  Water Supply | in press | 2021

Short-term water demand predictions coupling an Artificial Neural Network model and a Genetic Algorithm

Majid Gholami Shirkoohi, Mouna Doghri and Sophie Duchesne

ABSTRACT

The application of artificial neural network (ANN) models for short-term (15 min) urban water demand predictions is evaluated. Optimization of the ANN model's hyperparameters with a Genetic Algorithm (GA) and use of a growing window approach for training the model are also evaluated. The results are compared to those of commonly used time series models, namely the Autoregressive Integrated Moving Average (ARIMA) model and a pattern-based model. The evaluations are based on data sets from two Canadian cities, providing 15 minute water consumption records over respectively 5 years and 23 months, with a respective mean water demand of 14,560 and 887 m^3/d. The GA optimized ANN model performed better than the other models, with Nash-Sutcliffe Efficiencies of 0.91 and 0.83, and Relative Root Mean Square Errors of 6 and 16% for City 1 and City 2, respectively. The results of this study indicate that the optimization of the hyperparameters of an ANN model can lead to better 15 min urban water demand predictions, which are useful for many real time control applications, such as dynamic pressure control.

Key words | artificial intelligence, artificial neural network, genetic algorithm, hyperparameter optimization, short term, time series model, urban water demand prediction

Majid Gholami Shirkoohi
Mouna Doghri†
Sophie Duchesne (corresponding author)
Research Centre on Water, Earth, and the Environment,
Institut National de la Recherche Scientifique (INRS),
490 rue de La Couronne, Quebec City, QC, Canada
G1 K 9A9
E-mail: *sophie.duchesne@ete.inrs.ca*

†Current address: Environment and Climate Change Canada, 1550 Avenue d'Estimauville, Quebec City, QC, Canada, G1 J 0C3.

HIGHLIGHTS

- ANN models were used for short-term (15 min) urban water demand predictions.
- The hyperparameters of the ANN model were optimized with a genetic algorithm for better model performance.
- The results of the ANN approach were compared to an ARIMA and a pattern-based models for two different datasets.
- The performance results proved GA optimized ANN model as an efficient approach for short-term UWD predictions.

INTRODUCTION

Forecasting urban water demand (UWD) is a crucial issue to ensure the better design, operation, and management of water distribution systems (WDSs). While long-term forecasting is mainly required for planning and design, short-term forecasting is particularly used for operation and management. More specifically, in the most recent applications of real-time control (RTC), knowledge of the near future fluctuations in consumption is required (e.g. Pascual *et al.* 2013; Kang 2014; Doghri *et al.* 2020).

The UWD is a complex and nonlinear function of different factors such as time, socio-economic factors, climatic

doi: 10.2166/ws.2021.049

and meteorological variables, and the cost of the supplied water (Ghiassi et al. 2008; Odan & Reis 2012; Hussien et al. 2016). UWD prediction models can be univariate, where only the UWD records are considered as inputs for the models (e.g. Alvisi et al. 2007; Alvisi & Franchini 2014; Romano & Kapelan 2014). Aditionally, the models can be multivariate, with some of the influencing factors mentioned above also being considered as model inputs to predict the UWD (e.g. Zhou et al. 2002; Herrera et al. 2010; Adamowski et al. 2012; Odan & Reis 2012; Tiwari & Adamowski 2013; Bakker et al. 2014; Tian et al. 2016). Both univariate and multivariate models can be based on various modeling approaches, from which the more commonly applied are pattern-based models (e.g. Alvisi et al. 2007; Bakker et al. 2013, 2014; Gagliardi et al. 2017; Pacchin et al. 2017, 2019), regression analyses (e.g. Adamowski & Karapataki 2010; Bakker et al. 2014), classical time series models (e.g. Bougadis et al. 2005; Caiado 2010; Arandia et al. 2015; Chen & Boccelli 2018; Viccione et al. 2019; Guarnaccia et al. 2020; Mu et al. 2020), artificial intelligence (AI) methods, such as artificial neural networks (ANN) (e.g. Cutore et al. 2008; Firat et al. 2010; Romano & Kapelan 2014; Gagliardi et al. 2017; Pacchin et al. 2019; Pesantez et al. 2020), random forest (e.g. Herrera et al. 2010; Mu et al. 2020; Nasser et al. 2020; Pesantez et al. 2020), support vector machines (SVM) (e.g. Candelieri et al. 2019; Pesantez et al. 2020), or fuzzy logic and neuro-fuzzy models (e.g. Vijayalaksmi & Babu 2015; Jithish & Sankaran 2017). More recently, hybrid models combining two of the previously cited approaches have also been proposed (e.g. Quevedo et al. 2010; Alvisi & Franchini 2014; Suhartono et al. 2018). While many factors impact the performance of UWD prediction models, Sebri (2016) found that forecasting accuracy of urban water demand is significantly influenced by demand periodicity, forecast horizon, forecasting method, model specification and some study specific characteristics such as the sample size, the publication year and the development level of the country on which the study was conducted. House-Peters & Chang (2011) and Donkor et al. (2014) presented an overview of the different existing models adapted to UWD forecasts up to 2010, while Ghalehkhondabi et al. (2017) published a review of the papers concerning specifically soft computing methods applied to UWD from 2005 to 2015. It was reported that using many explanatory variables for multivariate UWD prediction models poses a great challenge in terms of collecting and keeping track of the data since explanatory variables must have sufficiently long records if they are to be used as independent variables for developing forecasting models (House-Peters & Chang 2011; Donkor et al. 2014). Ghalehkhondabi et al. (2017) found that although it is still very difficult to pick a single method as the overall best, ANNs have been superior in many cases in short-term UWD forecasting. This can be derived from the inherent capability of ANNs in terms of analyzing the non-linear data. Also, limited number of applications of metaheuristics (such as evolutionary algorithms) in water demand forecasting was identified as one of the incentives for potential future research direction (Ghalehkhondabi et al. 2017).

