modifying the wavelet representation to [*Mallat*, 1999]

$$\psi_{j,k}(t) = s_0^{-j/2}\psi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right), \tag{7}$$

where j and k are integers and $s_0 > 1$ is a fixed dilation step. The effect of discretizing the wavelet is that the time-space scale is now sampled at discrete levels. A value of $s_0 = 2$ was chosen so that the sampling of the frequency axis corresponds to dyadic sampling [*Cannas et al.*, 2006]. The mother wavelet ψ is the "a Trous" wavelet transform algorithm. A translation factor of $\tau = 1$ was selected to ensure the dyadic sampling in the time axis [*Cannas et al.*, 2006]. This power of two logarithmic scaling of the translations and dilations is known as dyadic grid arrangement [*Cannas et al.*, 2006].

2.6. Coupled Wavelet and Artificial Neural Networks (WA-ANNs)

[24] Coupled wavelet and artificial neural network (WA-ANN) models are ANN models which use, as inputs, subseries components which are derived from the use of the discrete wavelet transform (DWT) on the original time series data. The DWT was used in this study since its use involves less computational effort. The decomposition of the time series into multilevels of "details" effectively diagnoses the main frequency component of the signal and abstract local information of the time series [*Minu et al.*, 2010]. The development of the WA-ANN models followed the procedures undertaken by *Cannas et al.* [2006], *Partal* [2009], and *Adamowski and Chan* [2011]. The details of the WA-ANN coupling process is described in section 4.4.

2.7. Model Performance Comparison

[25] The performance of different models can be assessed in terms of goodness of fit after each of the model structures are calibrated using the training/validation data set and testing data set. The coefficient of determination (R^2) measures the degree of correlation among the observed and predicted values. R^2 values range from 0 to 1, with 1 indicating a perfect relationship between the data and the line drawn through them, and 0 representing no statistical correlation between the data and a line. The root mean square error (*RMSE*) evaluates the variance of errors independently of the sample size. *RMSE* indicates the discrepancy between the observed and forecasted values. A perfect fit between observed and forecasted values would have a *RMSE* of 0. The relative root mean square error (*RRMSE*) is an additional indicator of model fit. A lower *RRMSE* signifies a smaller discrepancy relative to the predicted value. *RRMSE* is suggestive of a good model fit. The Nash–Sutcliffe model efficiency coefficient (E) is used to assess the predictive power of hydrological models. An efficiency of one ($E = 1$) corresponds to a perfect match of forecasted data to the observed data. An efficiency of zero ($E = 0$) indicates that the model predictions are as accurate as the mean of the observed data.

3. Study Areas and Data

[26] The City of Montreal provides drinking water to around 1.8 million people in six pressure zones using a

total of 2760 km of primary and secondary distribution network systems. The primary network consists of 440 km of reinforced concrete pipes. Their average age is approximately 47 years old, and 27% of them were constructed before 1920 [*Aubertin*, 2002]. The median age of the secondary network is 52 years old, compared to 40 years for the average Canadian municipality [*Aubertin*, 2002].

[27] The Atwater and Des Baillets plants are the two largest water treatment plants for the City of Montreal, and provide approximately 85% of the total urban water supply in Montreal. The Atwater plant is 100 years old and supplies approximately 680,000 m^3 d^{-1} of potable water, while the Des Baillets plant, built in 1978, supplies approximately 1,159,000 m^3 d^{-1} of potable water.

[28] It has been estimated that Montreal produces approximately 1300 liters per person per day compared to 600 liters per person per day in Toronto [*Aubertin et al.*, 2002; *Rubin*, 2005]. This very high level of production is the result of a number of factors, including the poor condition of the water infrastructure, lack of public awareness, low water consumption fees, and defective equipment in many industries, businesses, and institutions.

[29] In some cities in Canada, such as Ottawa, studies have shown that a major fraction of the water consumption in the summer can be attributed to outdoor water use, which essentially consists of the watering of lawns and gardens (e.g., *Adamowski*, 2008). It is thought that the water demand process in such situations is mainly driven by the maximum air temperature with the rainfall occurrences interrupting the process to cause transient drops in the water use. However, apart from the present study, this issue has not really been explored in Montreal.

[30] The primary consumers of Montreal's potable water are the industrial and commercial sectors, which account for 33% of total water usage [*City of Montreal*, 2010]. Domestic and municipal users account for approximately 30% of total use. In addition, it is estimated that the number of pipe breaks currently amounts to 500 per year, and approximately 40% of Montreal's water is wasted because of old leaky pipes [*Aubertin*, 2002]. Consumption is expected to continue to rise due to increasing urbanization and higher water demand, which will be exacerbated by the aging condition of the water infrastructure. The high water consumption rate has resulted in high operational costs in terms of water infrastructure maintenance and water treatment processes.

[31] To address the above concerns, the City of Montreal publicly committed to launching one of the largest pilot projects in its history in the spring of 2002, namely the "Effective Management of Water" throughout the city [*City of Montreal*, 2010]. In 2007, the Mayor of Montreal signed the framework agreement for water conservation of the Great Lakes and St. Lawrence Cities Initiative, a binational coalition to promote the protection of the Great Lakes and St. Lawrence River basin. Through this agreement, the city committed to reducing potable water production by 15% by 2015 compared with 2000 [*City of Montreal*, 2010]. To achieve the objective, the City of Montreal and related stakeholders have, among other initiatives, committed to reducing water consumption by sector of activity and establishing effective consumption targets, installing water meters in commercial buildings, and increasing public education among citizens about the importance of sustainable water use [*City of Montreal*, 2010]. A highly accurate short-term water demand forecasting model for the city of Montreal will help optimize the operational efficiency and management of the city water supply system. This study proposes the use of a highly accurate new method of daily urban water demand forecasting based on coupled wavelet-neural networks.

[32] Many variables influence water demand, most of which can be grouped into two classes: socioeconomic and climatic variables. Studies have demonstrated that socioeconomic variables are responsible for the long-term effects on water demand, while climatic variables are mainly responsible for short-term seasonal variations in water demand [*Miaou*, 1990]. Accordingly, in this study, MLR, MNLR, ARIMA, ANN, and WA-ANN models were developed using past daily urban water demands in addition to meteorological parameters. More specifically, the data used in this study consisted of daily total precipitation (mm), daily maximum temperature (°C), and daily total urban water demand (m^3), all during the summer period. The daily maximum temperature and total precipitation from May 2001 to August 2009 were obtained from the national climate data and information archive on the website of Environment Canada. The total daily water consumption data for Montreal from May 2001 to August 2009 was provided by the City of Montreal.

