DOI: 10.1515/jwld-2016-0004

JOURNAL OF WATER AND LAND DEVELOPMENT
2016, No. 28 (I–III): 37–52
PL ISSN 1429–7426
Available (PDF): http://www.itp.edu.pl/wydawnictwo/journal; http://www.degruyter.com/view/j/jwld

Received 04.11.2015
Reviewed 05.01.2016
Accepted 27.01.2016

A – study design
B – data collection
C – statistical analysis
D – data interpretation
E – manuscript preparation
F – literature search

Water demand forecasting using extreme learning machines

Mukesh TIWARI[1) ABCDEF], Jan ADAMOWSKI[2) ABCDEF], Kazimierz ADAMOWSKI[3) A]

[1)] Anand Agricultural University, Department of Soil and Water Engineering, College of Agricultural and Technology, Gujarat, India; e-mail: tiwari.iitkgp@gmail.com
[2)] McGill University, Faculty of Agricultural and Environmental Sciences, Department of Bioresource Engineering, Quebec, Canada, H9X 3V9; e-mail: jan.adamowski@mcgill.ca
[3)] University of Ottawa, Department of Civil Engineering, Canada; e-mail: kadamowski@hotmail.com

For citation: Tiwari M., Adamowski J., Adamowski K. 2016. Water demand forecasting using extreme learning machines. Journal of Water and Land Development. No. 28 p. 37–52.

Abstract

The capacity of recently-developed extreme learning machine (ELM) modelling approaches in forecasting daily urban water demand from limited data, alone or in concert with wavelet analysis (W) or bootstrap (B) methods (*i.e.*, ELM, ELM_W, ELM_B), was assessed, and compared to that of equivalent traditional artificial neural network-based models (*i.e.*, ANN, ANN_W, ANN_B). The urban water demand forecasting models were developed using 3-year water demand and climate datasets for the city of Calgary, Alberta, Canada. While the hybrid ELM_B and ANN_B models provided satisfactory 1-day lead-time forecasts of similar accuracy, the ANN_W and ELM_W models provided greater accuracy, with the ELM_W model outperforming the ANN_W model. Significant improvement in peak urban water demand prediction was only achieved with the ELM_W model. The superiority of the ELM_W model over both the ANN_W or ANN_B models demonstrated the significant role of wavelet transformation in improving the overall performance of the urban water demand model.

Key words: *artificial neural networks, bootstrap, Canada, extreme learning machines, uncertainty, water demand forecasting, wavelets*

INTRODUCTION

Attributable to demographic expansion and industrial development, the rapid rise in worldwide urban water consumption has placed potable water distribution systems under stress. Given the advent of climate change, these problems will likely become more acute in the future [SAADAT *et al.* 2011; ARAGHI *et al.* 2015]. Accurate forecasting of short-term water demand can contribute to the efficient operation and management of urban water supply systems, resulting in demand being met efficiently and sustainably [ADAMOWSKI *et al.* 2013; CAMPISI-PINTO *et al.* 2012; TIWARI, ADAMOWSKI 2014]. The estimation of future urban water demand is therefore essential to the sustainable planning of regional water-supply systems [ADAMOWSKI *et al.* 2013; TIWARI, ADAMOWSKI 2015a, b; ZHOU *et al.* 2002]. Given increases in the diverse components of urban water demand (*e.g.*, residential, public, industrial and commercial use – HANEMANN [1998]), water stress and scarcity have become critical issues [ADAMOWSKI *et al.* 2012a, b, c; 2013; DAVIS, KIEFER 2005; GOYAL *et al.* 2014; HAIDARY *et al.* 2013; KAYAGA, SMOUT 2008]. Furthermore, multiple-scale interactions between individuals and natural systems create a further range of urban water demand management challenges [HOUSE-PETERS, CHANG 2011]. Short-term urban water de-

mand forecasts play a significant role in the optimal operation of pumps, wells, and reservoirs, as well as in informing decisions regarding balanced water resource allocation in the face of urgent water needs [HERRERA *et al.* 2010; JAIN, ORMSBEE 2002; KAME'ENUI 2003]. Urban water is generally allocated according to the experience of operators and average water demand; however, accurate and reliable forecasts of short-term demand can help operators provide water in a more efficient and sustainable manner [ZHOU *et al.* 2002].

This study focuses on fast, efficient methods for short-term (1-day lead time) urban water demand forecasts, in an effort to achieve accurate and reliable forecasts of water demand for the City of Calgary, Alberta, Canada. Traditionally, linear regression, trend-extrapolation, and time-series techniques have been applied in forecasting water resources operations variables, particularly in the domain of urban water demand [ADAMOWSKI *et al.* 2012a; BELAYNEH *et al.* 2014]. Short-term water demand data generally exhibits nonlinear and nonstationary behaviour [GHIASSI *et al.* 2008] at multiple spatial and temporal scales [HOUSE-PETERS, CHANG 2011]. Non-stationarity, such as that attributable to seasonal variations and trends, significantly lowers modelling accuracy for time series, generally leading to poor predictions in operational applications [ADAMOWSKI *et al.* 2009; 2010; FRANCESCO, BERND 2000; NALLEY *et al.* 2012; 2013; PINGALE *et al.* 2014; RATHINASAMY *et al.* 2013; 2015]. Wavelet transformation, a time–frequency representation of a signal present at many different intervals in the time domain, provides considerable information about the physical structure of the time series data [BELAYNEH *et al.* 2014; DAUBECHIES 1990; KARRAN *et al.* 2014; NOURANI *et al.* 2014]. Wavelet analysis uses a mother wavelet function to decompose non-stationary data into multiple scale-specific time series [NALLEY *et al.* 2012; 2013] and thereby helps to distinguish among daily, weekly, and seasonal cycles inherent in water demand.

Wavelet-transformation-based artificial neural networks (ANN_W) have been found to be more accurate than multiple linear regression, time-series or regular artificial neural network (ANN) models in forecasting regional drought [KIM, VALDES 2003], rainfall-runoff [ANCTIL, TAPE 2004; NOURANI *et al.* 2009], monthly and daily streamflow [ADAMOWSKI 2007; ADAMOWSKI, SUN 2010; KIŞI 2008; 2009; MAHESWARAN, KHOSA 2012; NAYAK *et al.* 2013; TIWARI *et al.* 2013], monthly groundwater levels [ADAMOWSKI, CHAN 2011], and short-term urban water demand [ADAMOWSKI *et al.* 2012a, b, c; TIWARI, ADAMOWSKI 2013]. Other hybrid methods proposed in the hydrological forecasting literature are the bootstrap-based ANN (ANN_B) [TIWARI, CHATTERJEE 2010a, b], fuzzy neural networks [ALVISI, FRANCHINI 2011], and grey neural networks [ALVISI, FRANCHINI 2012]. Besides reducing uncertainty in the variance by mimicking randomness [EFRON, TIBSHIRANI 1993], ANN_B models are simpler and easier to use in addressing uncertainty in an operational setting compared to Bayesian approaches [ISUKAPALLI, GEORGOPOULOS 2001]. Several studies have shown ANN_B models to outperform standard ANN models [ABRAHART 2003; HAN *et al.* 2007; JEONG, KIM 2005; JIA, CULVER 2006; SHARMA, TIWARI 2009; SRIVASTAV *et al.* 2007; TIWARI, CHATTERJEE, 2010a]. Both ANN_W and ANN_B hybrid approaches can be combined to form a wavelet-bootstrap-ANN (ANN_{WB}) model with the potential ability to achieve greater accuracy and reliability in real time water demand forecasting. However, this has not been undertaken to date.

In the present study a wavelet-extreme learning machine (ELM_W) [HUANG *et al*. 2006; 2015] based water demand model, developed with limited data [limited years of dataset), is proposed. The ELM is a fast three-step model designed to use a Single Layer Feedforward Neural Network with hidden neurons and randomly chosen weights. The hidden layer learns patterns from distinct observations and therefore requires no parameter tuning, only a predefined network. The ELM is free from the complications faced by gradient-based algorithms (*e.g*., learning rate, learning epochs and local minima) [ACHARYA *et al*. 2014; BELAYNEH, ADAMOWSKI 2014; ŞAHIN *et al*. 2014]. Despite their widespread use, ANNs suffer from difficulty in training predictors and may not, therefore, produce a unique solution over various runs due to different weights [COULIBALY, EVORA 2007; KHAN, COULIBALY 2006].

As a result, the present study sought to explore, for the first time, the use of ELM, ELM_B and ELM_W water demand forecasting models for short-term urban water demand forecasting for the city of Calgary (Alberta Canada), and compare their performance to that of previously applied ANN, ANN_B, and ANN_W models [TIWARI, ADAMOWSKI 2015a, b]. As only three years of urban water demand data were available for calibration and validation of the models, a secondary aim of this study was also to explore how these methods fared in situations with limited data. The input variables applied in this study consisted of average daily water demand, maximum temperature and total precipitation.

THEORETICAL OVERVIEW

EXTREME LEARNING MACHINE

Owing to its prior application in hydrology [ACHARYA *et al*. 2014; DEO, ŞAHIN 2015a], the present study has extended the application of ELM algorithm-based models [HUANG *et al*. 2006] to forecasting daily urban water demand (UWD). Based on state-of-the-art single-layer feed-forward network algorithms, ELMs are similar to feed-forward back-propagation ANNs (ANN_{FFBP}) and least square support vector regression (LSSVR). However, compared to the latter algorithms, ELMs have greater ability to

solve regression problems efficiently in a short modelling time [HUANG *et al.* 2012] and show a relatively better predictive performance [ACHARYA *et al*. 2014; DEO, ŞAHIN 2015a, b]. Moreover, in ELMs the weights and hidden neuron selections are randomized, so that the output weights have a unique least-square solution solved by way of the Moore-Penrose generalized inverse matrix method [HUANG *et al.* 2006]. Consequently, the ELM is a simple, three-step procedure requiring no parameterization except the randomised determination of hidden neurons. Optimisation is performed by choosing activation functions for hidden nodes based on sigmoid, radial basis or hard limit equations [DEO, ŞAHIN 2015a, b; ŞAHIN 2012; ŞAHIN *et al*. 2013; 2014]. This yields distinct advantages over conventional models.