Pattern-based models rely on the identification of periodic patterns that characterize UWD over different periods of time, while classical time series models do not necessarily take this periodicity into account explicitly. The most popular univariate classical time series models are the autoregressive moving average (ARMA) model and its derivatives, such as autoregressive (AR), autoregressive integrated moving average (ARIMA), seasonal ARIMA, periodic ARMA, threshold AR, and fractionally integrated ARMA models (Adamowski & Karapataki 2010). As for AI models, they do not presume a specific model structure and are thus said to 'learn' from the data. Among those AI models, ANN models have been used in numerous studies to predict UWD during the last decade (see some examples above). ANN models have proven to be powerful for mapping the nonlinear trends of UWD, even in the case of possibly noisy multivariate time series (Ghalehkhondabi et al. 2017). Studies carried out in this field highlight the dominance of ANN over conventional techniques (Babel & Shinde 2011). Bougadis et al. (2005) showed that ANN models outperforms different regression and time series models for short-term peak water demand forecasting for data from the city of Ottawa, Canada. Similar results were reported by Adamowski & Karapataki (2010) and Adamowski et al. (2012) in favor of ANNs and wavelet transforms coupled with ANNs, compared to other conventional techniques for short-term water demand forecasts. The performance of ANN models is dependent on the set of explanatory variables fed into the model as input and on the forecasting horizon. Prediction accuracies over 98% were achieved by Babel & Shinde (2011) when using only the historic daily demand as the explanatory variable to forecast short-term UWD for the city of Bangkok, Thailand. However, they showed that meteorological, water utility and socioeconomic variables have a greater influence on medium-term (e.g. monthly) predictions. The benefit of univariate models was also reported by Odan & Reis (2012) where their ANN models for short-term UWD prediction did not require the use of weather variables, resulting in a simpler and faster model to train. Also, size of the data sets of case studies can affect the ANN model performance. As Gagliardi *et al.* (2017) showed, an ANN model applied to small districts, with a low number of users and more variability in water demands, can outperform a pattern-based model while in case of districts including a large number of users, the pattern-based model tends to be more efficient than the ANN one.

Evolutionary algorithms (Bäck 1996) have been used along with AI techniques for UWD prediction, either for the optimization of training algorithms (Rangel *et al.* 2016) or optimization of model hyperparameters (Chen 2009). Romano & Kapelan (2014) suggested to optimise the hyperparameters (e.g., number of hidden neurons and number of training cycles) and input structure (e.g. number of past demand values and additional explanatory variables to be used) of their ANN model with an evolutionary algorithm for the prediction of UWD 1 to 24 h ahead, and showed that this approach allows accurate forecasts of UWD. In their application of ANNs for the prediction of hourly UWD, Herrera *et al.* (2010) compared the growing and sliding window approaches to train their ANN model. For their case study, they found that the growing window approach lead to better results.

In terms of UWD, the definition of short-term predictions vary among authors. Although many of them consider predictions 1 h ahead as being short-term (e.g. Shvartser *et al.* 1993; Zhou *et al.* 2002), others qualify the daily, weekly or even monthly predictions as being short term (e.g. Jain & Ormsbee 2002; Bougadis *et al.* 2005; Yurdusev & Firat 2009). The forecast horizon depends on the purpose of the model application (Ghiassi *et al.* 2008; Bakker *et al.* 2013). More specifically for real time control applications, such as active pressure management, demand predictions should be provided a few times an hour (e.g. Creaco 2017; Doghri *et al.* 2020).

Few applications of UWD predictions at time steps lower that 1 h have been presented in the literature: ANN, support vector regression (SVR) and random forest for 10 min predictions in Nasser *et al.* (2020); a pattern-based model for 15 min predictions in Bakker *et al.* (2013); a seasonal ARIMA model for 15 min predictions in Arandia *et al.* (2015); and a pattern-based model combined to an ARIMA model for 10 min predictions in Quevedo *et al.* (2010). A few studies have compared the performances of prediction models on making short-term UWD forecasts (daily in Adamowski *et al.* 2012; Tiwari & Adamowski 2013; Bai *et al.* 2014; and hourly in Ghiassi *et al.* 2008; Herrera *et al.* 2010). A thorough search of the scientific literature showed that only two studies comparing the performances of different UWD prediction models for time steps lower than 1 h have already been published (Mu *et al.* 2020; Nasser *et al.* 2020). Mu *et al.* (2020) used a long short-term memory (LSTM) model to predict short-term UWD based on data with time steps of 15 min, 1 and 24 h. The performance of the LSTM-based model was compared with ARIMA, SVR, and random forest models. Nasser *et al.* (2020) compared the performance of different models for 10 min UWD predictions. However, this last study was limited to a few households (granular water demand) and only compares AI models among them. Short-term (10- or 15-min) time steps describe the UWD dynamics in details and can be required for some real time applications such and optimizing the exact timing of pumps switch (Bakker *et al.* 2013) or for dynamic pressure management (Creaco 2017; Doghri *et al.* 2020). These can help providing the required amount and pressure of water to urban areas at the lowest operation cost. Furthermore, a short-term 15-min time step UWD prediction can be used for leakage and energy analysis (Mu *et al.* 2020).

The main objective of this paper is to evaluate how the optimization of the hyperparameters of an ANN model and the use of a growing window approach for its training can improve the 15 min UWD predictions made with this model. The performance of the ANN model is also compared with those of a classical time series model (ARIMA) and of a pattern-based model (Bakker *et al.* 2013). To the best of the authors' knowledge, it is the first time that the performance of these three types of models in making predictions of UWD for time steps shorter than one hour is compared. Data sets provided by two Canadian cities are used for these evaluations.

METHODS

Datasets

The datasets analyzed in this study were collected from two cities in the province of Quebec (Canada). The data consisted of 15 min time step records of drinking UWD. The datasets were divided into training, validation and testing data for the ANN model. The training data were used to train the model and update the parameters. The validation data were used to select model parameters and stop the algorithm early based on minimum validation error which efficiently avoids model overfitting or underfitting. The testing data were used to evaluate the performance of the ANN model to the unseen data. The same testing data was used to evaluate the performance of the pattern-based model.

For the dataset from City 1, the records present the total drinking water produced in the treatment plant for a period of five years (2009–2013). The first four years of observation were used as training and validation samples for the ANN models (from January 1st, 2009, 00:00 AM to August 3rd, 2012, 17:45 PM for training and from August 3rd, 2012, 18:00 PM to December 31st, 2012, 23:45 PM for validation), and the remaining year (from January 1st, 2013, 00:00 AM to December 31st, 2013, 11:45 PM) was used as a test set to assess the accuracy of all prediction models (ANN, ARIMA and pattern-based). The average recorded consumption for this dataset is about 14,560 m^3/d with a standard deviation of about 3,090 m^3/d for 15 min time steps. The dataset from City 2 presents the total water provided to a district metered area and covers a period of 23 months. The first period of observations was used as a training and validation set (from September 1st, 2012, 00:00 AM to November 1st, 2013, 23:45 PM for training and from November 2nd, 2013, 00:00 AM to December 31st, 2013, 23:45 PM for validation) for the ANN models and the remaining 7 months (from January 1st, 2014, 00:00 AM to July 21st, 2014, 23:15 PM) were used as a test set to assess the accuracy of all prediction models (ANN, ARIMA and pattern-based). The average consumption, in this case, is about 887 m^3/d with a standard deviation of about 311 m^3/d for 15 min time steps. The time series for the two cities are shown in Figure 1 where training, validation and test sets are identified for each city, while their mean daily patterns are shown in Figure 2.