4. Model Development

4.1. MLR Models

[33] The MLR models for urban water demand forecasting were developed using the Microsoft Excel software program. The MLR models were trained and tested based on different combinations of the following input variables: the maximum temperature, the total precipitation, and the urban water demand. For each variable, data from the current day, from the previous day, from two days before, from three days before, and from four days before were explored. All MLR models were first trained using the data in the training set (May 2001 to May 2008), and then tested using the testing set (May 2008 to August 2009) and compared using the statistical measures of goodness of fit. It was determined that the best MLR model for one day ahead water demand forecasting is a function of the maximum temperature from the current day and the previous day, and the urban water demand from the current day and previous day.

4.2. Multiple Nonlinear Regression

[34] The MNLR models for urban water demand forecasting were developed using the Microsoft Excel XLSTAT software program [*XLSTAT*, 2011]. Different combinations of the following input variables were used to train and test the MNLR models: the maximum temperature, the total precipitation, and the urban water demand. For each variable, data from the current day, from the previous day, and from two, three, and four days before was explored. All MNLR models were first trained using the data in the training set (May 2001 to August 2007), and then tested using the testing set (May 2008 to August 2009) and compared using the statistical measures of goodness of

fit. It was determined that the best MNLR model for one day ahead water demand forecasting is a function of the following variables: maximum temperature (taken for the current day and for one, two and three days ago) and urban water demand (taken for the previous day, and for two and three days ago).

4.3. ARIMA Models

[35] In this study, the ARIMA models for urban water demand forecasting were developed using the NumXL software program [*Spider Financial*, 2011]. The NumXL is a Microsoft Excel add-in for time series analysis. The first step is to determine the stationarity of the input data series via the autocorrelation function (ACF). It was found that the City of Montreal urban water demand data series were not stationary. ARIMA models require the input data to have a constant mean, variance, and autocorrelation through time [*Box and Jenkins*, 1976]. Therefore, the input data series were transformed into a stationary model through a differencing process. The development of the ARIMA models in this study followed the methodology used by *Adamowski* [2008b]. The number of nonseasonal differences (d) was set to 1 to ensure the stationarity of the data. Following this, parameter estimation in the ARIMA models was performed, and then urban water demand was forecasted using the ARIMA models. All ARIMA models were first trained using the data in the training set (May 2001 to May 2008), and then tested using the testing set (May 2008 to August 2009), and compared using the statistical measures of goodness of fit. Among the various ARIMA models that were developed the model that had the best forecasting performance was a (2, 1, 3) ARIMA model.

4.4. ANN Models

[36] To develop an ANN model, the primary objective is to optimize the architecture of the ANN that captures the relationship between the input and output variables. The process of determining the number of neurons in the input and output layers is based on the input and output variables considered to model the physical process. The number of neurons in the hidden layer can be optimized using the available data through the use of a trial and error procedure. Other methods such as the one proposed by *Sudheer et al.* [2003] can also be used to determine the optimum number of neurons in the hidden layer.

[37] The regular ANN models (i.e., those not using wavelet decomposed input data) consisted of an input layer, one single hidden layer, and one output layer consisting of one node denoting the targeted daily water demand. Sigmoid and linear activation functions were used for the hidden and output nodes, respectively. The ANN models were trained and tested based on different combinations of input variables and the number of neurons in the model's hidden layer. The input nodes consisted of various combinations of the following variables: the maximum temperature, the total precipitation, and the urban water demand. For each variable, data from the current day, from the previous day, from two days before, from three days before, and from four days before was explored. Each ANN model was tested on a trial and error basis for the optimum number of neurons in the hidden layer (found to be between 2 and 5).

[38] All of the ANN models were first trained using the data in the training set (May 2001 to August 2007) to obtain the optimized set of connection strengths and then tested using the testing data set (May 2008 to August 2009) and compared using the statistical measures of goodness of fit.

4.5. WA-ANN Models

[39] The original data (daily urban water demand, daily maximum temperature, and daily total precipitation) was decomposed into series of approximation and details (DWs) using a modified version of the "a Trous" algorithm (so that future data values are not used in the calculation). To do this, the original time series were first decomposed into an approximation and accompanying detail signal. The decomposition process was then iterated with successive approximation signals being decomposed in turn, so that the original time series was broken down into many lower resolution components.

[40] In this study, three wavelet decomposition levels were selected. Each time subseries (DW) plays a distinct role in the original times series and has different effects on the original urban water demand series [*Wang and Ding*, 2003]. The correlation coefficients between each time subseries and the corresponding original data are presented in Table 1. The different correlation coefficients indicate the differing effect that each individual DW component has on the original series, and provide information to help determine the effective wavelet components on urban water demand. Selected DW components can then be added to each other to increase the performance of the ANN model [*Partal*, 2009; *Adamowski and Chan*, 2011]. From Table 1, the selected DW components are DW2 and DW3 for total precipitation, DW1, DW2, DW3 for maximum temperature, and DW1, DW2, DW3 for urban water demand. For each, the approximate series was also included. The above were used to form a new series for each variable. This dominant subseries selection process is the same as the procedure used by *Partal and Kisi* [2007], *Kisi* [2008], and *Kisi* [2009]. The DW selection process allowed most of the noisy data to be filtered out and facilitated the extraction of quasiperiodic and periodic signals in the original data time series.

[41] For the WA-ANN models, the ANN networks that were developed consisted of an input layer, a single hidden layer, and one output layer consisting of one node denoting the targeted water demand. Sigmoid and linear activation functions were used for the hidden and output nodes, respectively. The input nodes consisted of various combinations of the following variables: the new summed wavelet decomposed series of the maximum temperature, the total precipitation, and the urban water demand from the

Table 1. The Correlation Coefficients Between Each Discrete Wavelet Subseries and the Original Data for the City of Montreal

Discrete Wavelet Subseries	Total Precipitation Data	Maximum Temperature Data	Urban Water Demand Data
DW 1	0.024	0.232	0.260
DW 2	0.244	0.346	0.319
DW 3	0.275	0.405	0.284

current day, from the previous day, from two days before, from three days before, and from four days before. Each model was tested on a "trial and error" basis to determine the optimum number of neurons in the hidden layer based on different combinations of time subseries in the model's input layer and the number of neurons in the model's hidden layer [*Adamowski and Chan*, 2011]. The optimum number of neurons was found to be between 2 and 5.

[42] All of the WA-ANN models were first trained using the data in the training set (May 2001 to August 2007) to obtain the optimized set of connection strengths and then tested using the testing data set (May 2008 to August 2009), and compared using the statistical measures of goodness of fit.

Table 3. Variables of the Best Model for Each Forecasting Method[a]

Model	Variables
MLR	WD(t), WD(t-1), T(t), T(t-1)
MNLR	WD(t), WD(t-1), WD(t-2), WD(t-3), T(t), T(t-1), T(t-2), T(t-3)
ANN	WD(t), WD(t-1), WD(t-2), T(t), T(t-1)
WA-ANN	WD(t), WD(t-1), WD(t-2), WD(t-3), T(t), T(t-1)

[a]MLR, multiple linear regression; MLNR, multiple nonlinear regression; ARIMA, autoregressive integrated moving average; ANN, artificial neural networks; WA-ANN, coupled wavelet-neural networks; WD(t), water demand at time t; T(t), temperature at time t.