Figure 1(a) illustrates the general ELM modelling framework. Mathematically, for a set of predictive samples (X_i, t_i) where i = 1, 2, …, N with inputs, $X_i = [x_{i1}, x_{i2}, \ldots, x_{im}] \in \mathbf{R}^m$ and $t_i \in$ R, the ELM architecture consists of L random hidden neurons with activation function, $g(x)$ such that [HUANG *et al*. 2006]:

$$\sum_{z=1}^{z=L} \beta_z \, g_z(X_i) = \sum_{z=1}^{z=L} \beta_z \, g(a_z X_i + c_z) = O_i \quad (1)$$

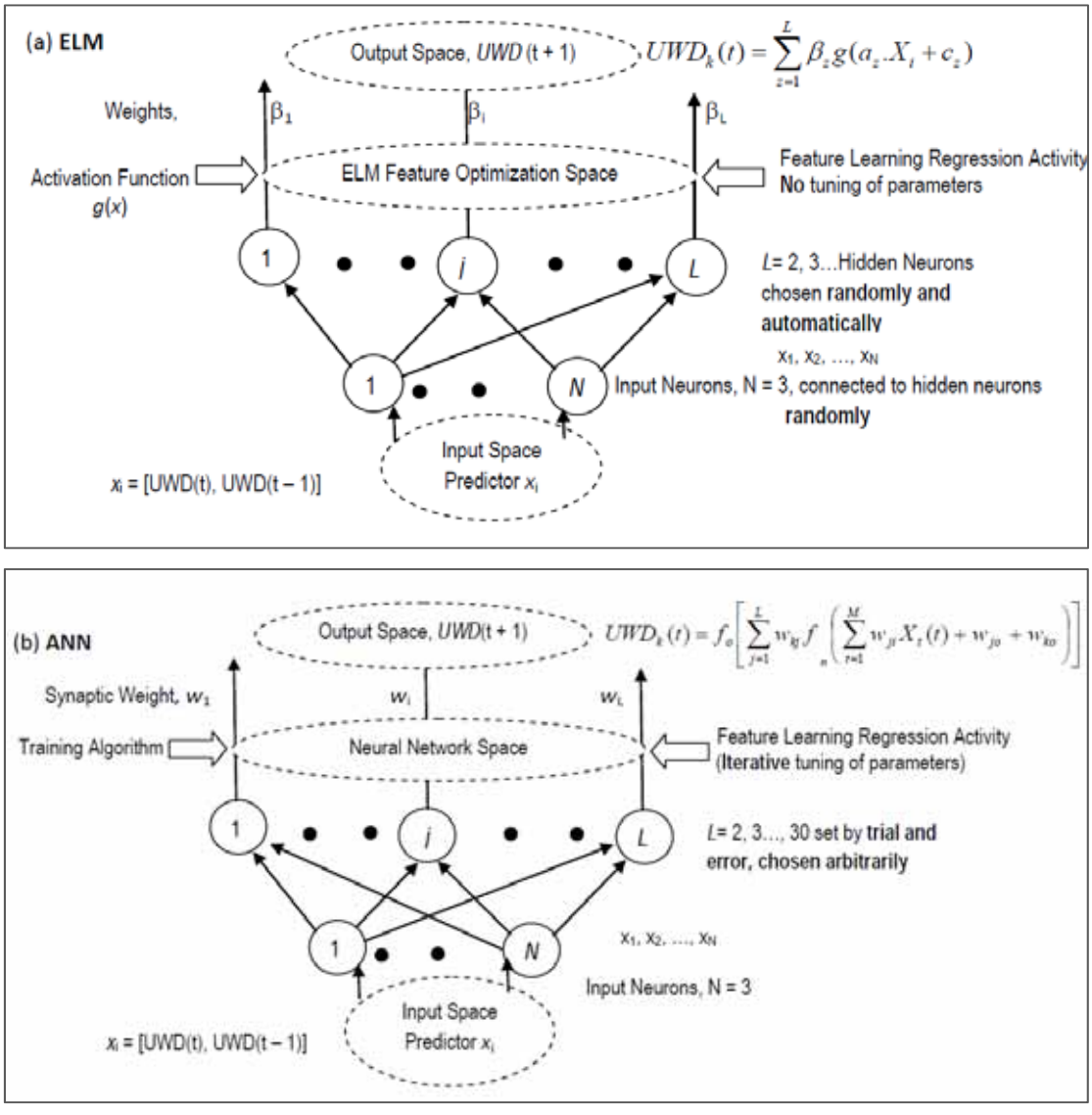

Fig. 1. An illustration of (a) extreme learning machine (ELM) and (b) artificial neural network (ANN) modelling frameworks used for prediction of the UWD; source: own study

Where $a_z = [a_{z1}, a_{z2}, \cdots, a_{zm}]^T$ is the weight vector connecting the z^{th} hidden neuron, the input neuron, $\beta_z = [\beta_z]^T$ is the weight vector connecting the z^{th} hidden neuron and the output neuron, O_i (which is the same dimension as the target function, T_i), c_z is the threshold of the z^{th} hidden neuron, and $a_z X_i$ denotes the inner product. The output neurons, which are considered to be linear in this study, do not require any transformative equation.It has been proven by HUANG *et al.* [2006] that Single-hidden-layer feed-forward ANNs with L hidden nodes and an activation function $g(X)$ have the proven capacity to approximate N training pair samples with zero error, $\sum_{i=1}^{N} \|O_i - T_i\| = 0$, Thus there must exist β_z, a_z and b_z such that:

$$\sum_{z=1}^{z=L} \beta_z \, g(a_z X_i + c_z) = T_i \tag{2}$$

Which can be written in a compact form as $\mathbf{H}\beta = \mathbf{T}$,

$$H(a_1,...a_L, C_1,...C_L, X_1,...,X_L) = \begin{bmatrix} (a_1 X_1 + c_1) & \cdots & (a_L X_1 + c_L) \\ \vdots & \cdots & \vdots \\ (a_1 X_N + c_1) & \cdots & (a_L X_N + c_L) \end{bmatrix}_{N \times L} \tag{3}$$

and

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times 1} \text{ and } T = \begin{bmatrix} T_1^T \\ \vdots \\ T_L^T \end{bmatrix}_{N \times 1} \tag{4}$$

It is noteworthy that the input weights and hidden neuron biases are randomly generated in the ELM model, which is different from the ANN that requires iterative tuning of parameters, and thus, requires greater modelling time [DEO, ŞAHIN 2015a, b; ŞAHIN *et al*. 2013; 2014]. The training algorithm is used to find least squares solutions to the system of equations $\mathbf{H}\beta = \mathbf{T}$ and the parameter β can directly be determined as

$$\hat{\beta} = H^{\dagger} T \tag{5}$$

where the $\hat{\beta}$ is the smallest least-square linear system solved using the $H^{\dagger}$ as the Moore-Penrose generalized inverse of H.

In order to develop an ELM model using a set of predictive samples (X_i), the forecasts of the *UWD*, $UWD_k(t)$ are given by:

$$UWD_k(t) = \sum_{z=1}^{L} \beta_z g(a_z X_i + c_z) \tag{6}$$

In the present study, the time-series forecasts denoted by $UWD_k(t)$ were generated using activation functions $g(X)$ described by the logarithmic sigmoid, $\psi(X)$ and the output function, $\chi(X)$ equations [VOGL *et al*. 1988] as seen in equations (7) and (8).

$$\psi(X) = \frac{1}{1 + e^{-X}} \tag{7}$$

$$\chi(X) = linear(X) \tag{8}$$

ARTIFICIAL NEURAL NETWORK

A well-established class of nonlinear modelling techniques mimicking the biological functions of the human brain [MCCULLOCH, PITTS 1943], ANN models served as a benchmark for ELM model performance in predicting UWD in this study. Basically, an ANN represents a highly interconnected framework that sends information from an input to an output layer through weighted connections and functional neurons to facilitate nonlinear mapping of the predictive dataset to high-dimensional hyper-planes, as demonstrated in Fig. 1(b). This allows the separation of data patterns, formation of idealised models and subsequent UWD predictions.

Widely applied in hydrology, the popular ANN_{FFBP} class of ANN models, equipped with multilayer perceptron functional neurons [ABBOT, MAROHASY 2012; ADAMOWSKI *et al*. 2012c; DEO, ŞAHIN 2015a; KESKIN, TERZI 2006; MEKANIK *et al*. 2013] was used in the present study. The ANN architecture is designed to successively update the model parameters (weighted connections and neuron biases) to drive the empirical error to a set tolerance through each iteration (epochs) of forward passing of updated parameters and backward propagation of the errors to tune them.

For a set of predictive (input) sample (X_i, t_i) where i = 1, 2, …, N denotes the sequence of inputs, $X_i = [x_{i1}, x_{i2}, \ldots, x_{im}] \in \mathbf{R}^m$ and $t_i \in$ R, the FFBP-ANN model formulated is written as [KIM, VALDÉS 2003]:

$$UWD_k(t) = f_o\left[\sum_{j=1}^{L} w_{kj} f_n\left(\sum_{t=1}^{M} w_{ji} X_t(t) + w_{jo} + w_{ko}\right)\right] \tag{9}$$

where, L (determined iteratively rather than randomly as with the ELM model) is the number of hidden neurons, $X_i(t)$ is the i^{th} input variable at the time-step, w_{ji} is the weight that connects the i^{th} neuron in input layer and the j^{th} neuron in the hidden layer and w_{jo} is the bias for the hidden j^{th} hidden neuron.

In literature, second-order training methods with Levenberg-Marquardt (LM) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton backpropagation algorithms are used [DENNIS, SCHNABEL 1996; MARQUARDT 1963]. The algorithm is used to minimize the mean squared error of the predicted and observed UWD [TIWARI, ADAMOWSKI 2013]. In our study, an LM algorithm that uses an approximation to the Hessian matrix was used as follows [DEO, ŞAHIN 2015a]:

$$x_{k+1} = x_k - \left[J^T J + \mu I\right]^{-1} J^T e \tag{10}$$

where **J** is the Jacobian matrix calculated using standard backpropagation techniques and is less complex than computing the Hessian matrix [MARQUARDT 1963]. The **J** contains first derivatives of network errors with respect to the weights and biases, **e** is a vector of errors, μ is the combination coefficient and ***I*** is the identity matrix.

DISCRETE WAVELET TRANSFORMATION

A primary purpose of this study was to demonstrate the effectiveness of ELM_W models for urban water demand forecasting. In general, wavelet de-

composition is a multi-resolution tool for pre-processing of non-stationary signals. This is similar to short-time Fourier transformation as a windowing technique in which the time-series are decomposed into the shifted and scaled versions of a wavelet, termed the mother wavelet. These can serve in extracting frequency-based information from current time-series that can then be used to predict future time-series. Assuming a continuous time-series $X(t)$ as an input vector where $t \in [\infty, -\infty]$, a wavelet function $\psi(\eta)$ that depends on a non-dimensional time parameter η is defined as:

$$\psi(\eta) = \Psi(\tau, s) = s^{-\frac{1}{2}} \psi\left(\frac{t-\tau}{s}\right) \tag{11}$$

where t = time, τ = time step in which the window function is iterated and $s \in [0, \infty]$ is the wavelet scale. The term $\psi(\eta)$ must have zero mean and localized in both the time and Fourier space.