To remove the outliers from both databases, the *filloutliers* function in MATLAB (ver. R2019a) has been used, which defines outliers as points outside three standard deviations from the mean and replace the outlier with the nearest element that is not an outlier. Finally, 128,167 data points were used as training/validation sets and 35,028 as test sets for City 1. Whereas for City 2, a total of 66,145 data points were divided into 46,752 and 19,393 data points for training/validation and test sets, respectively. As illustrated in Figure 1, there were about 1,200 time steps in the dataset for City 2 with constant values (from time step 8,800 to time step 10,000). As these data were used only for the training of the ANN models and rather negligible compared to the 46,752 data points, they were not excluded from the data set. Also, our preliminary results indicated that performance indices for the ANN models are not dependent on the presence or exclusion of this part of the data set.

ANN model

Artificial neural networks

ANNs have been proven to have excellent predictive ability in various domains (Sun & Zhang 2002; Nastos *et al.* 2013; Mislan *et al.* 2015; Cheng & Bao 2020).

One of the most popular ANNs in the scope of UWD prediction are feedforward neural networks (FNNs) (Hamed *et al.* 2004). The layers of the ANN can be fully or partially connected and for the purpose of forecasting, the weights should be adapted accordingly in a process called training.

A number of optimization algorithms can be used for the training process including gradient descent and Levenberg-Marquardt (Hamed *et al.* 2004; Ghalehkhondabi *et al.* 2017). Levenberg-Marquardt algorithm is often utilized for multilayer perceptron neural networks due to its faster convergence as it adopts the method of approximate second derivative (Singh *et al.* 2010).

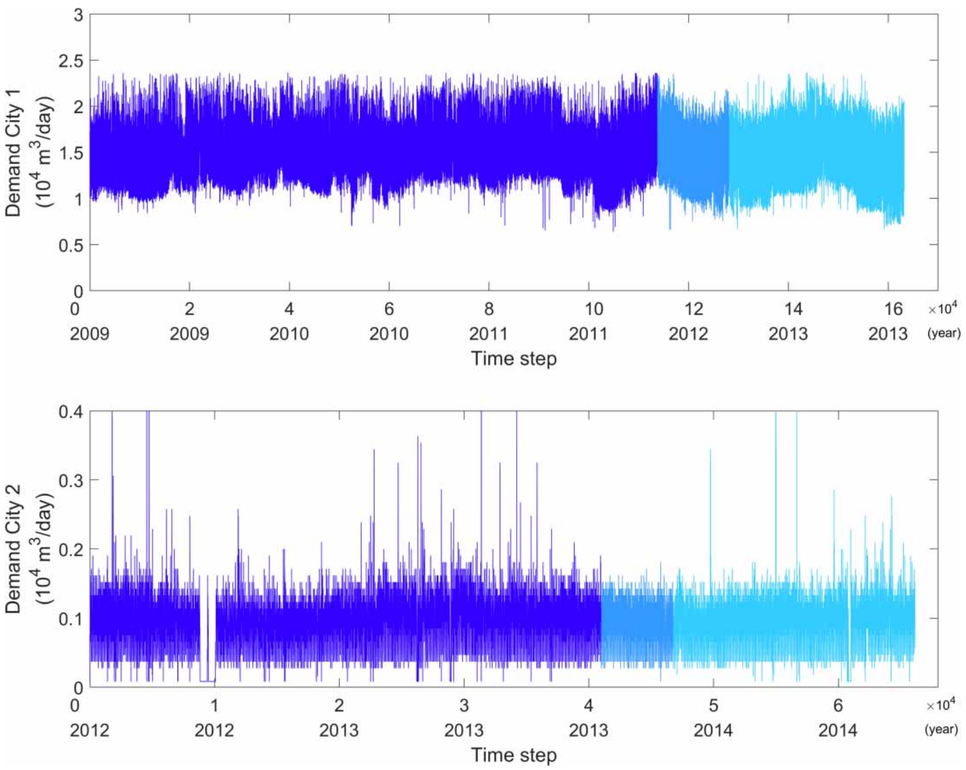

Figure 1 | Time series of the two datasets (training, validation and test subsets represesented respectively in dark blue, medium dark blue and light blue).

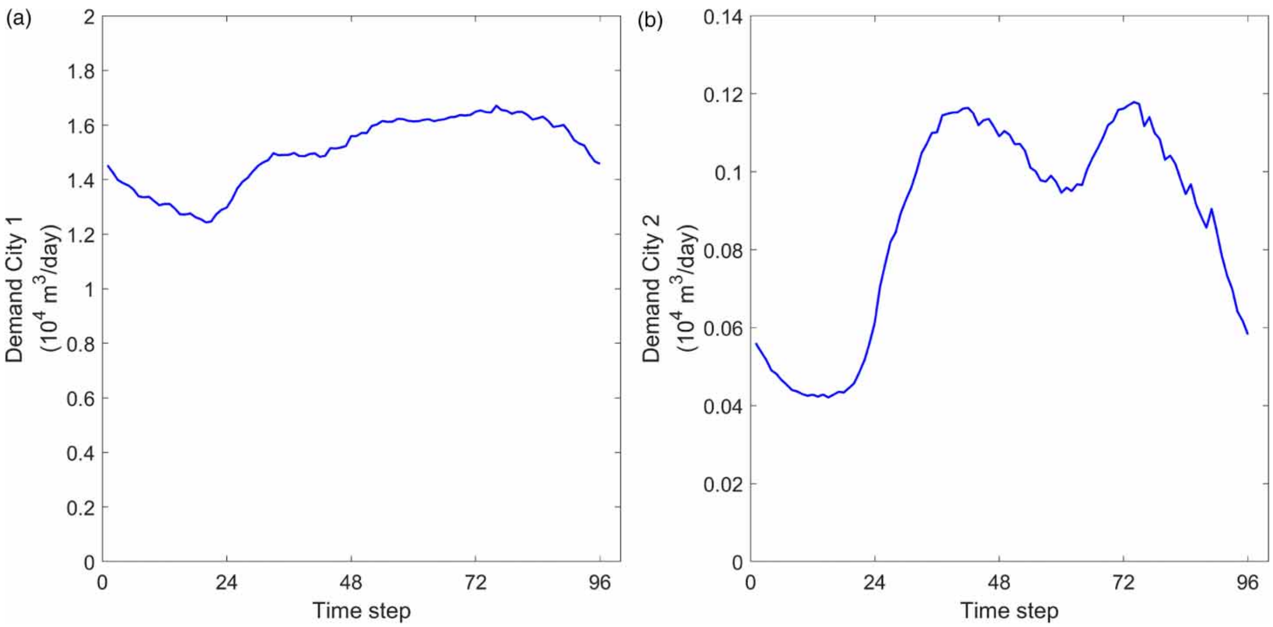

Figure 2 | Mean daily pattern of water consumption for the two data sets, (a) City 1; (b) City 2.