5. Results and Discussion

[43] The results obtained from the best model for each type of forecasting method are presented in Table 2, and the variables for the best model of each forecasting method are shown in Table 3. All the models were developed in the same way via an iterative procedure involving successively adding variables and keeping them if they improved the forecasting performance (while keeping the models as parsimonious as possible).

[44] The WA-ANN models were found to provide more accurate urban water demand forecasts than the MLR, MNLR, ARIMA, and ANN models for one day lead time forecasting for the City of Montreal, Canada. The best WA-ANN model was a function of the maximum temperature from the current day and the previous day, the water demand from the current day, from the previous day, from two days before, and from three days before. The best WA-ANN model had four neurons in the hidden layer. The best WA-ANN model, which had a testing coefficient of determination (R^2) of 0.919, and a testing Nash–Sutcliffe model efficiency coefficient (E) of 0.919, performed better than the best ANN model ($R^2 = 0.865$, $E = 0.864$), MNLR model ($R^2 = 0.838$, $E = 0.633$), MLR model ($R^2 = 0.786$, $E = 0.629$), and ARIMA model ($R^2 = 0.782$, $E = 0.778$). An R^2 and $E = 1$ correspond to a perfect match of forecasted data to the observed data in a model. Also, the best WA-ANN model, which had a testing root mean square error (*RMSE*) of 0.027, and a testing relative root mean square error (*RRMSE*) of 1.591%, was more accurate than the best ANN model ($RMSE = 0.035$, $RRMSE = 2.056\%$), MLNR model ($RMSE = 0.058$, $RRMSE = 3.323\%$), MLR model ($RMSE = 0.059$, $RRMSE = 3.491\%$), and ARIMA model ($RMSE = 0.045$, $RRMSE = 2.663\%$). The lower *RMSE* and *RRMSE* values in the best WA-ANN model indicate a smaller discrepancy relative to the forecasted value compared to other models.

Table 2. Comparison of the Best Models for Urban Water Demand Forecasting During the Training and Testing Period[a]

Performance Index	Best ANN Model	Best WA-ANN Model	Best MNLR Model	Best MLR Model	Best ARIMA Model
		Training Period			
R^2	0.792	**0.896**	0.848	0.760	0.758
E	0.789	**0.895**	0.848	0.761	0.762
RMSE	0.061	**0.043**	0.052	0.065	0.064
RRMSE (%)	3.333	**2.337**	2.797	3.532	3.525
		Testing Period			
R^2	0.865	**0.919**	0.838	0.786	0.782
E	0.864	**0.919**	0.633	0.629	0.778
RMSE	0.035	**0.027**	0.058	0.059	0.045
RRMSE (%)	2.056	**1.591**	3.323	3.491	2.663

[a]Note: R^2, coefficient of determination; *E*, Nash–Sutcliffe model efficiency coefficient; *RMSE*, root mean square error; *RRMSE*, relative root mean square error.

[45] Figures 2 to 6 compare the observed daily urban water demand in Montreal with the forecasted daily urban water demand for one day lead time during the testing period using the best WA-ANN, ANN, MNLR, MLR, and ARIMA models, respectively. It can be seen that the best ANN and ARIMA models slightly under-forecast during some high peak water demand periods, the best MNLR and MLR models over-forecast during most of the high peak water demand periods, while the best WA-ANN model provides the closest estimates to the corresponding observed daily water demand during most of the peak water demand periods.

[46] Figures 7 to 11 show the scatterplots comparing the observed and forecasted urban water demand for one day lead time forecasting during the testing period using the best WA-ANN, ANN, MLR, MNLR, and ARIMA models, respectively. Looking at the fit line equations (assuming that the equation is $y = a_0x + a_1$) of the scatterplots, the a_0 and a_1 coefficients for the WA-ANN model ($a_0 = 0.92$ and $a_1 = 0.14$) are closer to 1 and 0, respectively, than for the ANN model ($a_0 = 0.89$ and $a_1 = 0.19$), MNLR model ($a_0 = 0.85$ and $a_1 = 0.29$), MLR model ($a_0 = 0.80$ and $a_1 = 0.38$), and ARIMA model ($a_0 = 0.84$ and $a_1 = 0.28$). $a_0 = 1$ and $a_1 = 0$ represent the best fit line. The more the observed and forecasted data agree, the more the scatters concentrate in the vicinity of the identity line. The scatters fall on the identity line exactly when the observed and forecasted data sets are numerically equal. It can be seen that the WA-ANN model has less scattered estimates than other models.

[47] In this study it has been shown that the WA-ANN method performed better than the ANN, MLR, MNLR, and ARIMA methods in forecasting water demand. The WA-ANN models likely performed better than the ARIMA models and MLR models because the ARIMA technique did not include climatic variables during the modeling process, while the MLR method can only

Training Period

Performance Index	Best ANN model	Best WA-ANN model	Best MNLR model	Best MLR model	Best ARIMA model
R^2	0.792	**0.896**	0.848	0.760	0.758
E	0.789	**0.895**	0.848	0.761	0.762
RMSE	0.061	**0.043**	0.052	0.065	0.064
RRMSE (%)	3.333	**2.337**	2.797	3.532	3.525

Testing Period

Performance Index	Best ANN model	Best WA-ANN model	Best MNLR model	Best MLR model	Best ARIMA model
R^2	0.865	**0.919**	0.838	0.786	0.782
E	0.864	**0.919**	0.633	0.629	0.778
RMSE	0.035	**0.027**	0.058	0.059	0.045
RRMSE (%)	2.056	**1.591**	3.323	3.491	2.663

Figure 2. Comparison of forecasted versus observed water demand using the best WA-ANN model for one day ahead forecasting during the testing period (a) summer 2008 and (b) summer 2009.

