Predicting water demand: a review of the methods employed and future possibilities

Gustavo de Souza Groppo, Marcelo Azevedo Costa and Marcelo Libânio

ABSTRACT

The balance between water supply and demand requires efficient water supply systems management techniques. This balance is achieved through operational actions, many of which require the application of forecasting concepts and tools. In this article, recent research on urban water demand forecasting employing artificial intelligence are reviewed, aiming to present the 'state of the art' on the subject and provide some guidance regarding methods and models to researches and professional sanitation companies. The review covers the models developed using standard statistical techniques, such as linear regression or time-series analysis, or techniques based on Soft Computing. This review shows that the studies are, mostly, focused on the management of the operating systems. There is, therefore, room for long-term forecasts. It is worth noting that there is no global model that surpasses all the methods for all cases, being necessary to study each region separately, evaluating the strengths of each model or the combination of methods. The use of statistical applications of Machine Learning and Artificial Intelligence methodologies has grown considerably in recent years. However, there is still room for improvement with regard to water demand forecasting.

Key words | artificial intelligence, machine learning, soft computing, water demand prediction

Gustavo de Souza Groppo
Applied Economics,
São Paulo University;
Sanitary Engineering,
Federal University of Minas Gerais;
The Companhia de Saneamento de Minas Gerais COPASA-MG
and
Rua Mar de Espanha,
525 CEP: 30.330-900,
Belo Horizonte - MG,
Brazil

Marcelo Azevedo Costa
Electrical Engineering,
UFMG;
Harvard Medical School,
Harvard Pilgrim Health Care,
USA;
Linköping University,
Sweden
and
Department of Production Engineering,
Federal University of Minas Gerais

Marcelo Libânio (corresponding author)
Environment and Water Resources,
UFMG;
Hydraulics and Sanitation,
São Paulo University;
University of Alberta,
Canada
and
Department of Sanitary and Environmental Engineering,
Federal University of Minas Gerais
E-mail: *marcelo@desa.ufmg.br*

ABBREVIATIONS

ACO	Ant Colony Optimization
ACPSO	Adaptive Chaos Particle Swarm Optimization
AFS	Adaptative Fourier Series
AMALGAM	Multi-Algorithm Genetically Adaptative Method
ANFIS	Adaptative Neuro Fuzzy Inference System
ANN	Artificial Neural Networks
AP	Average Price
ARIMA	Autoregressive Integrated Moving Averages
B	Bootstrap
B-ANN	Bootstrap Artificial Neural Networks
B-ELM	Bootstrap Extreme Learning Machine
BPCA	Bayesian Principal Components Analysis
CCNN	Cascade Correlation Neural Networks
CMF	Cumulative Weighting Mean Fuzzy
CNN or ConvNet	Convolutional Neural Network
DAN2	Dynamic Artificial Neural Networks
DAN2-H	Dynamic Artificial Neural Networks Hybrid