In this work, three ANN models are developed. In all cases, datasets were decomposed to the trend and cyclical components, and were used as inputs to the ANN models. For each time step prediction, the two last time steps were utilized. The first model is based on initially performing one time training, using the entire training and validation data set, and then obtaining the performance results on the whole test data set with the trained network (Single-ANN). The second model is considering, for each time step in the test data set, all the data before that time step as a training and validation data set to predict that specific time step. In other words, the growing window concept is utilized in which all the data from the first to current time steps are used to train a network for predicting the next 15 min time step (Multi-ANN). Finally, as a third approach, genetic algorithm (GA) is used to optimize the ANN hyperparameters.

Indeed, the performance of an ANN model depends on the optimization of its hyperparameters, which define the topology and learning options of a neural network. The number of hidden layers and neurons in each hidden layer, learning rate, cost function, regularization parameter, learning algorithm and maximum validation failure are considered as ANN hyperparameters. In the third approach presented here, a GA is utilized for the hyperparameter optimization of the neural networks.

Genetic algorithm

Genetic algorithm is a strategy of evolutionary computation search algorithms which states that individuals in a population who are best fitted are more likely to survive and reproduce.

In GA, a chromosome is a set of parameters which define a proposed solution to the problem that the GA is trying to solve by searching through the space of possible chromosome values. In this work, the hyperparameters of a network are the chromosome of one individual. The major steps of a GA are described in Whitley (1994).

In the case studies presented here, the three hyperparameters (genes) of chromosomes and their range of values are (i) the number of hidden neurons (1 to 20), (ii) the LM algorithm parameter (μ) (0 to 1), and (iii) the maximum validation failures (1 to 20).

The relative root mean square error (RRMSE) is used as the fitness function, and three selection operators are employed to select the most fitted individuals as the first and second parents to go through the crossover, namely Roulette wheel, Tournament, and Random selection (Zhong *et al.* 2005). A crossover percentage of 0.8 and a mutation percentage of 0.3 are used. The number of individuals in an initial population is set to 20. Also, the maximum number of iterations is used as stoppage criterion for GA optimization and is set to 50. The flowchart of the novel ANN model with GA optimization for network hyperparameters is presented in Figure 3.

ARIMA MODEL

ARIMA model (Box & Jenkins 1976) is a commonly and widely used model to make forecasts for a large range of time steps. It showed satisfactory results in UWD applications (e.g. Bougadis *et al.* 2005; Ghiassi *et al.* 2008; Caiado 2010; Tiwari & Adamowski 2013; Chen & Boccelli 2018). ARIMA models require the input data to have

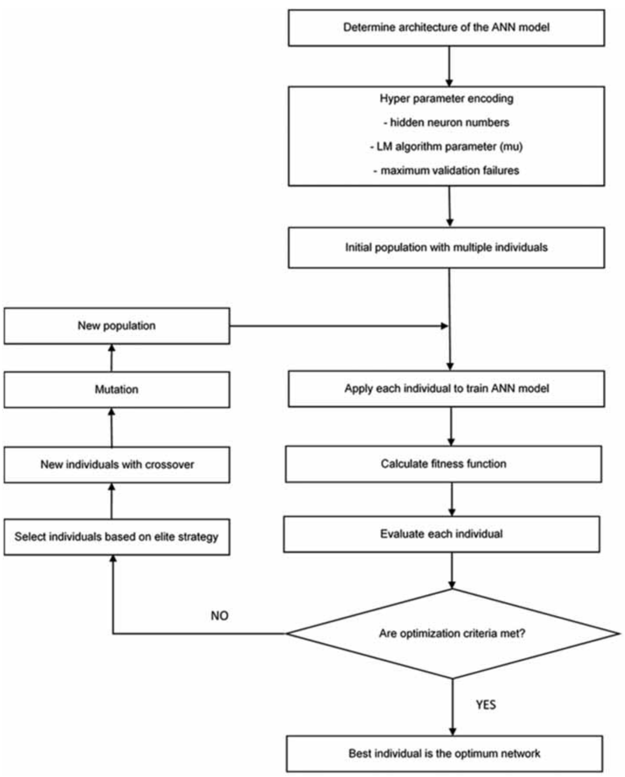

Figure 3 | Flowchart of ANN model with GA optimization for network hyperparameters.

constant mean, variance, and autocorrelation through time (Box & Jenkins 1976). They allow treating a non-stationary series by the elimination of the trend through successive differentiations of the time series data. In the case of water consumption records, one differentiation of the dataset is generally enough to satisfy this stationarity condition. The model is defined as follows:

$$C_t = C_{t-1} + \sum_{j=1}^{p_1} (\gamma_j W_{j-1}) - \sum_{j=1}^{p_2} (\theta_j \varepsilon_{t-j}) + \varepsilon_t \quad (5)$$

where: C_t is the observed value of the time series at time step t; the first sum-term represents an autoregressive model (AR) of order p_1; the second sum-term represents a moving average model (MA) of order p_2; W is the differentiated series; γ_j and θ_j are the parameters of the AR and the MA models to be calibrated, respectively; and ε_t is a random perturbation or white noise. The model is referred to as an ARIMA(p_1, d,p_2), where d represents the order of differentiation of the original dataset, and the values of the parameters p_1 and p_2 are estimated following the pre-analysis of the dataset.

The autocorrelation function (ACF) and the partial autocorrelation function (PACF), as defined by Box & Jenkins (1976), were used to identify the most appropriate time series model for the dataset (Yang *et al.* 2013; Arandia *et al.* 2015). For both case studies presented in this paper, the ACFs decayed slowly with increasing time lags. The PACFs showed a large spike in the first lags and cutoff to 0 after lags 25 and 17 for Cities 1 and 2, respectively. The above observations suggested a non-stationary process for the two datasets. An example of the ACF for 672 lags (time step of data equal to 15 min) is presented in the Supplementary Material (Figure S-1) for the dataset from City 2, in which the periodic behaviour of the water consumption can be seen. With a cycle of positive and then negative values every 96 lags, the observations show the correlation between the data and exhibit the daily seasonality of the water consumption data.

Values of p_1, d and p_2 were determined by analysing the data of the training and validation sets. Through the differentiation process, the trend was removed from the autocorrelation functions and the datasets transformed into stationary series. Indeed, ACFs and PACFs of the differentiated series tend more quickly to reach a value near zero than those of the original series (see Figures S-2 to S-5 in the Supplementary Material). Various orders of models have been tested for the AR and MA processes, however, results are not presented herein for brevity. It was concluded that ARIMA(2,1,1) provides satisfying results for both case studies (City 1 and City 2) and was further adopted for the following studies. The test set used to evaluate the performance of the ANN model is the same as the one used for the other models (see section 2.5 below). Values of the γ_j and θ_j parameters were calibrated at each time step using the previous 192 data (i.e. a total of 2 days of data) with the *estimate* function in MATLAB, which is based on the maximum likelihood.