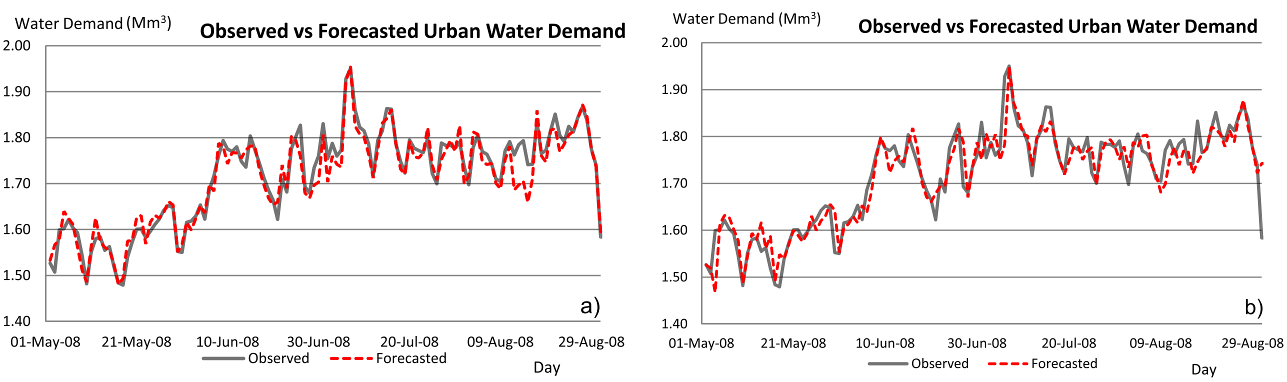

Figure 3. Comparison of forecasted versus observed urban water demand using the best ANN model for one day ahead forecasting during the testing period (a) summer 2008 and (b) summer 2009.

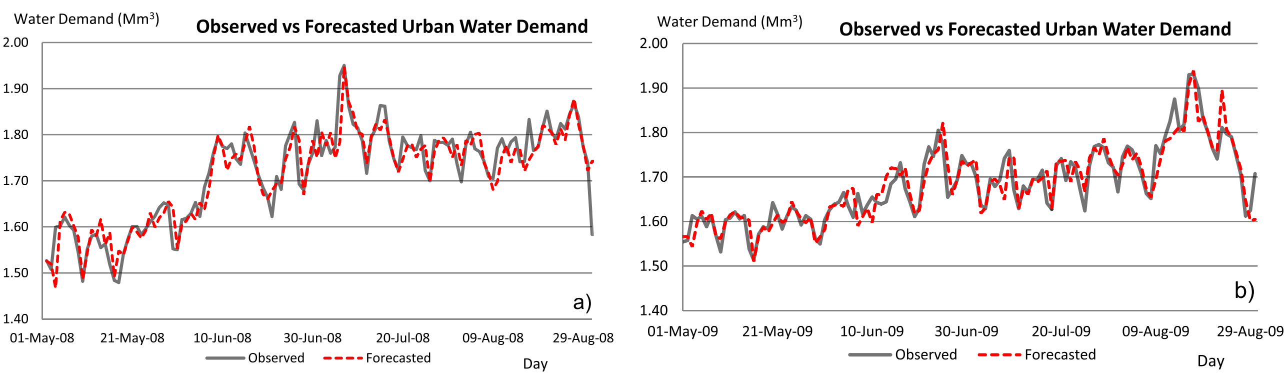

Figure 4. Comparison of forecasted versus observed water demand using the best MLR model for one day ahead forecasting during the testing period (a) summer 2008 and (b) summer 2009.

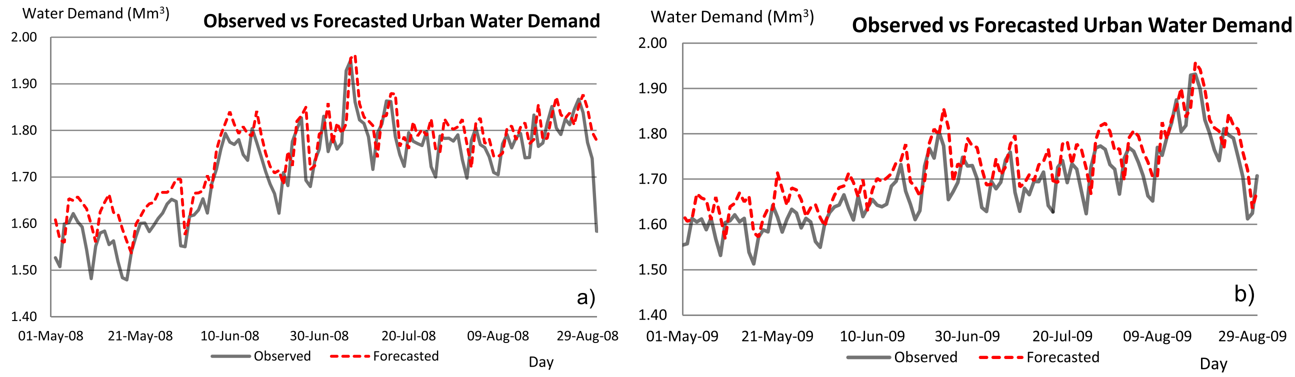

Figure 5. Comparison of forecasted versus observed water demand using the best MLNR model for one day ahead forecasting during the testing period (a) summer 2008 and (b) summer 2009.

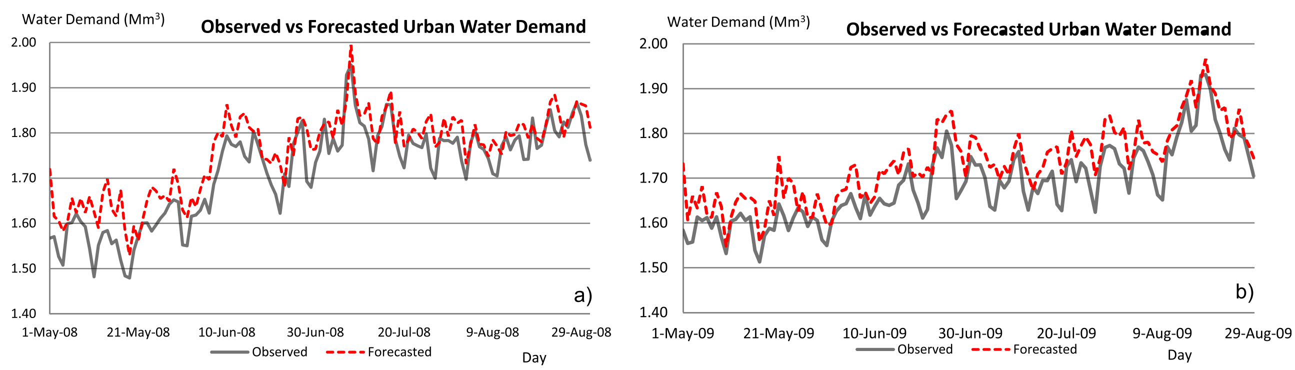

Figure 6. Comparison of forecasted versus observed water demand using the best ARIMA model for one day ahead forecasting during the testing period (a) summer 2008 and (b) summer 2009.