The discrete wavelet transformation (DWT) selects translation and location parameters for the input signal. Subsequently, discrete wavelets coefficients (DWCs) are obtained that represent the minimum number of components needed to reflect the time-series according to the mother wavelet. Several families of wavelets have proven useful in a range of applications [MALLAT 1989]. For practical applications, hydrologists use wavelets to analyse a discrete rather than a continuous signal. This discrete wavelet is of the form:

$$g_{i,j}(t) = \frac{1}{\sqrt{a_o^i}} \left(\frac{t - j b_o a_o^l}{a_o^i}\right) \tag{12}$$

where i and j are the integer values, and b_0 and a_0 are the location parameter and the specified fined dilation step, respectively. Common values for a_0 and b_0 are 2 and 1, respectively [SEHGAL *et al.* 2014; TIWARI, ADAMOWSKI 2013]. The discrete wavelet transform involves selecting scales and positions based on powers of two, (called the dyadic scales and translations). The dyadic wavelet can be compressed as follows:

$$g_{i,j} = 2^{-i/2} g(2^{-i} h - j) \tag{13}$$

$$T_{i,j} = 2^{-i/2} \sum_{h=0}^{j-1} g(2^{-i} h - j) x_h \tag{14}$$

where $T_{(i, j)}$ is the wavelet coefficient for the discrete wavelet of scale $a=2^i$ and the location $b = 2^i j$. Equation (14) considers a finite time series, x_h, h = 0, 1, 2, …, j–1; and j is an integer power of 2, i.e., $j = 2^i$.

The inverse discrete transform is given by:

$$x_h = \overline{T} + \sum_{i=1}^{l} \sum_{j=0}^{2^{i-j}} T_{i,j} 2^{-i/2} g(2^{-i} h - j) = \overline{T} + \sum_{i=1}^{l} W_i(t) \tag{15}$$

where $\overline{T}$ is called the approximation sub-signal at level i, and $W_i(t)$ represents the details of the sub-signals at level i = 1, 2, …, l.

The wavelet coefficients, $W_i(t)$ (i = 1, 2, …, l) provide the details of the signal, which can capture small features of interpretational values in the inputted dataset. The residual term, $\overline{T}$, represents the background information of the data. The wavelet is robust since it does not include any potentially erroneous assumption or parametric testing procedure. Because of the simplicity of $W_1(t)$, $W_2(t)$ …, $W_l(t)$, the relevant characteristics in the hydrologic dataset (e.g. periods, hidden periods, dependence and jumps) can be diagnosed through these discrete wavelet components [TIWARI, CHATTERJEE 2011]. Consequently, the prediction accuracy of drought models are improved.

BOOTSTRAP TECHNIQUE

There are three sources of uncertainty that affect the output of the ANN and ANN_W models: parameter uncertainty, sub-optimal training and insufficient input variables. Bootstrapping is a computational, data-driven simulation method that can be used to assess uncertainty by measuring the variance σ_s^2 of S, the bootstrap resample. Bootstrap samples are generated through an intensive resampling with replacement method. These samples or realizations provide a better understanding of the mean and variability of the original data, and thus of its unknown distribution or process, thereby reducing uncertainty [EFRON 1979; EFRON, TIBISHIRANI 1993].

Assume a population T with unknown probability distribution F, where $t_i = (x_i, y_i)$ is a realization drawn independently and identically distributed (i.i.d.) from T, x_i is a input vector and y_i is the corresponding output vector, and n is the size of original dataset. In this case:

$$T_n = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \tag{16}$$

is a bootstrap resample obtained from an empirical distribution function, F with a mass of $1/n$ for each t_1, t_2, …, t_n. Similarly, a set of bootstrap samples such as T^1, T^2, …, T^s, …, T^S can be produced, where s is a particular bootstrap resample, whereas S is the total number of bootstrap resamples. In such a case the total number of bootstrap samples, S, usually ranges from between 50 and 200 samples [EFRON 1979].

In this study, several bootstrap resamples were generated and used to train several different ANN and ANN_W models. Ensemble forecasts were obtained and designated ANN_B and ANN_{WB}, respectively. For each T^s, an ANN and ANN_W model was developed and trained using all n observations. The output of ANN models was represented in terms of bootstrap resamples and corresponding optimized weights as $f_{NN}(x_i, w_s)$ where x_i was the input data pattern, and w_s was the optimized weights of the ANN model for a particular bootstrap resample s. The performance of

both the models was then evaluated using a set A_s. Then the generalization error denoted as $\hat{E}_0$ was estimated (*e.g.*, ANN model) as [TWOMEY, SMITH 1998]:

$$\hat{E}_0 = \frac{\sum_{s=1}^{S}\sum_{i\in A_s}\left[y_i - f_{NN}(x_i, w_s)\right]^2}{\sum_{s=1}^{S}\#(A_s)} \tag{17}$$

A_s is a set of observation pairs $t_i = (x_i, y_i)$ that were not included in generating the bootstrap resamples or it is the set of data patterns in the testing data set or the dataset not included in the development of different resamples. S is the number of bootstrap samples generated from the training dataset.

Finally, the estimate $y(x)$ of the ANN_B, ANN_W and ANN_B were presented as the average of the S bootstrapped estimates of the corresponding ANN model (*e.g.*, ANN_B) as [TIWARI, CHATTERJEE 2011]:

$$\hat{y}(x_i) = \frac{\sum_{s=1}^{S} f_{NN}(x_i, w_s)}{S} \tag{18}$$

and the variance using S resample was given by:

$$\sigma_S^2(x_i) = \frac{\sum_{s=1}^{S}\left[f_{NN}(x_i, w_s) - \hat{y}(x_i)\right]^2}{S-1} \tag{19}$$

A number of forecasts obtained with ANN and ANN_W models, trained with multiple realizations of the training dataset, served to generate a 95% confidence interval (CI), *i.e.*, two tailed $\alpha = 0.05$ significance level. These CIs indicated the frequency with which the CIs would contain the true value in the repeated application of the model. A $100\cdot(1-\alpha)$ percent CI covering the overall UWD water demand $\hat{y}(x)$ can be estimated as [EFRON, TIBSHIRANI 1993]:

$$\begin{aligned}&\text{CI} = (\text{LB}, \text{UB}) = \\ &= [\hat{y}(x) + t_{n-p}^{\alpha/2}\sigma_s(x), \hat{y}(x) - t_{n-p}^{\alpha/2}\sigma_s(x)\end{aligned} \tag{20}$$

where n is the total number of water demand observations, p is the total number of parameters in the NN and WNN models, $t_{n-p}^{\alpha/2}$ is the $\alpha/2$ percentile for the Student distribution, with $n - p$ degrees of freedom, UB is the upper bound, LB is the lower bound, and $\sigma_S(x)$ is the standard deviation of s bootstrapped estimates.

PERFORMANCE INDICES

The developed models' performances were evaluated using five statistical indices: the coefficient of determination (R^2), root mean square error (*RMSE*), persistence index (*PI*), mean absolute error (*MAE*), and peak percentage deviation (P_{dv}), as defined below [DAWSON *et al.* 2007].

(i) The coefficient of determination (R^2) performance index is a squared ratio of the combined dispersion of two time series to the total dispersion of the observed and modelled time series. It presents the overall agreement between observed and modelled time series and varies from 0 for a poor model to 1 for a perfect model.

The coefficient of determination (R^2) is expressed as:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2 \sum_{i=1}^{n}(P_i - \overline{P})^2}}\right)^2 \tag{21}$$

where n is the number of data points, O_i and P_i are the i^{th} observed and i^{th} forecasted UWD values, respectively, and O and P are the observed and forecasted UWD means, respectively.

(ii) The root mean square error (*RMSE*) is a measure of overall performance across the entire range of the dataset and provides a good measure of model performance for high flows [KARUNANITHI *et al.* 1994], as it is sensitive to small differences in model performance and exhibits high sensitivity to the larger errors occurring for higher magnitudes. It is expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \tag{22}$$

The $RMSE \geq 0$, and shows a perfect model fit $RMSE = 0$.

(iii) Percentage deviation in peak (P_{dv}). It is defined as:

$$P_{dv} = 100\sum\frac{P^{peak} - O^{peak}}{O^{peak}} \tag{23}$$

where, O^{peak} and P^{peak} are the peak of observed and forecasted water demand, respectively.

(iv) The mean absolute error (*MAE*) measures overall agreement between observed and forecasted values, but is not weighted towards higher or lower magnitude events. It evaluates all deviations from the observed values equally, without considering sign. So $MAE \geq 0$, with $MAE = 0$ representing a perfect model fit to observed values. It is expressed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|O_i - P_i\right| \tag{24}$$

(v) The persistence index (*PI*) is one minus the ratio of the sum square error (*SSE*) to the same *SSE* obtained when the last observed value itself is considered as the forecasted value for a particular lead time. The more *PI* exceeds zero the greater the model's accuracy; however, if $PI = 0$ then the model has performed no better than a one parameter 'no knowledge' model, while if $PI < 0$ the model has performed more poorly than a 'no knowledge' model. It is essentially a comparison between the model under study and a simple naïve persistence model. Thus, to estimate PI the predicted water demand at time t (P_i) is considered as the observed water demand at time t-j (O_{i-j}), where j indicates the lead time selected for the water demand forecast. *PI* is therefore expressed as:

$$PI = 1 - \frac{\sum_t (O_i - P_i)^2}{\sum_t (O_i - O_{i-1})^2} = 1 - \frac{SSE}{SSE_{naive}} \tag{25}$$

MATERIALS AND METHODOLOGY

STUDY AREAS AND DATA PARTITIONING

With a population of approximately 1.1 million people, Calgary is amongst the largest cities in Canada [City of Calgary 2011]. The Bearspaw Plant treats water from the Bow River primarily to supply the northern half of the city, while the Glenmore Plant treats water from the Elbow River and supplies the southern portion of the city [City of Calgary 2011]. Each plant supplies about half of Calgary's total drinking water needs, and the 4600 km distribution system is interconnected through transmission mains. Since 1980, the city has invested in maintenance of the network by replacing corroded pipes with PVC and by adding cathodic protection on pipes to reduce the rate of corrosion. As a result, emergency repairs have been reduced by 73% [City of Calgary 2011].

In 2010, the total per capita water demand in Calgary was 406 $L \cdot d^{-1}$, while residential use was 257 $L \cdot d^{-1}$, less than the Canadian average of 343 $L \cdot d^{-1}$ but still greater than other Prairie cities [Environmental and Safety Management 2010]. The Government of Alberta announced in 2006 that new water licenses for the Bow River Basin would no longer be granted, which has led to an increased awareness and need for water conservation and a demand for reduction measures in the midst of continuing population growth and climate change. For example, the Calgary City Council has adopted a goal of reducing total per capita use to 350 $L \cdot d^{-1}$ by 2033, metering all residential homes by the end of 2014, and maintaining peak demand below $0.95 \cdot 10^9$ L through to 2032 [Environmental and Safety Management 2010].