Pattern-based model

The model developed by Bakker *et al.* (2013) is the pattern-based model chosen for the applications presented here. This forecasting method combines the daily average estimation of the UWD with the demand pattern to provide 15 min predictions over the following 24 h (only the 15 min ahead predictions are presented and discussed in this paper). The model analyzes the historical data series to determine different factors ($f_{dotw,typ,i}$: typical day of the week factor and $f_{qtr,typ,i,j}$: typical 15 min time step factor). By exploring the available database, the method defines the specific factors for the seven ordinary days of the week and for each particular day of the year, and the 15 min demand pattern corresponding for each one of these days.

The model makes the prediction of the UWD for the next day with a 15 min time step. The main steps of the model are summarized in Equations (6) and (7). The method is as follows: (i) the model computes the value of the mean demand for the next day (Q_i) based on the mean water demands of the previous two days (Q_{i-1} and Q_{i-2}), divided by the corresponding typical day of the week factors and making the more recent day four times more important than the older demand (corresponding weighing constants set at 0.8 and 0.2); then (ii) the mean demand of day i is discretized in a set of 96 values, namely predictions for each 15 min time step over the next 24 h. The latter step was performed by the multiplication of the mean demand with the corresponding typical 15-min time step factors.

$$Q_i = f_{dotw,typ,i}\left(0.8\frac{Q_{i-1}}{f_{dotw,typ,i-1}} + 0.2\frac{Q_{i-2}}{f_{dotw,typ,i-2}}\right) \quad (6)$$

$$Q_{i,j} = Q_i * \mathrm{f_{qtr,typ,i.j}} \quad (7)$$

The model was coded using MATLAB 2014a software by considering, as most as possible, the steps and parameters described in Bakker et al. (2013) without any specific consideration of the sprinkle demand. The same method as the one described in Bakker et al. (2013) was applied for the factors calculation. The model will thereafter be called the fully adaptive forecasting (FAF) model. The test set used to evaluate the performance of the ANN model is the same as the one used for the other models (see Performance indicators below).

Performance indicators

The accuracies of the different models were evaluated using the following three statistical indices, namely the Relative Root Mean Square Error (RRMSE), the Nash-Sutcliffe Model Efficiency coefficient (E) and the Mean Absolute Percentage Error (MAPE) for the test sets of City 1 and City 2 (i.e. respectively from 1 January 2013, 00:00 AM to 31 December 2013, 11:45 PM, and from 1 January 2014, 00:00 AM to 21 July 2014, 23:15 PM; see Figure 1). The measures selected to compare the forecasted and measured values are the most commonly used by researchers addressing UWD forecasting (Adamowski et al. 2012; Bakker et al. 2013) and they all generate dimensionless outputs. The equations used to compute the values of these indices are given in Table 1, where N is the total number of forecasted values, C_t is the measured value at time t, $\hat{C}_t$ is the forecasted value at time t, and $\bar{C}$ is the mean of the measured values.

Table 1 | Performance indices used

Indicators	Mathematical formulation	Range of values	Values of perfect agreement
Relative Root Mean Square Error (RRMSE)	$\frac{\sqrt{\frac{1}{N}\sum_{t=1}^{N}(C_t - \hat{C}_t)^2}}{\bar{C}}100\%$	$[0, +\infty[$	0
Mean Absolute Percentage Error (MAPE)	$\frac{100}{N}\sum_{t=1}^{N}\left\lvert\frac{C_t - \hat{C}_t}{C_t}\right\rvert$	$[0, +\infty[$	0
Nash-Sutcliffe Efficiency (E)	$1 - \frac{\sum_{t=1}^{N}(C_t - \hat{C}_t)^2}{\sum_{t=1}^{N}(C_t - \bar{C})^2}$	$]-\infty, 1]$	1

Table 2 | Summary of results

Model	RRMSE (%)	MAPE (%)	E
City 1			
Multiple ANN training	6.46	4.29	0.91
Single ANN training	6.48	4.28	0.91
GA optimized ANN	**6.35**	**4.15**	**0.91**
ARIMA	7.76	5.20	0.87
FAF	13.25	8.94	0.61
City 2			
Model	**RRMSE (%)**	**MAPE (%)**	**E**
Multiple ANN training	19.00	14.49	0.77
Single ANN training	17.77	14.16	0.80
GA optimized ANN	**16.23**	**13.27**	**0.83**
ARIMA	20.09	14.88	0.74
FAF	23.92	14.79	0.63

RESULTS AND DISCUSSION

The performance indices (RRMSE, MAPE and E) of the studied models for the test sets of the two case studies are presented in Table 2. Examples of results are illustrated for two specific days in Figure 4 for City 1 and City 2.

The ANN hyperparameter values obtained by GA optimization were as follows: number of hidden neurons = 18, LM parameter $\mu = 0.4632$, and maximum validation failures = 9 for City 1; and number of hidden neurons = 16, LM parameter $\mu = 0.4307$, and maximum validation failures = 3 for City 2. As can be seen in Table 3, the