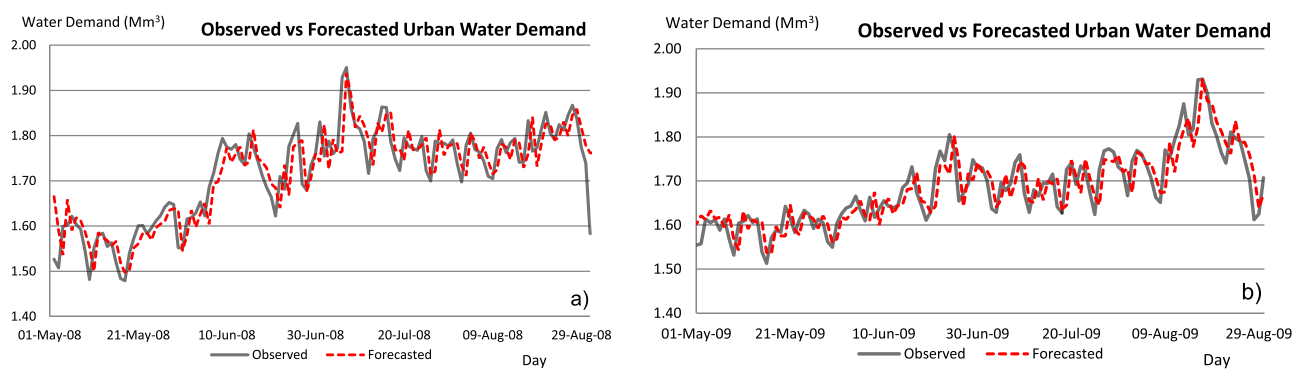

Figure 7. Scatterplot comparing observed and forecasted urban water demand using the best WA-ANN model for one day ahead forecasting during the testing period.

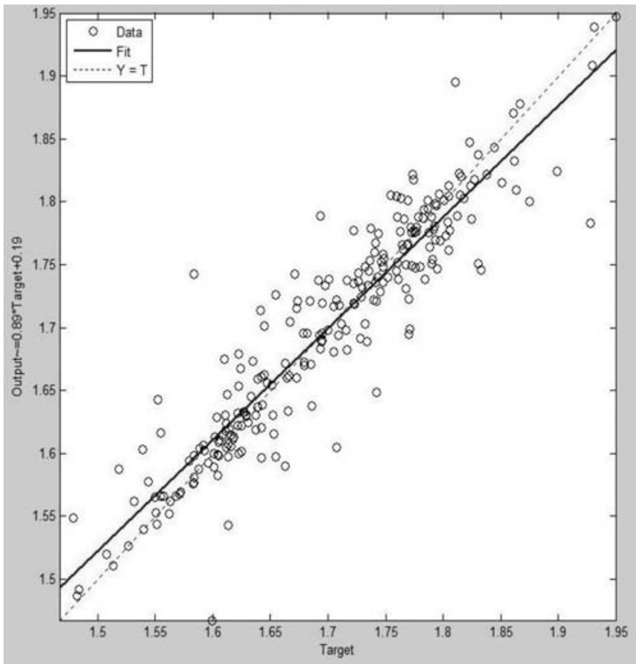

Figure 8. Scatterplot comparing observed and forecasted urban water demand using the best ANN model for one day ahead forecasting during the testing period.

capture relationships of a prespecified functional form, and as such it is not able to accurately predict nonlinear water demand. MNLR, on the contrary, is able to process nonlinear data and this likely resulted in more precise forecasts since the data that was used was nonlinear. However, the MNLR method does not explicitly address the nonstationarity inherent in the data. In addition, the

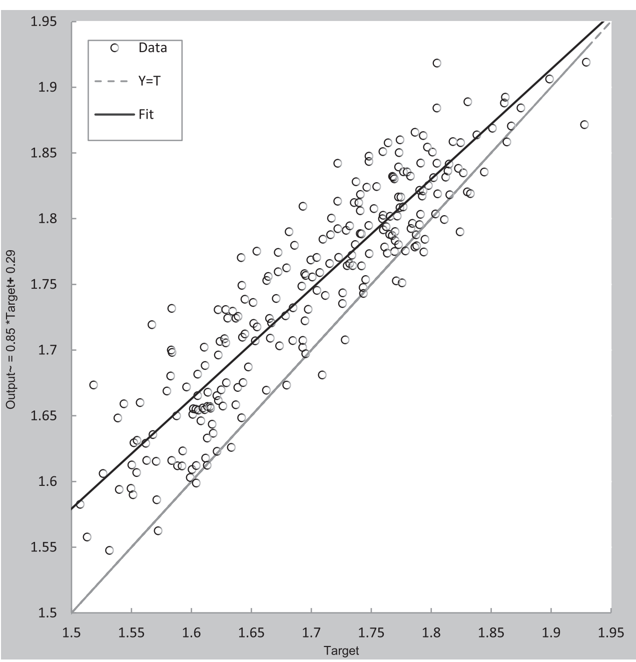

Figure 9. Scatterplot comparing observed and forecasted urban water demand using the best MNLR model for one day ahead forecasting during the testing period.

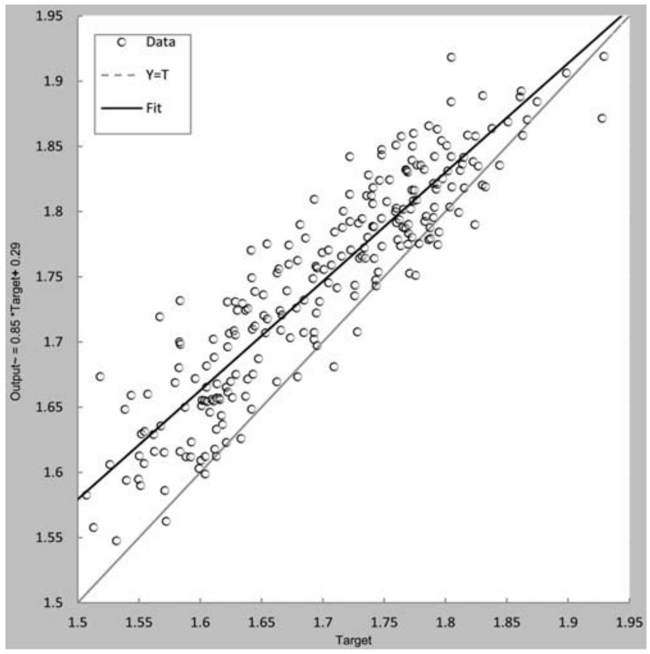

Figure 10. Scatterplot comparing observed and forecasted urban water demand using the best MLR model for one day ahead forecasting during the testing period.