doi: 10.2166/ws.2019.122

DFS	Demand Forecasting System
DMA	District Measures Areas
E-ANN	Evolutionary Artificial Neural Networks
EC	Evolutionary Computation
EDBD	Extended Delta Bar Delta
EKF	Extended Kalman Filter
EKF-PG	Extended Kalman Filter with Programming Genetic
ELM	Extreme Learning Machine
EMD	Empirical Mode Decomposition
EMD-ANN	Empirical mode Decomposition Artificial Neural Networks
ESNN	Neural State Echo
FFNN	Feedforward Neural Networks
FIS	Fuzzy Inference Systems
FL	Fuzzy Logic
FLR	Fuzzy Linear Regression
FRNN	Totally Recurrent Neural Network
FTS	Fuzzy Takagi-Sugeno
GA	Genetic Algorithms
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GRASP	Agile Adaptative Randomization
GRNN	Generalized Regression Neural Networks
GSAA	Genetic Simulated Annealing Algorithm
HW	Holt Winters
IMF	Intrinsic Function
K-NN	K-Nearest Neighbors
LF-DFNN	Local Feedback Dynamic Neural Network
LRGF	Local Recurrent Global Feedforward
LS-SVM	Least Square Support Vector Machine
LSTM	Long Short-Term Memory
LTF	Linear Transfer Function
MAPE	Mean Absolute Percentage Error
MARS	Multivariate Adaptative Regression Splines
MFIS	Mandami Fuzzy Inference Systems
ML	Machine Learning
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
MNLR	Multiple Non-Linear Regression
MP	Marginal Price
MSE	Mean Square Error
MSE	Mean Squared Error
MS-RVR	Multi-Scale Relevance Vector Regressor
NC	Neural Computing
NRLM	Non-linear Regression Multiple
NSI	Nash-Sutcliffe
PPR	Projection Pursuit Regression
PR	Probability Ratio
RBNN	Radial Basis Neural Networks
RF	Random Forest
RVR	Relevance Vector Regressor
RW	Random Walk
RWD	Random Walk with Trend
SAR	Spatial Autoregressive
SARMA	Spatial Autoregressive Moving Average
SC	Soft Computing
SEM	Spatial Error
SVD	Singular Value Decomposition
SVR	Support Vector Regression
TDNN	Time Delayed Neural Network
TS-GRNN	Time Series Generalized Regression Neural Networks
VAR	Vector Autoregressive
W	Wavelet
W-ANN	Wavelet Artificial Neural Networks
WB-ANN	Wavelet Bootstrap Artificial Neural Networks
WDF	Water Demand Forecasting module
W-ELM	Wavelet Extreme Learning Machine
WPatt	Weighted Pattern
WSS	Water Supply Systems
WSZ	Water Supply Zone
YWS	Yorkshire Water Services
RRMSE	Relative Root Mean Square Error
R^2	Coefficient of Determination
AREP	Average Relative Error Percentge
NRMSE	Normalized Root Mean Square Error
AARE	Average Absolute Relative Error
max ARE	Maximum Absolute Relative Error
R	Correlation Coefficient
ts	Threshold Statisitic
HM *p*-value	Henriksson and Merton Probability Value
U-Statistc	Theil Inequality Coefficient
SEP	Standart Error of Prediction

PI	Persistence Index
RMS	Absolute Error Mensure
NMSE	Normal Mean Square Error
LM	Lagrange Multiplier
LR	Lagrange Robust
Moran I	Moran I Statistic
Pdv	Percentage Deviation in Peak

INTRODUCTION

Making a sufficient amount of potable water available represents a great challenge, especially in large cities. Cities have grown without appropriate planning, resulting in removal of vegetation cover and rendering soil impermeable. This has caused hydrological and meteorological alterations, such as increasing air temperature, evapotranspiration and flood risk. Consequently, water has become less available, due to either the pollution or devastation of protected water source areas. Thus, the available water supply sources are far from urban centers, making water exploration increasingly more expensive.

Dealing with these problems requires efficient Water Supply Systems (WSS) management techniques, in order to maintain a balance between supply and demand. Maintenance of this balance is achieved through operational actions, many of which require application of forecasting tools.

Due to the importance of water demand forecasting, many researchers and professionals have recently started to study it, as described in Ghalehkhondabi *et al.* (2017). The number of published articles has increased exponentially over the last 20 years, as shown in the Web of Science (Thomson Reuters). This increase may reflect the growing scarcity of water resources and the growing importance of water demand management.

Therefore, managers are concerned with planning WSS to meet water demand at lower operating costs. According to Ghiassi *et al.* (2008), the optimization of operations can result in substantial savings of 25% to 30% in operating costs, due to the reduction of costs with electricity and treatment inputs. Similar findings are presented by Odan (2013).