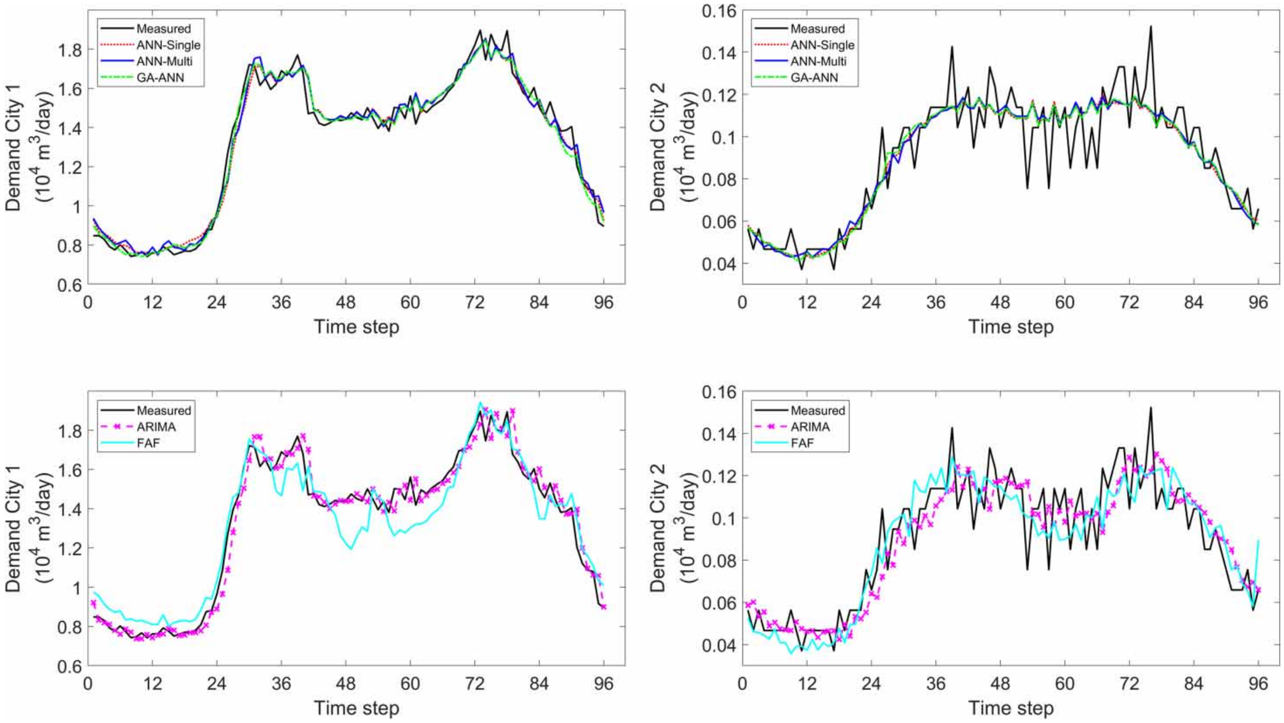

Figure 4 | Forecasts versus observations of different models (one day sample from test set).

ANN models provide better 15 min UWD predictions (i.e. lower RRMSE, lower MAPE, and higher E values) than the ARIMA and FAF models for both case studies. Moreover, optimization of the hyperparameters of ANN with GA allows for refining of the predictions. It is also worth mentioning that the statistical indicators for City 1 are always better than those for City 2. It can be thought, as suggested by previous studies (e.g. Maidment & Miaou 1986; Bakker *et al.* 2013; Gagliardi *et al.* 2017), that the performance of UWD predictions vary depending on the size of the water distribution area. Indeed, when the size of the area increases the total consumption also increases, and its fluctuations are generally mitigated (see Figure 2), making future values easier to predict. It can additionally be observed that that the multiple ANN training approach has led to better results as compared to the single ANN training for City 1 than City 2. Again, it seems that the size of the case study has a direct impact on the performance of the models. However, the GA optimized ANN showed the best performance in both cities compared to other models and proves its reliability in different circumstances.

The presented results show that ANN models outperform the classical time series and pattern-based models in forecasting short-term UWD. Results in Figure 4 also show that the ARIMA model provides predictions that are delayed from the measured data. This is typical of low order ARIMA models, in which the projections represent a weighting of the last observations. Better performance of ANN models over time series and linear and nonlinear regression methods have also been presented by Jain *et al.* (2001) and Adamowski *et al.* (2012). There are varying degrees of nonlinearity in UWD data that make them difficult to be handled by linear methods. The strengths of ANN models can rely on their inherent ability to capture the nonlinearities related to the UWD time series.

Another finding of this study was that hyperparameter optimization of the ANN model could enhance its prediction performance. This supports the findings of Romano & Kapelan (2014) for the prediction of UWD 1 to 24 h ahead. These authors reported Nash-Sutcliffe Efficiencies higher than 0.9 for their adaptive ANN models for both daily and hourly forecasting. Their optimization procedure included six decision variables, namely the number of hidden neurons, the number of training cycles, the training algorithm regularization factor, the lag size, and the use of not of the time of the day and the day of the week.

CONCLUSIONS AND PERSPECTIVES

The 15 min UWD predictions obtained by different models were compared in this paper, based on data collected from two different Canadian water distribution systems. Considering the intention of developing real-time control tools: i) the models that were compared were exclusively univariate time series models, using only the records of previous UWD as input data to predict the future demand, and ii) the comparison of the models was based on their ability to provide short-term forecasts of UWD. An original model combining ANN and GA, for the optimization of the ANN model hyperparameters, was proposed, showing its superiority in providing more accurate 15 min UWD predictions than ARIMA and a pattern-based models. However, although the ANN model provided better 15 min UWD predictions than the ARIMA and the tested pattern-based models for the two presented case studies, many authors showed that pattern-based models provide better predictions for longer lead times (e.g. from about 3 h or more in Doghri 2019). Moreover, the efficiency of the GA optimized ANN model should be verified with other consumption datasets and different forecast horizons, in order to validate the obtained results and to generalize these findings. Finally, since coupled wavelet-neural network models (WA-ANNs) have shown good potential to predict UWD (Adamowski *et al.* 2012), a comparative study between the GA optimized ANN and WA-ANN models for UWD prediction could be useful.

ACKNOWLEDGEMENTS

The data used in this paper were provided by two cities from the province of Quebec, Canada, which are not named here for confidentiality reasons, but are gratefully acknowledged. Financial support to Majid Gholami Shirkoohi and to Mouna Doghri was provided by the National Sciences and Engineering Research Council of Canada (Gholami and Doghri) and the CREATE TEDGIEER program (Gholami).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Adamowski, J. & Karapataki, C. 2010 Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. *Journal of Hydrologic Engineering* **15** (10), 729–743. doi: 10.1061/(ASCE)HE.1943-5584.0000245.

Adamowski, J. F., Chan, H. F., Prasher, S. O., Ozga-Zielinski, B. & Sliusarieva, A. 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 Resources Research* **48** (1), W01528. doi: 10.1029/2010wr009945.

Alvisi, S. & Franchini, M. 2014 Assessment of the predictive uncertainty within the framework of water demand forecasting by using the model conditional processor. *Procedia Engineering* **89**, 893–900. doi: 10.1016/j.proeng.2014.11.522.

Alvisi, S., Franchini, M. & Marinelli, A. 2007 A short-term, pattern-based model for water-demand forecasting. *Journal of Hydroinformatics* **9** (1), 39–50. doi: 10.2166/hydro.2006.016.

Arandia, E., Ba, A., Eck, B. & McKenna, S. 2015 Tailoring seasonal time series models to forecast short-term water demand. *Journal of Water Resources Planning and Management* **142** (3), 04015067. doi: 10.1061/(ASCE)WR.1943-5452.0000591.

Babel, M. S. & Shinde, V. R. 2011 Identifying prominent explanatory variables for water demand prediction using Artificial Neural Networks: a case study of Bangkok. *Water Resources Management* **25** (6), 1653–1676. doi: 10.1007/s11269-010-9766-x.

Bäck, T. 1996 *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press, New York, NY.

Bai, Y., Wang, P., Li, C., Xie, J. & Wang, Y. 2014 A multi-scale relevance vector regression approach for daily urban water demand forecasting. *Journal of Hydrology* **517**, 236–245. doi: 10.1016/j.jhydrol.2014.05.033.

Bakker, M., Vreeburg, J. H. G., van Schagen, K. M. & Rietveld, L. C. 2013 A fully adaptive forecasting model for short-term drinking water demand. *Environmental Modelling & Software* **48**, 141–151. doi: 10.1016/j.envsoft.2013.06.012.