WA-ANN models likely provided more accurate results than the ANN models because the wavelet transforms provide useful decompositions of the original time series, and the wavelet-transformed data improves the ANN forecasting performance by capturing detail information on various wavelet resolution levels. Through the use of wavelet analysis, a water demand series can be decomposed into

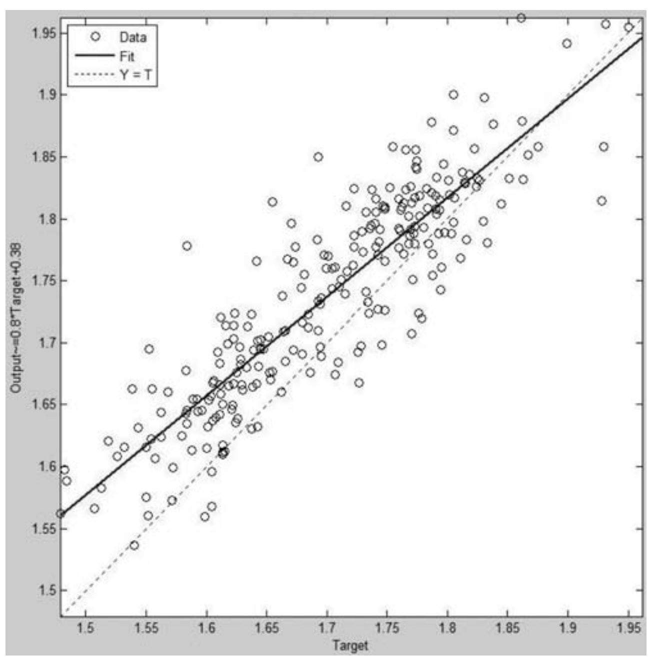

Figure 11. Scatterplot comparing observed and forecasted urban water demand using the best ARIMA model for one day ahead forecasting during the testing period.

a few selected component series that carry most of the information which can then be selectively used in forecasting. This allows most of the noisy data to be removed and it facilitates the extraction of quasiperiodic and periodic signals in the water demand time series.

[48] The accurate WA-ANN urban water demand forecasting method can help optimize a water supply system's operational efficiency by supplying the minimum amount of water necessary. The optimization of operations can also result in substantial savings (as high as 25%–30% of total operating costs) by reducing electricity costs and the use of chemicals in the treatment process [*Ghiassi et al.*, 2008]. In the long term, the proposed WA-ANN forecasting method for urban water demand management could be useful in the management, planning, and evaluation of existing water systems, water conservation initiatives, water pricing policies, and drought condition analyses.

[49] The results of this study indicate that the water demand process during the summer months in Montreal is mainly driven by the maximum air temperature (and less so by precipitation). However, this issue should be studied in greater detail in future studies. This study also showed that the WA-ANN method provides more accurate results than the traditional methods of ANN, ARIMA, MLR, and MNLR when used in short-term urban water demand forecasting. However, none of the forecasting models studied in this research account for uncertainty in the data. Failure to account for data uncertainty in models could lead to incorrect forecasts. The bootstrap method is a simple algorithm which can be used to account for uncertainty. The combination of WA-bootstrap-ANN has been successfully used in other areas of hydrological forecasting [*Tiwari and Chatterjee*, 2010a, 2010b], and the possible use of this method in water demand forecasting appears promising and will be studied further by the authors. In addition, there is little research available on the MNLR method in water demand forecasting. Although it was not the best method in this study, this method also showed promising results and should be studied further. And finally, the combination of wavelet transforms and SVMs have provided good forecasting results in hydrology [*Kisi and Cimen*, 2011]. In light of this, a comparative study between WA-ANN and WA-SVM applied to urban water demand forecasting would be useful.

6. Conclusions

[50] The potential of coupled wavelet-neural network models (WA-ANNs) for daily urban water demand forecasting during the summer months was investigated in this study. The study site was the City of Montreal, Canada. The coupled wavelet-neural network models were developed by combining two methods, namely discrete wavelet transforms and artificial neural networks. The WA-ANN models were compared to ANN, MLR, MNLR, and ARIMA models for urban water demand forecasting with a one day lead time. It was determined that the WA-ANN models provided more accurate results than the other models that were tested. Through the use of wavelet analysis, the water demand series was decomposed into a few selected component series that carried most of the information which was then selectively used in forecasting. This allowed most of the noisy data to be removed and it facilitated the extraction of quasiperiodic and periodic signals in the water demand time series.

[51] In reference to the original aims of this study, it was determined that the use of selected wavelet decomposed subseries as inputs to ANN models helps provide very accurate forecasts of daily urban water demand. The results of this study indicate that coupled wavelet-neural network models are a promising new method of short-term water demand forecasting that warrant further exploration. Possible future research stemming from this study include: exploring the application of coupled wavelet-neural network models for urban water demand forecasting for different lead times (weekly, monthly); analyzing holiday water consumption effects and other socioeconomic factors; comparing the use of different types of continuous (Morlet and Mexican Hat) and discrete (Daubechies) mother wavelets in the wavelet decomposition phase of the wavelet-neural network forecasting method; developing wavelet-bootstrap-neural network urban water demand forecasting models to provide ensemble forecasts; and comparing the wavelet-neural network method with other new methods such as support vector machines with localized multiple kernel learning.

Notation

ANN	artificial neural network
ARIMA	autoregressive integrated moving average
CWT	continuous wavelet transform
DF	degree of freedom
DW	decomposed wavelet subseries
DWT	discrete wavelet transform
E	Nash–Sutcliffe model efficiency coefficient
G	activation function
I_i	input value to node i of the input layer
MLR	multiple linear regression
MNLR	multiple nonlinear regression
N	number of data points used
O_k	output at node k of the output layer
R^2	coefficient of determination
RMSE	root mean square error
RRMSE	relative root mean square error
SEE	sum of squared errors
V_j	hidden value to node j of the hidden layer
WA	wavelet transform
WA-ANN	wavelet-neural network
s	scale parameter
$x(t)$	signal
$\bar{y}$	i mean value taken over N
y_i	observed peak weekly water demand
$\hat{y}_i$	forecasted peak weekly water demand
τ	translation parameter
*	complex conjugate
$\psi(t)$	mother wavelet

[52] **Acknowledgments.** Financial support for this study was provided by a FQRNT grant held by Jan Adamowski. Data and advice were provided by Michel Merette, the Director of the Drinking Water Production Department for the City of Montreal. His help is greatly appreciated.

References

Adamowski, J. (2007), Development of a short-term river flood forecasting method based on wavelet analysis, *J. Hydrol.*, *353*(3–4), 247–266, doi:10.1016/j.jhydrol.2008.02.013.

Adamowski, J. (2008a), Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis, *J. Hydrol.*, *353*, 247–266, doi:10.1016/j.jhydrol.2008.02.013.

Adamowski, J. (2008b), Peak daily water demand forecast modeling using artificial neural networks, *J. Water Resour. Plan. Manage.*, *134*(2), 119–128, doi:10.1061/(ASCE)0733-9496(2008)134:2(19).

Adamowski, J. (2008c), River flow forecasting using wavelet and cross-wavelet transform models, *J. Hydrol. Processes*, *22*, 4877–4891, doi:10.1002/hyp.7107.