The average summer high in Calgary is 20°C, with a historic extreme high of 36°C, and the average winter low is –13°C, with a historic extreme low of –45°C. Annual rainfall in Calgary is about 320 mm, with a recorded extreme daily rainfall of 95.3 mm. Annual snowfall averages around 125 cm, with an extreme daily snowfall of 48.4 cm [Environment Canada 2010].

The data obtained from the City of Calgary consisted of average daily water demand, maximum temperature, and total precipitation, compiled from 25.03.2004 to 31.12.2006. Additional data were not available. For the development of the models, the data was divided into three sets: one for training the models, one for cross-validation to check that the models did not over-fit, and one for testing the performance of the developed models. The details of the data partitioning are shown in Table 1.

Table 1. Partitioning of data for artificial neural network (ANN) model development

Dataset	Period	Number of data patterns
Training	25.03.2004 to 24.03.2005	365
Cross-validation	25.03.2005 to 24.03.2006	365
Testing	25.03.2006 to 31.12.2006	282

Source: own study.

INPUT SELECTION AND DROUGHT MODEL DEVELOPMENT

ANN Model Development

Selection of significant input variables and identifying optimal model structure are two important steps in ANN model development. Correlation statistics (*e.g.*, cross-correlation, auto-correlation and partial auto-correlation) along with a trial and error approach were employed to obtain prior knowledge of input variables dynamics. In this procedure information at different lag times of daily UWD, daily maximum temperature and daily total precipitation were considered. Following this, the optimal network geometry for the ANN model was identified by trial and error, and the number of hidden neurons that produced the lowest generalization error, ranging between 1 and 15, was considered to be the optimal structure [JIA, CULVER 2006]. ANN models were initially developed using the significant inputs that were log-transformed and linearly scaled to a range of 0 to 1 [CAMPOLO *et al.* 1999]. A second-order training method, the Levenberg–Marquardt optimization method was used to minimize the mean square error (*MSE*) between the forecasted and observed *UWD* values.

ELM, ELM_B and ELM_W Model Development

Based on an earlier study that demonstrated the practical use of ELM models for drought forecasting in eastern Australia [DEO, ŞAHIN 2015a, b], a 3-layer network containing input, feature optimisation and output spaces was employed (Fig. 1a, b). The ELM model employed in the current study was developed using the logarithmic sigmoid activation function.

Initially, the ELM model was randomly executed ~50–1000 times to explore input layer weights, weights and optimal nodes in the hidden layer and model biases yielding the smallest MSE. This resulted in ~100 randomisations yielding a stable solution. In each case, the CPU time consumed to run urban water demand models was recorded.

Also shown to be an effective tool in drought forecasting [DEO, ŞAHIN 2015b], an ANN model was developed as a benchmark. The ANN model was stopped early when processing the validation dataset to avoid overtraining or over-fitting [ADAMOWSKI 2008a, b; ADAMOWSKI *et al.* 2012a; TIWARI, ADAMOWSKI 2013]. During this process the MSE was monitored at iterations of the training and during cross-validation phases. The training was stopped when the MSE reached a minimum [BISHOP 1995]. As with previous studies [DEO, ŞAHIN 2015a, b; TIWARI, ADAMOWSKI 2013], the fast and efficient second-order Levenberg–Marquardt training algorithm was employed in the ANN model.

A robust ELM model was developed by considering different input variables and optimization parameters. Two further hybrid models were developed: bootstrap-based ELM (ELM_B), and wavelet-based ELM (ELM_W). MATLAB® (v.7.10.0] code was written to develop all the wavelet models, while bootstrap resamples were generated using an Excel add-in (Bootstrap.xla) [BARRETO, HOWLAND 2006]. The ensemble of roughly 200 ELM_B models were developed for each bootstrap resample dataset, and all the 200 forecasts were later combined to generate an ensemble of all these forecasts. To further improve the ELM, we applied discrete wavelet transformation (DWT) on the predictor signals to achieve a time-scale representation of the localized and transient phenomenon at different scales in the data series [ADAMOWSKI 2008a, b; ADAMOWSKI, CHAN 2011; ADAMOWSKI *et al.* 2012a, b; KIM, VALDÉS 2003; TIWARI, ADAMOWSKI 2013]. The DWT process aimed to achieve a time-scale realisation of both the localized and the transient phenomena at various frequencies. The frequency content and temporal variations was analysed by effectively decomposing inputs into discrete wavelet coefficients (DWCs) to make the non-stationarity obvious and the model more responsive to the variations in frequencies of the input data [TIWARI, ADAMOWSKI 2013].

The utilized wavelet function was adopted from a family of the Daubechies mother wavelet [NOURANI *et al.* 2009; WU *et al.* 2009] whereby the DWT process operated as two sets of functions with a high-pass and a low-pass filter. The predictor variables were passed through the high- and low-pass filters to acquire detail (db1, db2, db3) in terms of high frequency components and approximation coefficients (A3) in terms of low frequency components of the signal. As the performance of the db5 wavelet with three levels of decomposition provided the best performance, for illustrative purposes only 3 levels of decomposition (db1, db2, db3] and 1 approximation (A3) are presented for the UWD data over the tested period (Fig. 2). The low-frequency components reflected by A3 showed the broad-scale patterns in the predictor dataset including its periodicity and trends, and was closely in-phase with the predictor signal, whereas the high-frequency components (db1, db2, db3] appeared to replicate greater details of the subtle but significant patterns in the UWD time-series [KÜÇÜK, AĞIRALI˙OĞLU 2006].

Though earlier studies have demonstrated a better performance of wavelet-based models, the ways in which the wavelet sub time-series are included in model development can vary greatly. Some studies have used all of their wavelet sub-series [ADAMOWSKI, SUN 2010; NOURANI *et al.* 2009; WANG, DING 2003] whereas others have removed the db1 sub-series and added the remaining series, considering the former series as noise due to its low correlation with their original data [KISI, CIMEN 2011; PARTAL, KIŞI 2007; RAJAEE *et al.* 2010]. However, in some studies, new wavelet time-series were developed by adding up the effective DWCs based on regression correlation [TIWARI, CHATTERJEE 2010b; 2011]. As it is wise not to completely rely on a model based on a particular wavelet series that captures some phenomena at the expense of others [RATHINASAMY *et al.* 2013], we considered each wavelet function in terms of its own strengths in capturing stochastic characteristics and physical structure of the hydrological dataset.

ANN_B and ANN_W Model Development

ANN_B models were developed in similar manner as ELM_B models. The ANN_W models were developed by inputting the wavelet sub time series produced using DWCs. In this study, wavelet functions from the Daubechies family of wavelets [NOURANI *et al.* 2009; WU *et al.* 2009] were used, and three levels of decomposition were considered based on the following formula [NOURANI *et al.* 2008]:

$$L = \text{int}\,[\log(n)] \tag{26}$$

where, L is the number of decomposition levels, and n is the number of time series data.

The number of datasets $n = 1012$ yield a value of $L = 3$, leading to three levels of decomposition (d1, d2, and d3) and approximation (A3) for the data (Fig. 2). The effective DWCs were determined using the correlation coefficients between each wavelet component and the observed UWD. The correlation between the original daily time series for Calgary and corresponding different wavelet sub-time series are shown in Table 2.

In earlier studies [ADAMOWSKI, SUN 2010; KIŞI 2010; TIWARI, CHATTERJEE 2010a; 2011], the significant wavelet sub-time series of a particular time series was used and added to generate a new time series,

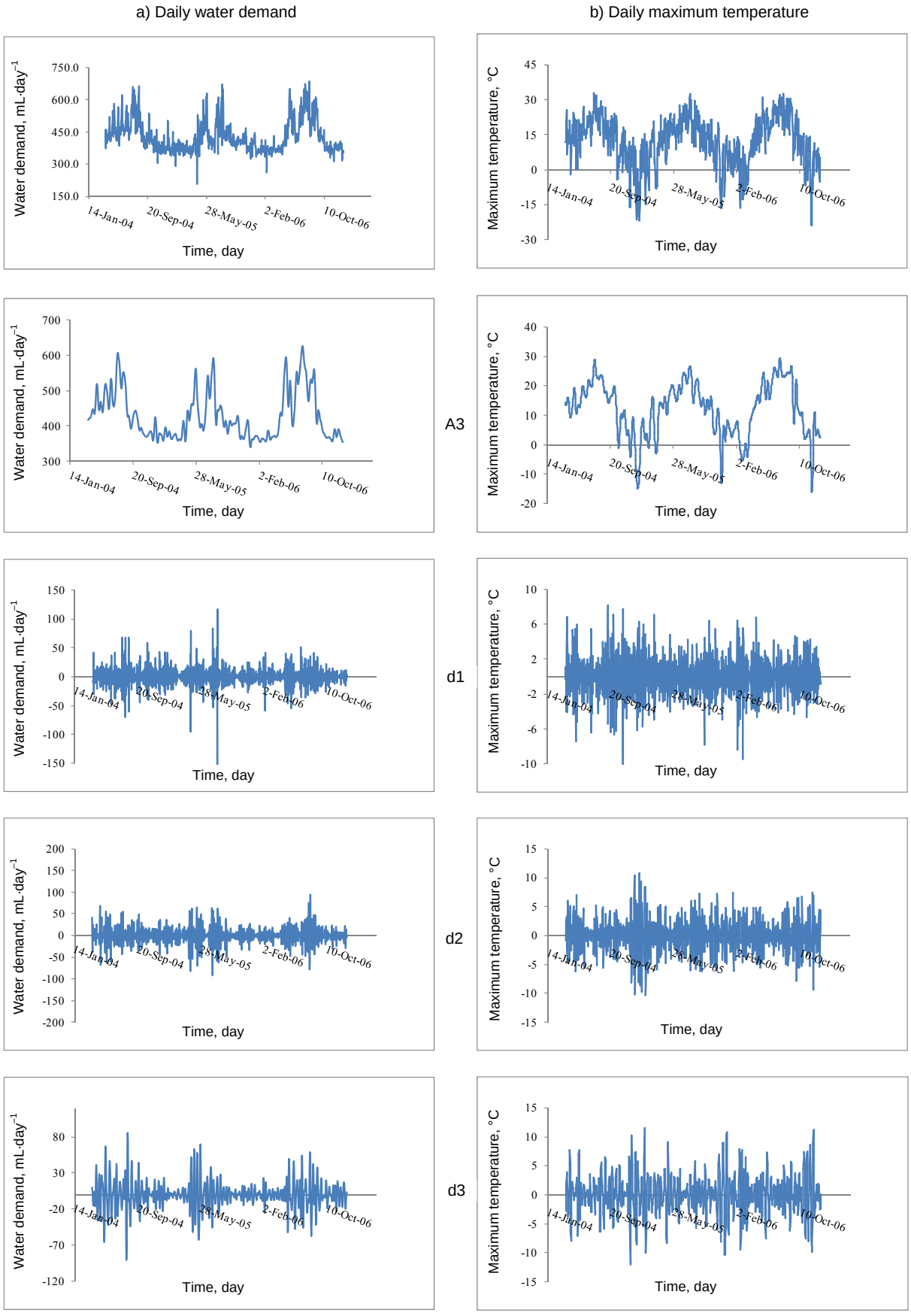

Fig. 2. Wavelet sub-time series of the (a) daily water demand and (b) daily maximum temperature of Calgary from 24 March, 2004 to 31 December, 2006; A3 = approximation, d1, d2, d3 = details; source: own study

Table 2. Correlations between different wavelet sub-time series and the original time series

Wavelet sub-time series	Water demand	Maximum temperature	Total precipitation
A3	0.40	0.32	0.01
d1	0.10	0.02	–0.06
d2	0.15	0.06	–0.13
d3	0.16	0.06	–0.10

Source: own study.

becoming new inputs with which to develop the ANN_W model. In this study, a threshold allowable correlation level of 0.1 was used in determining the inclusion and use of all DWCs in the model development process. Moreover, considering that all the wavelet components play a different role in the original time series, ANN_W models were developed using all components as separate inputs. As with the development of the ANN and ANN_B models, data partitioning for training, cross-validation, and testing was done in a manner similar to that employed for ANN_W models.