Bakker, M., van Duist, H., van Schagen, K., Vreeburg, J. & Rietveld, L. 2014 Improving the performance of water demand forecasting models by using weather input. *Procedia Engineering* **70**, 93–102. doi: 10.1016/j.proeng.2014.02.012.

Bougadis, J., Adamowski, K. & Diduch, R. 2005 Short-term municipal water demand forecasting. *Hydrological Processes* **19** (1), 137–148. doi: 10.1002/hyp.5763.

Box, G. E. & Jenkins, G. M. 1976 *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, CA.

Caiado, J. 2010 Performance of combined double seasonal univariate time series models for forecasting water demand. *Journal of Hydrologic Engineering* **15**, 215–222. doi: 10.1061/(ASCE)HE.1943-5584.0000182.

Candelieri, A., Giordani, I., Archetti, F., Barkalov, K., Meyerov, I., Polovinkin, A., Sysoyev, A. & Zolotykh, N. 2019 Tuning hyperparameters of a SVM-based water demand forecasting system through parallel global optimization. *Computers & Operations Research* **106**, 202–209. doi: 10.1016/j.cor.2018.01.013.

Chen, X. 2009 Prediction of urban water demand based on GA-SVM. In: *FCC '09: Proceedings of the 2009 ETP International Conference on Future Computer and Communication*. IEEE Computer Society, Washington, DC, pp. 285–288. doi:10.1109/FCC.2009.82.

Chen, J. & Boccelli, D. L. 2018 Forecasting hourly water demands with seasonal autoregressive models for real-time application. *Water Resources Research* **54** (2), 879–894. doi:10.1002/2017WR022007.

Cheng, L. & Bao, Y. 2020 Short-term power load forecasting based on empirical mode decomposition and deep neural network. In: *Proceedings of Purple Mountain Forum 2019 – International Forum on Smart Grid Protection and Control. Lecture Notes in Electrical Engineering*, Vol. 585 (Y. Xue, Y. Zheng & S. Rahman eds). Springer, Singapore. doi: 10.1007/978-981-13-9783-7_62.

Creaco, E. 2017 Exploring numerically the benefits of water discharge prediction for the remote RTC of WDNs. *Water* **9** (12), 961. doi: 10.3390/w9120961.

Cutore, P., Campisano, A., Kapelan, Z., Modica, C. & Savic, D. 2008 Probabilistic prediction of urban water consumption using the SCEM-UA algorithm. *Urban Water Journal* **5** (2), 125–132. doi: 10.1080/15730620701754434.

Doghri, M. 2019 *Gestion de la pression dans les réseaux de distribution d'eau potable en vue de la réduction des fuites : évaluation dans le contexte des réseaux nord-américains (Pressure Management in Drinking Water Distribution Networks to Reduce Leakage: Assessment in the Context of North American Networks). PhD thesis*, Institut national de la recherche scientifique, Quebec, Canada.

Doghri, M., Duchesne, S., Poulin, A. & Villeneuve, J.-P. 2020 Comparative study of pressure control modes impact on water distribution system performance. *Water Resources Management* **34** (1), 231–244. doi: 10.1007/s11269-019-02436-z.

Donkor, E. A., Mazzuchi, T. A., Soyer, R. & Roberson, J. A. 2014 Urban water demand forecasting: review of methods and models. *Journal of Water Resources Planning and Management* **140** (2), 146–159. doi: 10.1061/(ASCE)WR.1943-5452.0000314.

Firat, M., Turan, M. E. & Yurdusev, M. A. 2010 Comparative analysis of neural network techniques for predicting water consumption time series. *Journal of Hydrology* **384**, 46–51. doi: 10.1016/j.jhydrol.2010.01.005.

Gagliardi, F., Alvisi, S., Franchini, M. & Guidorzi, M. 2017 A comparison between pattern-based and neural network short-term water demand forecasting models. *Water Science and Technology – Water Supply* **17** (5), 1426–1435. doi: 10.2166/ws.2017.045.

Ghalehkhondabi, I., Ardjmand, E., Young II, W. A. & Weckman, G. A. 2017 Water demand forecasting: review of soft computing methods. *Environmental Monitoring and Assessment* **189** (7), 313. doi: 10.1007/s10661-017-6030-3.

Ghiassi, M., Zimbra, D. & Saidane, H. 2008 Urban water demand forecasting with a dynamic artificial neural network model. *Journal of Water Resources Planning and Management* **134** (2), 138–146. doi: 10.1061/(ASCE)0733-9496(2008)134:2(138).

Guarnaccia, C., Tepedino, C., Viccione, G. & Quartieri, J. 2020 Short-term forecasting of tank water levels serving urban water distribution networks with ARIMA models. In: *Frontiers in Water-Energy-Nexus – Nature-Based Solutions, Advanced Technologies and Best Practices for Environmental Sustainability. Advances in Science, Technology & Innovation (IEREK Interdisciplinary Series for Sustainable Development)* (V. Naddeo, M. Balakrishnan & K. H. Choo eds). Springer, Cham. doi:10.1007/978-3-030-13068-8_6.

Hamed, M. M., Khalafallah, M. G. & Hassanien, E. A. 2004 Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling & Software* **19** (10), 919–928. doi: 10.1016/j.envsoft.2003.10.005.

Herrera, M., Torgo, L., Izquierdo, J. & Pérez-García, R. 2010 Predictive models for forecasting hourly urban water demand. *Journal of Hydrology* **387** (1–2), 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. *Water Resources Research* **47** (5), W05401. doi: 10.1029/2010WR009624.

Hussien, W. A., Memon, F. A. & Savic, D. A. 2016 Assessing and modelling the influence of household characteristics on per capita water consumption. *Water Resources Management* **30**, 2931–2955. doi:10.1007/s11269-016-1314-x.

Jain, A. & Ormsbee, L. 2002 Short-term water demand forecast modeling techniques conventional methods versus AI. *Journal of the American Water Works Association* **94** (7), 64–72. doi: 10.1002/j.1551-8833.2002.tb09507.x.

Jain, A., Kumar Varshne, A. & Chandra Joshi, U. 2001 Short-term water demand forecast modelling at IIT Kanpur using Artificial Neural Networks. *Water Resources Management* **15** (5), 299–321. doi: 10.1023/A:1014415503476.

Jithish, J. & Sankaran, S. 2017 A neuro-fuzzy approach for domestic water usage prediction. In IEEE Region 10 Symposium (TENSYMP 2017) – IEEE International Symposium on Technologies for Smart Cities, 14–16 July 2017, Kochi, Kerala, India, pp. 1–5. doi: 10.1109/TENCONSpring.2017.8070087.

Kang, D. 2014 Real-time optimal control of water distribution systems. *Procedia Engineering* **70**, 917–923. doi: 10.1016/j.proeng.2014.02.102.