Adamowski, J., and H. F. Chan (2011), A wavelet neural network conjunction model for groundwater level forecasting, *J. Hydrol.*, *407*, 28–40, doi:10.1016/j.jhydrol.2011.06.013.

Adamowski, J., and C. Karapataki (2008), Comparison of multivariate regression and artificial neural networks for peak urban water demand forecasting: the evaluation of different ANN learning algorithms, *J. Hydrol. Eng.*, *15*(10), 729–743, doi:10.1061/(ASCE)HE.1943-5584.0000245.

Adamowski, J., and K. Sun (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*, 85–90, doi:10.1016/j.jhydrol.2010.06.033.

Anderson, R., T. Miller, and M. Washburn (1980), Water savings from lawn watering restrictions during a drought year in Fort Collins, Colorado, *Water Resour. Bull.*, *16*(4), 642–645, doi:10.1111/j.1752-1688.1980.tb02443.

Asefa, T., M. Kemblowski, M. McKee, and A. Khalil (2006), Multi-time scale stream flow predictions: The support vector machines approach, *J. Hydrol.*, *318*(1–4), 7–16, doi:10.1016/j.jhydrol.2005.06.001.

Aubertin, L., A. Aubin, G. Pelletier, D. Curodeau, M. Osseyrane, and P. Lavallee (2002), Identifying and Prioritizing Infrastructure Rehabilitation, *North American Society for Trenchless Technology*, Liverpool, NY.

Banerjee, P., R. K. Prasad, and V. S. Singh (2009), Forecasting of groundwater level in hard rock region using artificial neural network, *Environ. Geol.*, *58*(6), 1239–1246, doi:10.1007/s00254-008-1619-z.

Box, G. E. P., and G. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, 2nd ed., San Francisco: Holden-Day.

Cannas, B., A. Fanni, L. See, and G. Sias (2006), Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning, *Phys. Chem. Earth*, *31*, 1164–1171, doi:10.1016/j.oce.2006.03.020.

Cassuto, A. E., and S. Ryan (1979), Effect of price on the residential demand for water within an agency, *Water Resour. Bull.*, *15*(2), 345–353, doi:10.1111/j.1752-1688.1979.tb00337.

Cimen, M., and O. Kisi (2009), Comparison of two different data driven techniques in modeling surface water level fluctuations of Lake Van, Turkey, *J. Hydrol.*, *378*(3–4), 253–262, doi:10.1016/j.jhydrol.2009.09.029.

City of Montreal (2010), *The Montreal Community Sustainable Development Plan 2010-2015*, Montreal, QC.

Cogger, K. O. (2010), Nonlinear regression methods: A survey and extensions, *Intell. Syst. Acc. Fin. Manage. 17*, 19–39, doi:10.1002/isaf.311.

Dziegielewski, B., and D. Baumann (1992), Tapping alternatives: The benefits of managing urban water demands, *Environment*, *34*(9), 6–11, 35–41.

Emiroglu, M., O. Bilhan, and O. Kisi (2011), Neural networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel, *Expert Syst. Appl.*, *38*(1), 867–874, doi:10.1016/j.eswa.2010.07.058.

Environment Canada (2010), Environmental Trends, *CESI*, *1*(3).

Ghiassi, M., D. K. Zimbra, and H. Saidane (2008), Urban water demand forecasting with a dynamic artificial neural network model, *J. Water Resour. Plan. Manage.*, *134*(2), 138–146, doi10.1061/(ASCE)0733-9496(2008)134:2(138).

Hagan, M. T., and M. Menhaj (1994), Training feedforward networks with the Marquardt algorithm, *IEEE Trans. Neural Networks*, *5*(6), 989–993, doi:10.1109/72.329697.

Han, D., L. Chan, and N. Zhu (2007), Flood forecasting using support vector machines, *Journal of Hydroinformatics*, *9*(4), 267–276, doi:10.2166/hydro.2007.027.

Haykin, S. (1998), *Neural Networks—A Comprehensive Foundation*, Prentice-Hall, Upper Saddle River, NJ.

Hughes, T. (1980), Peak period design standards for small western U.S. water supply, *Water Resour. Bull.*, *16*(4), 661–667, doi:10.1111/j.1752-1688.1980.tb02446.

Ivakhnenko, A. G. (1970), Heuristic self-organization in problems of engineering cybernetics, *Automatica*, *6*(2), 207–219, doi:10.1016/0005-1098(70)90092-0.

Jain, A., and L. Ormsbee (2002), Short-term water demand forecast modeling techniques: Conventional methods versus AI, *J. Am. Water Works Assoc.*, *94*(7), 64–72.

Karul, C., S. Soyupak, A. F. Cilesiz, N. Akbay, and E. Germen (2000), Case studies on the use of neural networks in eutrophication modeling, *Ecol. Model.*, *134*(2–3), 145–152, doi:10.1016/S0304-3800(00)00360-4.

Khan, M. S., and P. Coulibaly (2006), Application of support vector machine in lake water level prediction, *J. Hydrol. Eng.*, *11*(3), 199–205, doi:10.1061/(ASCE)1084-0699(2006)11:3(199).

Kim, T. W., and J. B. Valdes (2003), Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks, *J. Hydrol. Eng.*, *8*(6), 319–328, doi:10.1061/(ASCE)1084-0699(2003)8:6(319).

Kisi, O. (2004), River flow modeling using artificial neural networks, *J. Hydrol. Eng.*, *9*(1), 60–63, doi:10.1061/(ASCE)1084-0699(2004)9:1(60).

Kisi, O. (2008), Stream flow forecasting using neuro-wavelet technique, *Hydrol. Processes*, *22*, 4142–4152, doi:10.1002/hyp.7014.

Kisi, O. (2009), Neural networks and wavelet conjunction model for intermittent streamflow forecasting, *J. Hydrol. Eng.*, *14*(8), 773–782, doi:10.1061/(ASCE)HE.1943-5584.0000053.

Kisi, O. and M. Cimen (2011), A wavelet-support vector machine conjunction model for monthly streamflow forecasting, *J. Hydrol.*, *399*(1–2), 132–140, doi:10.1016/j.jhydrol.2010.12.041.

Lu, W., W. Wang, A. Y. T. Leung, S. Lo, R. K. K. Yuen, Z. Xuand H. Fan (2002), Air pollutant parameter forecasting using support vector machines, *Neural Networks Proceedings of the 2002 International Joint Conference*, *1*, 630–635, doi:10.1109/IJCNN.2002.1005545.

Maas, T. (2003), What the experts think: Understanding urban water demand management in Canada. Retrieved from http://www.polisproject.org/publications/byauthor/Maas,%20Tony.

Maidment, D., and S. Miaou (1986), Daily water use in nine cities, *Water Resour. Res.*, *22*(6), 845–851, doi:10.1029/WR022i006p00845.