RESULTS AND GENERAL DISCUSSION

WATER DEMAND FORECASTING IN CALGARY USING ELM, B-ELM AND W-ELM MODELS

For 1 day lead-time UWD forecasting, the significant input variables were water demand at time t [*i.e.*, WatDemand(t)] and maximum temperature at time t [*i.e.*, MaxT(t)]. The statistical and graphical (scatter plots) assessment of ELM models' performance in UWD forecasting are presented in Table 3 and Figure 3a, respectively. Statistical performance metrics for the ELM model were generally satisfactory, indicating that the margin of difference between actual and forecasted UWD was relatively small. Considering the range of water demand in Calgary (313.38 mL·d^{-1} ≤ UWD ≤ 684.25 mL·d^{-1}), the performance of the best ELM models (*RMSE* = 33.02 mL·d^{-1}, can be viewed as satisfactory for 1-day lead-time UWD forecast.

Table 3. Water demand forecasting for 1 day lead times using the best models for testing dataset

Model	Lead time	R^2	*RMSE* mL·day^{-1}	P_{dv} %	*MAE* mL·day^{-1}
ELM	1	0.850	33.02	10.06	24.42
B-ELM	1	0.851	32.97	9.61	24.14
W-ELM	1	0.927	23.11	6.04	16.70

Source: own study.

This study further explored the capacity of bootstrap-based ensemble modelling along with ELM models for UWD forecasting. Some 200 bootstrap resamples of the training dataset were generated and ELM_B models were developed by averaging the forecasts obtained from 200 resultant ELM. The statistical and graphical (scatter plots) assessment of ELM_B models' performance in UWD forecasting are presented in Table 3 and Figure 3b, respectively. On the basis of these assessments the ELM_B could be seen to slightly outperform the ELM model. Given that the ELM_B model was developed using several realization of the training dataset it was expected to produce stable and reliable results even if the pattern of the training and testing datasets changed.

The efficacy in daily UWD forecasting of ELM_W models relative to other machine learning models was also considered. In order to develop the ELM_W models, all the time series datasets were decomposed into approximation and details and all the components were considered separately as inputs for ELM_W model development. The ELM_W models were developed using wavelet sub-time series derived from the dataset that produced the best ELM model for Calgary. Based on a trial and error process the best ELM_W models for 1-day lead-time forecasts were found with inputs a3(t), d3(t), d2(t), and d1(t) at time t of wavelet components A_3, d_1, d_2, d_3 of daily water demand (WatDemand), daily maximum temperature (Max$T$) and total precipitation (Tot$P$), respectively. The d_1 component showed a lesser correlation with the original water demand than the other components (*i.e.*, A_3, d_2, and d_3), indicating that d_1 components may be redundant/noise information contained in the original water demand time series (Tab. 2). However, it was found to not be altogether without importance during UWD forecasting, as it showed that all the components exhibit some important information about the physical characteristics of the original UWD time series. The statistical and graphical (scatter plots) assessment of ELM_W models' performance in UWD forecasting are presented in Table 3 and Figure 3c, respectively. The scatter plots revealed the ELM_W model to significantly outperform the unenhanced ELM model (R^2 = 0.927 *vs.* R^2 = 0.850, respectively) in UWD forecasting. This showed the supremacy of ELM_W models in UWD forecasting compared to ELM and ELM_B models.

WATER DEMAND FORECASTING IN CALGARY USING ANN, B-ANN AND W-ANN MODELS

In an earlier study TIWARI and ADAMOWSKI [2015] developed models for UWD forecasting by applying the same input variables for the same dataset length. One can therefore compare the performance of newly developed ELM, ELM_B and ELM_W models with the earlier applied ANN, ANN_B and ANN_W models. The structure and performance of the best ANN models for the UWD forecasting testing dataset for Calgary are shown in Table 4. For 1-day lead-time UWD forecasting, the significant input variables obtained were water demand at time t [*i.e.*, WatDemand(t)] and maximum temperature at time t [*i.e.*, Max$T$(t)]. The number of optimum hidden neurons was identified as 3.

The hydrographs and scatter plots of observed and forecasted UWD values using the best ANN models for 1 day lead-time UWD prediction (Fig. 4a) show that ANN models' performance to be acceptable, with forecasted values following the general trend

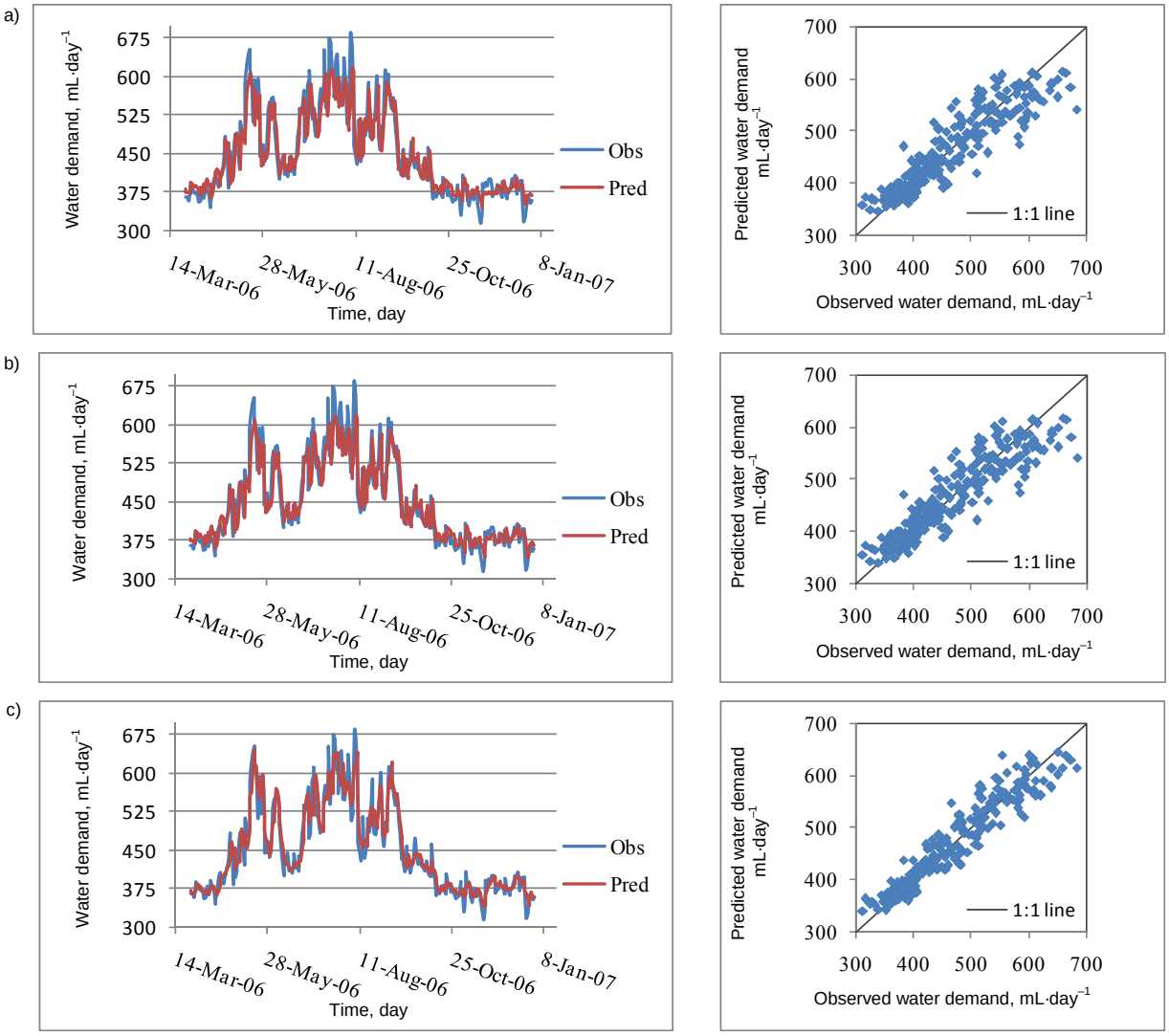

Fig. 3. Hydrographs and scatter plots for observed (Obs) and predicted (Pred) water demand in Calgary for 1 day lead time forecasts for the testing dataset using the best models: a) ELM, b) B-ELM, c) W-ELM; source: own study

Table 4. Water demand forecasting for 1 day lead times using the best models for testing dataset

Model	Lead time	R^2	*RMSE* mL·day^{-1}	P_{dv} %	*MAE* mL·day^{-1}
ANN	1	0.860	32.97	11.92	24.31
B-ANN	1	0.854	33.71	13.47	24.67
W-ANN	1	0.924	24.15	8.25	17.48

Source: own study.

of observed values and yielding an almost 1:1 regression line in forecasted *vs* observed scatter plots. The performance of ANN models was comparable to that of ELM models; however, the ELM models were more time efficient than the ANN models (ELM ≈ 2 sec, ELM_B ≈ 400 sec, ELM_W ≈ 1 sec, ANN ≈ 4 sec, ANN_B ≈ 800 sec and ANN_W ≈ 3 sec CPU time).