Maidment, D. R. & Miaou, S.-P. 1986 Daily water use in nine cities. *Water Resources Research* **22** (6), 845–851. doi: 10.1029/WR022i006p00845.

Mislan, M., Haviluddina, H., Hardwinarto, S., Sumaryono, S. & Aipassa, M. 2015 Rainfall monthly prediction based on Artificial Neural Network: A case study in Tenggarong Station, East Kalimantan - Indonesia. *Procedia Computer Science* **59**, 142–151. doi: 10.1016/j.procs.2015.07.528.

Mu, L., Zheng, F., Tao, R., Zhang, Q. & Kapelan, Z. 2020 Hourly and daily urban water demand predictions. *Journal of Water Resources Planning and Management* **146** (9), 05020017. doi: 10.1061/(ASCE)WR.1943-5452.0001276.

Nasser, A. A., Rashad, M. Z. & Hussein, S. E. 2020 A two-layer water demand prediction system in urban areas based on micro-services and LSTM Neural Networks. *IEEE Access* **8**, 147647–147661. doi: 10.1109/ACCESS.2020.3015655.

Nastos, P. T., Moustris, K. P., Larissi, I. K. & Paliatsos, A. G. 2013 Rain intensity forecast using Artificial Neural Networks in Athens, Greece. *Atmospheric Research* **119**, 153–160. doi: 10.1016/j.atmosres.2011.07.020.

Odan, F. K. & Reis, L. F. R. 2012 Hybrid water demand forecasting model associating Artificial Neural Network with Fourier series. *Journal of Water Resources Planning and Management* **138** (3), 245–256. doi:10.1061/(ASCE)WR.1943-5452.0000177.

Pacchin, E., Alvisi, S. & Franchini, M. 2017 A short-term water demand forecasting model using a moving window on previously observed data. *Water* **9** (3), 172. doi: 10.3390/w9030172.

Pacchin, E., Gagliardi, F., Alvisi, S. & Franchini, M. 2019 A comparison of short-term water demand forecasting models. *Water Resources Management* **33** (4), 1481–1497. doi: 10.1007/s11269-019-02213-y.

Pascual, J., Romera, J., Puig, V., Cembrano, G., Creus, R. & Minoves, M. 2013 Operational predictive optimal control of Barcelona water transport network. *Control Engineering Practice* **21** (8), 1020–1034. doi: 10.1016/j.conengprac.2013.01.009.

Pesantez, J. E., Berglund, E. Z. & Kaza, N. 2020 Smart meters data for modeling and forecasting water demand at the user-level. *Environmental Modelling & Software* **125** (March), 104633. doi:10.1016/j.envsoft.2020.104633.

Quevedo, J., Puig, V., Cembrano, G. & Blanch, J. 2010 Validation and reconstruction of flow meter data in the Barcelona water distribution network. *Control Engineering Practice* **18** (6), 640–651. doi.org/10.1016/j.conengprac.2010.03.003.

Rangel, H. R., Puig, V., Farias, R. L. & Flores, J. J. 2016 Short-term demand forecast using a bank of neural network models trained using genetic algorithms for the optimal management of drinking water networks. *Journal of Hydroinformatics* **19** (1), 1–16. doi: 10.2166/hydro.2016.199.

Romano, M. & Kapelan, Z. 2014 Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling & Software* **60**, 265–276. doi: 10.1016/j.envsoft.2014.06.016.

Sebri, M. 2016 Forecasting urban water demand: a meta-regression analysis. *Journal of Environmental Management* **183** (P3), 777–785. doi: 10.1016/j.jenvman.2016.09.032.

Shvartser, L., Shamir, U. & Feldman, M. 1993 Forecasting hourly water demands by pattern recognition approach. *Journal of Water Resources Planning and Management* **119** (6), 611–627. doi: 10.1061/(ASCE)0733-9496(1993)119:6(611).

Singh, K. P., Basant, N., Malik, M. & Jain, G. 2010 Modeling the performance of 'up-flow anaerobic sludge blanket' reactor based wastewater treatment plant using linear and nonlinear approaches – A case study. *Analytica Chimica Acta* **658** (1), 1–11. doi: 10.1016/j.aca.2009.11.001.

Suhartono, S., Isnawati, S., Salehah, N. A., Prastyo, D. D., Kuswanto, H. & Lee, M. H. 2018 Hybrid SSA-TSR-ARIMA for water demand forecasting. *International Journal of Advances in Intelligent Informatics* **4** (3), 238–250. doi: 10.26555/ijain.v4i3.275.

Sun, D. & Zhang, X. Y. 2002 Forecast model for stock market based on artificial neural networks. *Journal of Jilin University (Information Science Edition)* **20** (4), 68–70.

Tian, D., Martinez, C. J. & Asefa, T. 2016 Improving short-term urban water demand forecasts with reforecast analog ensembles. *Journal of Water Resources Planning and Management* **142** (6), 04016008. doi: 10.1061/(ASCE)WR.1943-5452.0000632.

Tiwari, M. K. & Adamowski, J. 2013 Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models. *Water Resources Research* **49** (10), 6486–6507. doi: 10.1002/wrcr.20517.

Viccione, G., Guarnaccia, C., Mancini, S. & Quartieri, J. 2019 On the use of ARIMA models for short-term water tank levels forecasting. *Water Supply* **20** (3), 787–799. doi:10.2166/ws.2019.190.

Vijayalaksmi, D. P. & Babu, K. S. J. 2015 Water supply system demand forecasting using adaptive neuro-fuzzy inference system. *Aquatic Procedia* **4**, 950–956. doi: 10.1016/j.aqpro.2015.02.119.

Whitley, D. 1994 A genetic algorithm tutorial. *Statistics and Computing* **4** (2), 65–85. doi: 10.1007/BF00175354.

Yang, Y., Wu, J., Chen, Y. & Li, C. 2013 A new strategy for short-term load forecasting. *Abstract and Applied Analysis* **2013**, 208964. doi: 10.1155/2013/208964.

Yurdusev, M. A. & Firat, M. 2009 Adaptive neuro fuzzy inference system approach for municipal water consumption modeling:

an application to Izmir, Turkey. *Journal of Hydrology* **365** (3), 225–234. doi: 10.1016/j.jhydrol.2008.11.036.

Zhong, J., Hu, X., Zhang, J. & Gu, M. 2005 Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms. In *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, Vienna, Austria, pp. 1115–1121. doi: 10.1109/CIMCA.2005.1631619.

Zhou, S. L., McMahon, T. A., Walton, A. & Lewis, J. 2002 Forecasting operational demand for an urban water supply zone. *Journal of Hydrology* **259** (1–4), 189–202. doi: 10.1016/S0022-1694(01)00582-0.

First received 30 November 2020; accepted in revised form 5 February 2021. Available online 18 February 2021