Maidment, D., and E. Parzen (1984), Monthly water use and its relationship to climatic variables in Texas, *Water Resour. Bull.*, *19*(8), 409–418.

Maidment, D., S. Miaou, and M. Crawford (1985), Transfer function models of daily urban water use, *Water Resour. Res.*, *21*(4), 425–432, doi:10.1029/WR021i004p00425.

Maier, H., A. Jain, G. C. Dandy, and K. P. Sudheer (2010), Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, *Environ. Model. Software*, *25*, 891–909.

Maity, R., P. P. Bhagwat, and A. Bhatnagar (2010), Potential of support vector regression for prediction of monthly streamflow using endogenous property, *Hydrol. Processes*, *24*(7), 917–923, doi:10.1002/hyp.7535.

Mallat S. (1999), *A Wavelet Tour of Signal Processing*, 2nd ed.. Academic, New York.

Miaou, S. (1990), A class of time series urban water demand models with non-linear climatic effects, *Water Resour. Res.*, *26*, 169–178.

Minu, K. K., M. C. Lineesh, and C. Jessy John (2010), Wavelet neural networks for nonlinear time series analysis, *Appl. Math. Sci.*, *4*(50), 2485–2495.

Miyagishi, K., M. Ohsako, and H. Ichihashi (1999), Temperature prediction from regional spectral model by neurofuzzy GMDH, *Proc. of The Second Asia-Pacific Conference on Industrial Engineering and Management Systems, Kanazawa, Japan*, 705–708.

Msiza, I. S., F. V. Nelwamondo, and T. Marwala (2008), Water demand prediction using artificial neural networks and support vector regression, *J. Comput.*, *3*(11), 1–8, doi:10.4304/jcp.3.11.1-8.

Ozbek, F. S., and H. Fidan (2009), Estimation of pesticides usage in the agricultural sector in Turkey using artificial neural network (ANN), *J. Animal Plant Sci.*, *4*(3), 373–378.

Partal, T. (2009), River flow forecasting using different artificial neural network algorithms and wavelet transform, *Can. J. Civil Eng.*, *36*(1), 26–38, doi:10.1139/L08-090.

Partal, T., and O. Kisi (2007), Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, *J. Hydrol.*, *342*(1–2), 199–212, doi:10.1016/j.jhydrol.2007.05.026.

Pedhazur, E. J. (1982), *Multiple Regression in Behavioral Research: Explanation and Prediction*, Holt, Rinehart and Winston, New York.

Pramanik, N., and R. K. Panda (2009), Application of neural network and adaptive neuro fuzzy inference systems for stream flow prediction, *Hydrol. Sci. J.*, *54*(2), 247–260, doi:10.1623/hysj.54.2.247.

Rajasekaran, S., S. Gayathri, and T.-L. Lee (2008), Support vector regression methodology for storm surge predictions, *Ocean Eng.*, *35*, 1578–1587, doi:10.1016/j.oceaneng.2008.08.004.

Rubin, D. (2005), *Down the drain: Dealing with Montreal's water wastage*, McGill Reporter, *37*(17), Montreal, QC.

Sahoo, G. B., and C. Ray (2006), Flow forecasting for a Hawaii stream using rating curves and neural networks, *J. Hydrol.*, *317*, 63–80, doi:10.1016/j.jhydrol.2005.05.008.

Sarycheva, L. (2003), Using GMDH in ecological and socio-economical monitoring problems, *Syst. Anal. Model. Simul.*, *43*(10), 1409–1414, doi:10.1080/02329290290024925.

Sethi, R. R., A. Kumar, S. P. Sharma, and H. C. Verma (2010), Prediction of water table depth in a hard rock basin by using artificial neural network, *Int. J. Water Resour. Environ. Eng.*, *2*(4), 95–102.

Shiri, J., and O. Kisi (2010), Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model, *J. Hydrol.*, *394*(3–4), 486–493.

Smith, J. (1988), A model of daily municipal water use for short-term forecasting, *Water Resour. Res.*, *24*(2), 201–206, doi:10.1029/WR024i002p00201.

Solomatine, D., and A. Ostfeld (2008), Data driven modeling: Some past experiences and new approaches, *J. Hydroinf.*, *10*(1), 3–22.

Spider Financial (2011), *NumXL Microsoft Excel add-in*, Spider Financial, Chicago.

Sreekanth, P. D., N. Geethanjali, P. D. Sreedevi, S. Ahmed, N. R. Kumar, and P. D. K. Jayanthi (2009), Forecasting groundwater level using artificial neural networks, *Curr. Sci.*, *96*(7), 933–939.

Sudheer, K. P., P. C. Nayak, and K. S. Ramasastri (2003), Improving peak flow estimates in artificial neural network river flow models, *Hydrol. Processes*, *17*(3), 677–686, doi:10.1002/hyp.5103.

Tiwari, M. K., and C. Chatterjee (2010a), A new wavelet-bootstrap-ANN hybrid model for daily discharge forecasting, *J. Hydroinf.*, *13*(3), 500–519, doi:10.2166/hydro.2010.142.

Tiwari, M. K., and C. Chatterjee (2010b), Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach, *J. Hydrol.*, *394*(3–4), 458–470, doi:10.1016/j.jhydrol.2010.10.001.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*, Springer, New York.

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

Wang, W. C., K. W. Chau, C. T. Cheng, and L. Qiu (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series, *J. Hydrol. 374*(3–4), 294–306.

Wong, J. S., Q. Zhang, and Y. D. Chen (2010), Statistical modeling of daily urban water consumption in Hong Kong: Trend, changing patterns, and forecast, *Water Resour. Res.*, *46*, W03506, doi:10.1029/2009WR008147.

XLSTAT (2011), *XLSTAT Microsoft Excel add-in*, Addinsoft SARL, Paris.

Yu, P.-S., S.-T. Chen, and I.-F. Chang (2006), Support vector regression for real-time flood stage forecasting, *J. Hydrol.*, *328*(3–4), 704–716, doi:10.1016/j.jhydrol.2006.01.021.

Zhou, S., T. McMahon, A. Walton, and J. Lewis (2000), Forecasting daily urban water demand: A case study of Melbourne, *J. Hydrol.*, *236*(3), 153–164, doi:10.1016/S0022-1694(00)00287-0.

J. Adamowski, H. Fung Chan, S. O. Prasher, and A. Sliusarieva, Department of Bioresource Engineering, McGill University, 21 111 Lakeshore Road, Ste. Anne de Bellevue, QC H9X 3V9, Canada. (jan.adamowski@mcgill.ca)

B. Ozga-Zielinski, Centre of Hydrology, National Research Institute, Institute of Meteorology and Water Management, ul. Podlesna 61, Warsaw PL-01-673, Poland.