Similar to ELM_B models, the ANN_B models were developed using bootstrap resamples of the training dataset used to develop the best ANN models. For each lead time, results from 200 ANN models developed from 200 bootstrap resamples were averaged to generate the ANN_B forecast. Based on observed *vs.* forecasted UWD scatter plots (Fig. 4b), and statistical performance indices (Tab. 4) for the testing dataset the performance of the ANN_B models to be comparable to that of ANN models for 1-day lead-time UWD forecasting, but their accuracy decreases significantly for longer lead-times. The performance of the ANN_B models was slightly less accurate than those of the ANN models. Likewise, the performance of ELM_B model was slightly better than that of the ANN_B model for UWD forecasting.

The statistical and graphical (scatter plots) assessment of ANN_W models' performance in UWD forecasting with 1-day lead-times are presented in Table 4 and Figure 4c, respectively. The best ANN_W model outperformed the best ANN and ANN_B models, demonstrating the ability of wavelet analysis to capture useful information from different periodic components (i.e. wavelet sub-time series). The ELM_W model performed slightly better than the ANN_W models for UWD forecasting.

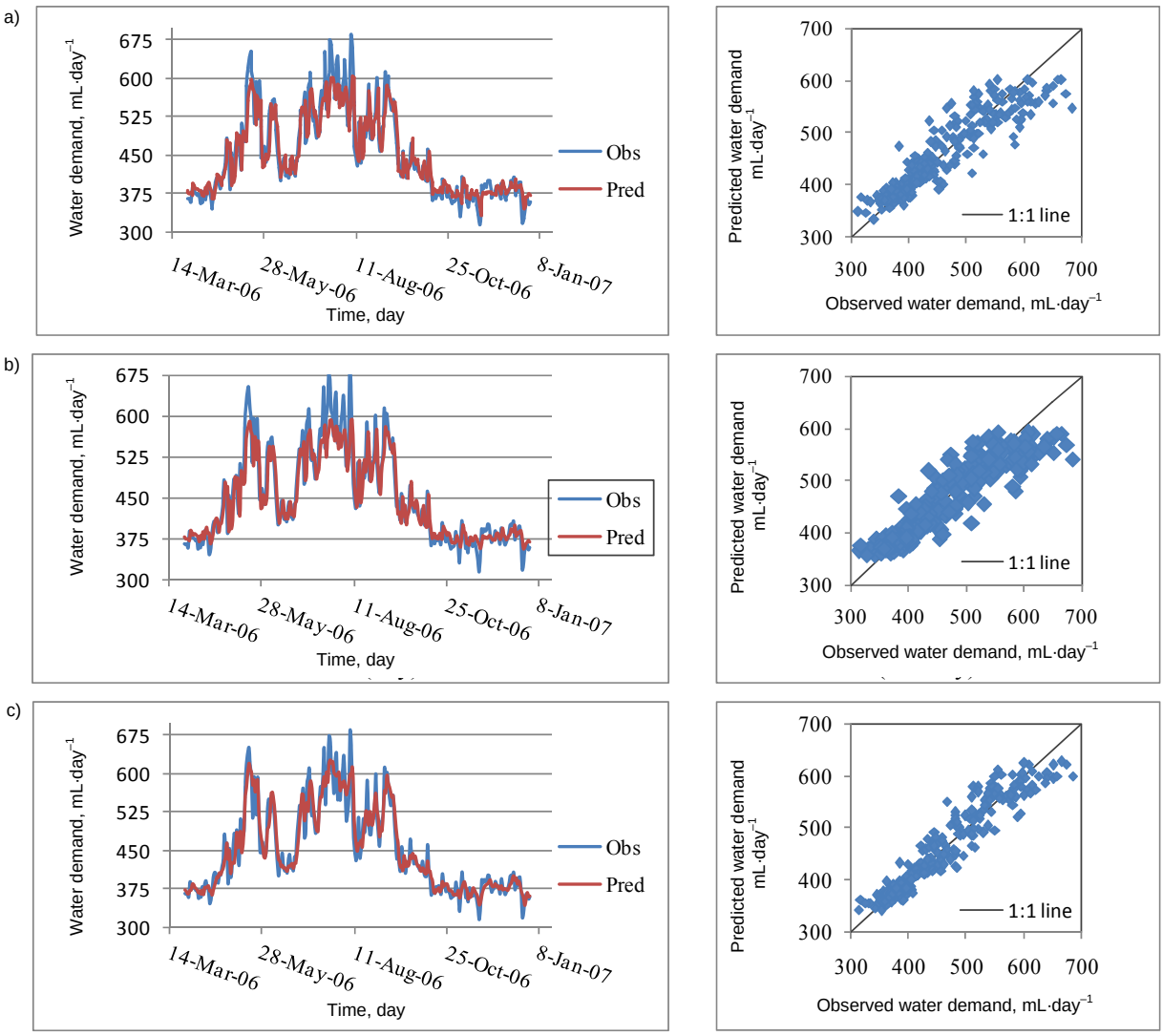

Fig. 4. Hydrographs and scatter plots for observed (Obs) and predicted (Pred) water demand in Calgary for 1 day lead time forecasts for the testing dataset using the best models: a) ANN, b) B-ANN, c) W-ANN; source: own study

CONCLUSIONS

Accurate and reliable UWD forecasting is necessary to help transition to more effective and sustainable urban water resources planning and management [BUTLER, ADAMOWSKI 2015; HALBE *et al.* 2013; 2014; INAM *et al.* 2015; KOLINJIVADI *et al.* 2014; STRAITH *et al.* 2014]. In this study, ELM_W models based on their capacities of wavelet transformation and ELM modeling techniques were employed to simulate the UWD in the city of Calgary, Canada. A limited yet more appropriate set of predictor variables were utilized. The feasibility of using ELM_W for UWD forecasting was compared to that of traditional ELM and ANN models, as well as ANN_B and ANN_W models. In this study ELM, ANN, ELM_W, ANN_W and ANN_B models were developed for 1-day lead time UWD forecasting for the city of Calgary, Canada. Based on five performance indices (R^2, P_{dv}, *RMSE*, *PI*, *MAE*), ELM_W models were found to perform considerably better than ANN, ELM, ANN_B, and ANN_W models. This highlights the ability of wavelet transformation to decompose time-series data with non-stationarity into discrete wavelet components, highlighting cyclic patterns and trends in the time series data at varying temporal and spatial scales and making the data readily usable in forecasting. Indeed, the margins of prediction errors were much smaller for the ELM_W and the model execution time was shorter compared to the other machine learning models considered in this work. Therefore, as a pioneer study on the application of ELM_W modelling to UWD prediction, this research clearly demonstrated the feasibility of wavelet-based modelling for UWD forecasting. Moreover, this study provides a promising advancement to machine learning models for UWD studies and the opportunity to explore the ELM and wavelet techniques in real-time UWD forecasting.

REFERENCES

ABBOT J., MAROHASY J. 2012. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. Advances in Atmospheric Sciences. Vol. 29. Iss. 4 p. 717–730. DOI: 10.1007/s00376-012-1259-9.

ABRAHART R.J. 2003. Neural network rainfall-runoff forecasting based on continuous resampling. Journal of Hydroinformatics. Vol. 5. Iss. 1 p. 51–61.

ACHARYA N., SHRIVASTAVA N.A., PANIGRAHI B., MOHANTY U. 2014. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. Climate Dynamics. Vol. 43. Iss. 5 p. 1303–1310. DOI: 10.1007/s00382-013-1942-2.

ADAMOWSKI J. 2007. Development of a short-term river flood forecasting method based on wavelet analysis. Warsaw, Poland. Polish Academy of Sciences.

ADAMOWSKI J.F. 2008a. Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. Journal of Hydrology. Vol. 353. Iss. 3 p. 247–266. DOI: 10.1016/j.jhydrol.2008.02.013.

ADAMOWSKI J.F. 2008b. Peak daily water demand forecast modeling using artificial neural networks. Journal of Water Resources Planning and Management. Vol. 134. Iss. 2. p. 119–128. DOI: 10.1061/(ASCE)0733-9496(2008)134:2(119).

ADAMOWSKI J., ADAMOWSKI K., BOUGADIS J. 2010. Influence of trend on short duration design storms. Water Resources Management. Vol. 24 p. 401–413.

ADAMOWSKI J., ADAMOWSKI K., PROKOPH A. 2013. A spectral analysis based methodology to detect climatological influences on daily urban water demand. Mathematical Geosciences. Vol. 45. Iss. 1 p. 49–68. DOI: 10.1007/s11004-012-9427-0.

ADAMOWSKI J., CHAN H.F. 2011. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology. Vol. 407. Iss. 1 p. 28–40. DOI: 10.1016/j.jhydrol.2011.06.013.

ADAMOWSKI J., CHAN H.F., PRASHER S.O., OZGA-ZIELINSKI B., SLIUSARIEVA A. 2012a. 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 Resources Research. Vol. 48 W01528 p. 1–14. DOI: 10.1029/2010WR 009945.

ADAMOWSKI J., CHAN H.F., PRASHER S.O., SHARDA V.N. 2012b. Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan microwatersheds with limited data. Journal of Hydroinformatics. Vol. 14. Iss. 3 p. 731–744. DOI: 10.2166/hydro.2011.044.

ADAMOWSKI J., PROKOPH A. 2013. Assessing the impacts of the urban heat island effect on streamflow patterns in Ottawa, Canada. Journal of Hydrology. Vol. 496 p. 225–237.

ADAMOWSKI J., PROKOPH A., ADAMOWSKI K. 2012c. Influence of the 11 year solar cycle on annual streamflow maxima in Southern Canada. Journal of Hydrology. Vol. 442 p. 55–62.

ADAMOWSKI J., SUN K. 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology. Vol. 390. Iss. 1 p. 85–91. DOI: 10.1016/j.jhydrol.2010.06.033.

ADAMOWSKI K., PROKOPH A., ADAMOWSKI J. 2009. Development of a new method of wavelet aided trend detection and estimation. Hydrological Processes. Vol. 23 p. 2686–2696.

ALVISI S., FRANCHINI M. 2011. Fuzzy neural networks for water level and discharge forecasting with uncertainty. Environmental Modelling and Software. Vol. 26. Iss. 4 p. 523–537. DOI: 10.1016/j.envsoft.2010.10.016.

ALVISI S., FRANCHINI M. 2012. Grey neural networks for river stage forecasting with uncertainty. Journal of Physics and Chemistry of the Earth. Vol. 42 p. 108–118. DOI: 10.1016/j.pce.2011.04.002.

ANCTIL F., TAPE D. 2004. An exploration of artificial neural network rainfall–runoff forecasting combined with wavelet decomposition. Journal of Environment Engineering and Science. Vol. 3 (Suppl.) p. 121–128. DOI: 10.1139/s03-071.

ARAGHI A., ADAMOWSKI J., NALLEY D., MALARD J. 2015 Using wavelet transforms to estimate surface temperature trends and dominant periodicities in Iran based on gridded reanalysis data. Journal of Atmospheric Research. Vol. 11 p. 52–72.

BARRETO H., HOWLAND F.M. 2006. Introductory econometrics: Using Monte Carlo simulation with Microsoft Excel. Cambridge, UK. Cambridge University Press.

BELAYNEH A., ADAMOWSKI J., KHALIL B., OZGA-ZIELINSKI B. 2014. Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural networks and wavelet support vector regression models. Journal of Hydrology. Vol. 508 p. 418–429. DOI: 10.1016/ j.jhydrol.2013.10.052.

BISHOP C.M. 1995. Neural Networks for Pattern Recognition. Oxford, UK, Clarendon Press.

BUTLER C., ADAMOWSKI J. 2015. Empowering marginalized communities in water resources management: Addressing inequitable practices in Participatory Model Building. Journal of Environmental Management. Vol. 153 p. 153–162.

CAMPISI-PINTO S., ADAMOWSKI J., ORON G. 2012. Forecasting urban water demand via wavelet-denoising and neural network models. Case study: city of Syracuse, Italy. Water Resources Management. Vol. 26. Iss. 12 p. 3539–3558. DOI: 10.1007/s11269-012-0089-y.

CAMPOLO M., ANDREUSSI P., SOLDATI A. 1999. River flood forecasting with a neural network model. Water Resources Research. Vol. 35. Iss. 4 p. 1191–1197. DOI: 10.1029/1998WR900086.

City of Calgary. Environmental and Safety Management 2010 [online]. 2010 State of the environment. The Report 4th ed. pp. 54. [Access 25.10.2015]. Available at: www.calgary.ca/UEP/ESM/Documents/ESM-Documents/2010-state-of-the-environment-report.PDF

City of Calgary 2011. Civic Census results. Calgary, AB: Election and Information Services.

COULIBALY P., EVORA N. 2007. Comparison of neural network methods for infilling missing daily weather records. Journal of Hydrology. Vol. 341. Iss. 1 p. 27–41. DOI: 10.1016/j.jhydrol.2007.04.020.

DAUBECHIES I. 1990. The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory. Vol. 36. Iss. 5 p. 6–7. DOI: 10.1109/18.57199.

DAVIS B., KIEFER J. 2005. Demanding a better way: The benefits of demand forecasting [online]. Public Works (Oct.). [Access 20.10.2015]. Available at: http://www.pwmag.com/water-conservation/demanding-a-better-way_1.aspx

DAWSON C.W., ABRAHART R.J., SEE L.M. 2007. HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. Envi-

ronmental Modeling and Software. Vol. 22. Iss. 7 p. 1034–1052. DOI: 10.1016/j.envsoft.2006.06.008.

DENNIS JR J.E., SCHNABEL R.B. 1996. Numerical methods for unconstrained optimization and nonlinear equations. Philadelphia: Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611971200.

DEO R.C., ŞAHIN M. 2015a. Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydrometeorological parameters and climate indices in eastern Australia. Atmospheric Research. Vol. 161 p. 65–81. DOI: 10.1016/j.atmosres.2015.03. 018.

DEO R.C., ŞAHIN M. 2015b. Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. Atmospheric Research. Vol. 153 p. 512–525. DOI: 10.1016/j.atmosres.2014.10.016.

EFRON B. 1979. Bootstrap methods: another look at the Jackknife. Annals of Statistics. Vol. 7. Iss. 1 p. 1–26. DOI: 10.1214/aos/1176344552.

EFRON B., TIBSHIRANI R.J. 1993. An introduction to the Bootstrap. Monographs on Statistics and Applied Probability. Book 57. London, UK. Chapman and Hall. ISBN 978-04-12042-31-7 pp. 456.

Environment Canada 2010 [online]. [Access 25.10.2015]. Available at: http://www.climate.weather office.gc.ca

FRANCESCO V., BERND F. 2000. Nonstationarity and data preprocessing for neural network predictions of an economic time series. Proceedings of International Joint Conference Neural Network. Vol. 5. p. 129–134. DOI: 10.1109/IJCNN.2000.861446.

GHIASSI M., ZIMBRA D.K., SAIDANE H. 2008. Urban water demand forecasting with a dynamic artificial neural network model. Journal of Water Resources Planning and Management. Vol. 134. Iss. 2 p. 138–146. DOI: 10.1109/IJCNN.2000.861446.

GOYAL M., BHARTI B., QUILTY J., ADAMOWSKI J., PANDEY A. 2014 Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert Systems with Applications. Vol. 41. Iss. 11 p. 5267–5276.

HAIDARY A., AMIRI B.J., ADAMOWSKI J., FOHRER N., NAKANE K. 2013. Assessing the impacts of four land use types on the water quality of wetlands in Japan. Water Resources Management. Vol. 27. Iss. 7. p. 2217–2229.

HALBE J., ADAMOWSKI J., BENNETT E., PAHL-WOSTL C., FARAHBAKHSH K. 2014. Functional organization analysis for the design of sustainable engineering systems. Ecological Engineering. Vol. 73 p. 80–91.

HALBE J., PAHL-WOSTL C., SENDZIMIR J., ADAMOWSKI J. 2013. Towards adaptive and integrated management paradigms to meet the challenges of water governance. Water Science and Technology: Water Supply. Vol. 67 p. 2651–2660.

HAN D., KWONG T., LI S. 2007. Uncertainties in real-time flood forecasting with neural networks. Hydrological Processes. Vol. 21 p. 223–228. DOI: 10.1002/hyp.6184.

HANEMANN W.M. 1998. Determinants of urban water use. In: Urban water demand management and planning. Eds. D. Baumann, J. Boland, M. Hanemann. New York. McGraw Hill p. 31–75.

HERRERA M., TORGO L., IZQUIERDO J., PÉREZ-GARCÍA R. 2010. Predictive models for forecasting hourly urban water demand. Journal of Hydrology. Vol. 387. Iss. 1–2 p. 141–150. DOI: 10.1016/j.jhydrol.2010.04.005.

HOUSE-PETERS L.A., CHANG H. 2011. Urban water demand modeling: Review of concepts, methods, and organizing principles. Journal of Water Resources Research. Vol. 47 W05401. DOI: 10.1029/2010WR009624.

HUANG G., HUANG G.-B., SONG S., YOU K. 2015. Trends in extreme learning machines: A review., Neural Networks. Vol. 61 p. 32–48. DOI: 10.1016/j.neunet.2014.10.001.

HUANG G.-B., ZHOU H., DING X., ZHANG R. 2012. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics. Part B. Cybernetics. Vol. 42. Iss. 2 p. 513–529. DOI: 10.1109/TSMCB.2011.2168604.

HUANG G.-B., ZHU Q.-Y., SIEW C.-K. 2006. Extreme learning machine: Theory and applications. Neurocomputing. Vol. 70. Iss. 1–3 p. 489–501. DOI: 10.1016/j.neucom.2005.12.126.

INAM A., ADAMOWSKI J., HALBE J., PRASHER S. 2015. Using causal loop diagrams for the initialization of stakeholder engagement in soil salinity management in agricultural watersheds in developing countries: A case study in the Rechna Doab watershed, Pakistan. Journal of Environmental Management. Vol. 152 p. 251–267.

ISUKAPALLI S., GEORGOPOULOS P. 2001. Computational methods for sensitivity and uncertainty analysis for environmental and biological models. National Exposure Research Laboratory USEPA. EPA/600/R-01/068 (NTIS PB2004-102518) Washington, DC: USEPA pp. 145.

JAIN A., ORMSBEE L.G. 2002. Short-term water demand forecast modeling techniques-conventional methods versus AI. Journal. (American Water Works Association). Vol. 94. Iss. 7 p. 64–72.

JEONG D., KIM Y. 2005. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. Hydrological Processe. Vol. 19. Iss. 19 p. 3819–3835. DOI: 10.1002/hyp.5983.

JIA Y., CULVER T.B. 2006. Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. Journal of Hydrology. Vol. 331. Iss. 3–4 p. 580–590. DOI: 10.1016/j.jhydrol.2006.06.005.

KAME'ENUI A. 2003. Water demand forecasting in the Puget Sound Region: Short and long-term models. M.Sc. thesis. Seattle, Washington, U.S.A. Department of Civil and Environmental Engineering, University of Washington pp. 172.

KARRAN D., MORIN E., ADAMOWSKI J. 2014. Multi-step streamflow forecasting using data-driven non-linear methods in contrasting climate regimes. Journal of Hydroinformatics. Vol. 16. Iss. 3 p. 671–689.

KARUNANITHI N., GRENNEY W., WHITLEY D., BOVEE K. 1994. Neural networks for river flow prediction. Journal of Computing in Civil Engineering. Vol. 8. Iss. 2 p. 201–220. DOI: 10.1061/(ASCE)0887-3801(1994)8:2(201).

KAYAGA S., SMOUT I. 2008. Water demand management: a key building block for sustainable urban water management [online]. Annual Conference of Allied Social Sciences Association: 2008 Annual Meeting, Transportation and Public Utilities Research Group, New Orleans, USA, 4th-6th January. [Access 20.10.2015]. Available at: https://dspace.lboro.ac.uk/ 2134/8240

KESKIN M.E., TERZI Ö. 2006. Artificial neural network models of daily pan evaporation. Journal of Hydrologic Engineering. Vol. 11. Iss. 1 p. 65–70. DOI: 10.1061/(ASCE)1084-0699(2006)11:1(65).

KHAN M.S., COULIBALY P. 2006. Application of support vector machine in lake water level prediction. Journal of Hydrologic Engineering. Vol. 11. Iss. 3 p. 199–205. DOI: 10.1061/(ASCE)1084-0699(2006)11:3(199).

KIM T.-W., VALDÉS J.B. 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. Journal of Hydrologic Engineering. Vol. 8. Iss. 6 p. 319–328. DOI: 10.1061/(ASCE)1084-0699(2003)8:6(319).

KISI O., CIMEN M. 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting, Journal of Hydrology. Vol. 399. Iss. 1–2 p. 132–140. DOI: 10.1016/j.jhydrol.2010.12.041.

KIŞI Ö. 2008. Stream flow forecasting using neuro-wavelet technique. Hydrological Processes. Vol. 22. Iss. 20 p. 4142–4152. DOI: 10.1002/hyp.7014.

KIŞI Ö. 2009. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. Journal of Hydrologic Engineering. Vol. 14. Iss. 8 p. 773–782. DOI: 10.1061/(ASCE)HE.1943-5584.0000053.

KIŞI Ö. 2010. Wavelet regression model for short-term streamflow forecasting. Journal of Hydrology. Vol. 389. Iss. 3–4 p. 344–353. DOI: 10.1016/j.jhydrol.2010.06.013.

KOLINJIVADI V., ADAMOWSKI J., KOSOY N. 2014. Recasting payments for ecosystem services (PES) in water resource management: A novel institutional approach. Ecosystem Services. Vol. 10 p. 144–154.

KÜÇÜK M., AĞIRALI˙ OĞLU N. 2006. Wavelet regression technique for streamflow prediction. Journal of Applied Statistics. Vol. 33. Iss. 9 p. 943–960. DOI: 10.1080/02664760600744298.

MAHESWARAN R., KHOSA R. 2012. Comparative study of different wavelets for hydrologic forecasting. Computers and Geosciences. Vol. 46 p. 284–295. DOI: 10.1016/j.cageo.2011.12.015.

MALLAT S.G. 1989. A theory for multi resolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 11. Iss. 7 p. 674–693. DOI: 10.1109/34.192463.

MARQUARDT D.W. 1963. An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics. Vol. 11. Iss. 2 p. 431–441. DOI: 10.1137/0111030.

MCCULLOCH W.S., PITTS W. 1943. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics. Vol. 5. Iss. 4 p. 115–133. DOI: 10.1007/BF02478259.

MEKANIK F., IMTEAZ M., GATO-TRINIDAD S., ELMAHDI A. 2013. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology. Vol. 503 p. 11–21. DOI: 10.1016/j.jhydrol.2013.08.035.

NALLEY D., ADAMOWSKI J., KHALIL B. 2012. Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). Journal of Hydrology. Vol. 475 p. 204–228. DOI: 10.1016/j.jhydrol.2012.09.049.

NALLEY D., ADAMOWSKI J., KHAIIL B., OZGA-ZIELINSKI B. 2013. Utilizing wavelet transforms and the Mann-Kendall test to assess trends in surface air temperature over southern Ontario and Québec. ASABE Paper 131620536. St. Joseph, MI: ASABE. DOI: 10.13031/aim.20131620536.

NAYAK P.C., VENKATESH B., KRISHNA B., JAIN S.K. 2013. Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach. Journal of Hydrology. Vol. 493 p. 57–67. DOI: 10.1016/j.jhydrol.2013.04.016.

NOURANI V., ALAMI M.T., AMINFAR M.H. 2008. A combined neural wavelet model for prediction of watershed precipitation, Ligvanchai, Iran. Engineering Applications of Artificial Intelligence. Vol. 22. Iss. 3 p. 466–472. DOI: 10.1016/j.engappai.2008.09.003.

NOURANI V., BAGHANAM A., ADAMOWSKI J., KISI O. 2014. Applications of hybrid wavelet-artificial intelligence models in hydrology: A review. Journal of Hydrology. Vol. 514 p. 358–377.

NOURANI V., KOMASI M., MANO A. 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resources Management. Vol. 23. Iss. 14 p. 2877–2894. DOI: 10.1007/s11269-009-9414-5.

PARTAL T., KIŞI Ö. 2007. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, Journal of Hydrology. Vol. 342. Iss. 1–2 p. 199–212. DOI: 10.1016/j.jhydrol.2007.05.026.

PINGALE S., KHARE D., JAT M., ADAMOWSKI J. 2014. Spatial and temporal trends of mean and extreme rainfall and temperature for the 33 urban centres of the arid and semi-arid state of Rajasthan, India. Journal of Atmospheric Research. Vol. 138 p. 73–90.

RAJAEE T., MIRBAGHERI S., NOURANI V., ALIKHANI A. 2010. Prediction of daily suspended sediment load using wavelet and neurofuzzy combined model, International Journal of Environmental Science and Technology. Vol. 7. Iss. 1 p. 93–110. DOI: 10.1007/BF03326121.

RATHINASAMY M., ADAMOWSKI J., KHOSA R. 2013. Multiscale streamflow forecasting using a new Bayesian Model Average based ensemble multi-wavelet Volterra nonlinear method. Journal of Hydrology. Vol. 507 p. 186–200. DOI: 10.1016/j.jhydrol.2013.09.025.

RATHINASAMY M., KHOSA R., ADAMOWSKI J., CH S., PARTHEEPAN G., ANAND J., NARSIMLU B. 2015. Wavelet-based multiscale performance analysis: An approach to assess and improve hydrological models. Water Resources Research. Vol. 50. Iss. 12 p. 9721–9737. DOI: 10.1002/2013WR014650.

SAADAT H., ADAMOWSKI J., BONNELL R., SHARIFI F., NAMDAR M., ALE-EBRAHIM S. 2011. Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. Journal of Photogrammetry and Remote Sensing. Vol. 66 p. 608–619.

SEHGAL V., SAHAY R.R., CHATTERJEE C. 2014. Effect of utilization of discrete wavelet components on flood forecasting performance of wavelet based ANFIS models. Water Resources Management. Vol. 28. Iss. 6 p. 1733–1749.

SHARMA S.K., TIWARI K.N. 2009. Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in upper Damodar Valley catchment. Journal of Hydrology. Vol. 374. Iss. 3–4 p. 209–222. DOI: 10.1016/j.jhydrol.2009.06.003.

SRIVASTAV R.K., SUDHEER K.P., CHAUBEY I. 2007. A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. Water Resources Research. Vol. 43. Iss. 10 p. 1–12. DOI: 10.1029/2006WR005352.

STRAITH D., ADAMOWSKI J., REILLY K. 2014. Exploring the attributes, strategies and contextual knowledge of champions of change in the Canadian water sector. Canadian Water Resources Journal. Vol. 39. Iss. 3 p. 255–269.

ŞAHIN M. 2012. Modelling of air temperature using remote sensing and artificial neural network in Turkey. Ad-

vances in Space Research. Vol. 50. Iss. 7 p. 973–985. DOI: 10.1016/j.asr.2012.06.021.

ŞAHIN M., KAYA Y., UYAR M. 2013. Comparison of ANN and MLR models for estimating solar radiation in Turkey using NOAA/AVHRR data. Advances in Space Research. Vol. 51. Iss. 5 p. 891–904. DOI: 10.1016/j.asr.2012.10.010.

ŞAHIN M., KAYA Y., UYAR M., YILDIRIM S. 2014. Application of extreme learning machine for estimating solar radiation from satellite data. International Journal of Energy Research. Vol. 38. Iss. 2 p. 205–212. DOI: 10.1002/er.3030.

TIWARI M., ADAMOWSKI J. 2014. Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models. Water Resources Research. Vol. 49. Iss. 10 p. 6486–6507.

TIWARI M.K., ADAMOWSKI J. 2015a. An ensemble wavelet bootstrap machine learning approach to water demand forecasting: A case study in the city of Calgary, Canada. Urban Water Journal. In Press. DOI: 10.1080/1573062X.2015.1084011.

TIWARI M., ADAMOWSKI J. 2015b. Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach. Journal of Water Resources Planning and Management. Vol. 141. Iss. 2.

TIWARI M.K., CHATTERJEE C. 2010a. Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). Journal of Hydrology. Vol. 382. Iss. 1–4 p. 20–33. DOI: 10.1016/j.jhydrol.2009.12.013.

TIWARI M.K., CHATTERJEE C. 2010b. Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN hybrid approach. Journal of Hydrology. Vol. 394 p. 458–470. DOI: 10.1016/j.jhydrol.2010.10.001.

TIWARI M., CHATTERJEE C. 2011. A new wavelet-bootstrap-ANN hybrid model for daily discharge forecasting. Journal of Hydroinformatics. Vol. 13. Iss. 3 p. 500–519. DOI: 10.2166/hydro.2010.142

TIWARI M.K., SONG K.Y., CHATTERJEE C., GUPTA M.M. 2013. Improving reliability of river flow forecasting using neural networks, wavelets and self-organizing maps. Journal of Hydroinformatics. Vol. 15. Iss. 2 p. 486–502. DOI: 10.2166/hydro.2012.130.

TWOMEY J., SMITH A. 1998. Bias and variance of validation methods for function approximation neural networks under conditions of sparse data. IEEE Transactions on Systems, Man, and Cybernetics. Part C: Applications and Reviews. Vol. 28. Iss. 3 p. 417–430. DOI: 10.1109/5326.704579.

VOGL T., MANGIS J., RIGLER A., ZINK W., ALKON D. 1988. Accelerating the convergence of the backpropagation method. Biological Cybernetics. Vol. 59. Iss. 4 p. 257–263. DOI: 10.1007/BF00332914.

WANG W., DING J. 2003. Wavelet network model and its application to the prediction of hydrology. Nature and Science. Vol. 1. Iss. 1 p. 67–71.

WU C., CHAU K., LI Y. 2009. Methods to improve neural network performance in daily flows prediction. Journal of Hydrology. Vol. 372. Iss. 1–4 p. 80–93. DOI: 10.1016/j.jhydrol.2009.03.038.

ZHOU S.L., MCMAHON T.A., WALTON A., LEWIS J. 2002. Forecasting operational demand for an urban water supply zone. Journal of Hydrology. Vol. 259. Iss. 1–4 p. 189–202. DOI: 10.1016/S0022-1694(01)00582-0.

Mukesh TIWARI, Jan ADAMOWSKI, Kazimierz ADAMOWSKI

Przewidywanie zapotrzebowania na wodę z użyciem technik uczenia maszynowego

STRESZCZENIE

Słowa kluczowe: *bootstrap, ekstremalne maszyny uczące się, falki, Kanada, niepewność, prognozowanie zapotrzebowania na wodę, sztuczne sieci neuronowe*

Oceniono zdolność modelowania z użyciem ekstremalnej maszyny uczącej się (ELM) stosowanej samodzielnie bądź w połączeniu z analizą falkową (W) lub metodami bootstrapowymi (B) (tzn. ELM, ELM_W, ELM_B) do przewidywania dobowego zapotrzebowania na wodę w mieście. Wyniki porównano z uzyskanymi tradycyjnymi metodami bazującymi na sztucznych sieciach neuronowych (tzn. ANN, ANN_W, ANN_B). Modele przewidujące zapotrzebowanie na wodę zbudowano z wykorzystaniem trzyletniego zapotrzebowania na wodę i zestawu danych klimatycznych dla miasta Calgary w kanadyjskiej prowincji Alberta. Hybrydowe modele ELM_B i ANN_B zapewniały satysfakcjonujące prognozy jednodniowe o podobnej dokładności, natomiast wyniki uzyskane z zastosowaniem modeli ELM_W i ANN_W były bardziej dokładne, przy czym model ELM_W okazał się lepszy niż ANN_W. Istotną poprawę prognozowania szczytowego zapotrzebowania na wodę w mieście uzyskano jedynie z zastosowaniem modelu ELM_W. Wyższość modelu ELM_W nad modelami ANN_W czy ANN_B dowodzi znaczącej roli transformacji falkowej w usprawnianiu działania modeli prognozujących zapotrzebowanie na wodę w mieście.