At the same time, there has been intensified evaluation of the existing infrastructure and expansion strategies (master plan and project/construction project study). The master plans aim at long-term investment, anticipating the water demands of the expected vegetative growth, geographic expansion and the socioeconomic and climatic variables that modify consumption behavior over time. Numerous factors affect the demanded water quantity. The most important are: climatic conditions such as temperature; precipitation and relative humidity; size of the population; water pressure in the network; losses in the system; price structure (residential, commercial, industrial and public); supply and water metering system (hydrometric); household income, size and outdoor space (Arbués *et al.* 2003; Tsutiya 2006; Wentz & Gober 2007; Scheich & Hillenbrand 2009; Nauges & Whittington 2010; Maria André and Carvalho, 2016).

After extensive literature review, Donkor *et al.* (2014) concluded that the application of a given method depends on the periodicity and the forecast horizon. According to these authors, neural networks were more likely to be used in short-term forecasting, while econometric methods and simulations were more frequently employed in the long-term forecast. Similar conclusions were found by Ghalehkhondabi *et al.* (2017). For these authors, Soft Computing (SC) methods were used mainly in the short-term forecast. According to Brentan *et al.* (2016), the use of statistical applications of Machine Learning and Artificial Intelligence methodologies in water demand forecasting has grown considerably in recent years.

The term Soft Computing (SC), also referred to as Computational Intelligence, is the combination of emerging problem-solving technologies such as Fuzzy Logic (FL), Neural Computing (NC), Genetic Algorithms (GA), Evolutionary Computation (EC), Machine Learning (ML), Probability Ratio (PR) and complementary hybrids. SC hybrid systems are further described in Bonissone (1997). Each of these technologies provides complementary reasoning and complex, real world problem-solving methods. A particularly effective combination is known as 'neuro-fuzzy systems'.

The objective of this paper is to present an extensive review of urban water demand forecasting methods to sanitation professionals and researchers, thereby providing basic guidelines for practical applications. The review covers the models developed using standard statistical techniques, such as linear regression or time-series analysis, or techniques based on SC.

Methods to support near-real-time Water Distribution Systems (WDS) management tasks, such as online pump programming and dynamic hydraulic modeling, have received less attention. Although dynamic systems are applied to water prediction problems, their application is quite limited when compared to other SC methods (Ghalehkhondabi *et al.* 2017).

This paper is organized as follows. Publications related to each investigated method are presented in Section 2. Discussions, addressing the results of the review are presented in Section 3. Final considerations and suggestions for future research are presented in Section 4.

DEMAND FORECASTING METHODS

Demand forecasting methods can be broadly classified as linear and non-linear (Zhang 2001). Linear methods use univariate time series analysis, such as exponential smoothing, Autoregressive Integrated Moving Averages (ARIMA) and linear regression methods (MLR) (eg, Adamowski & Karapataki 2010; Caiado 2010; Adamowski *et al.* 2012). Non-linear methods use non-linear regression methods (MNLR), Artificial Neural Networks (ANN) (eg, Ghiassi *et al.* 2008; Firat *et al.* 2009b; Firat *et al.* 2010; Adamowski & Karapataki 2010), FL (eg, Altunkaynak *et al.* 2005; Firat *et al.* 2009a), Support Vector Machine (SVM) (Peña-Guzmán *et al.* 2016), GA, expert systems (eg, Altunkaynak *et al.* 2005; Nasseri *et al.* 2011) and hybrid methods.

Artificial neural networks

ANN are very useful forecasting tools, due to several factors. The first is related to the requirement for a lower number of hypotheses, as compared to traditional statistical methods. A second factor relates to the generalization of results and the prediction of not observed data (Zhang *et al.* 1998). A third factor relates to the ability to deal with different degrees of non-linearity present in the water demand data. That is, they are able to model highly non-linear relations among the data and to estimate nonlinear functions with a high degree of precision. According to Adamowski *et al.* (2012), the ANN allow the use of historical series to predict future values of the possibly noisy multivariate time series.

Ghiassi *et al.* (2008) developed a Dynamic Artificial Neural Network (DAN2) method to predict water demand in a city of California. This dynamic method is a special case of Feedforward architecture (FFNN). The developed method performed better than the ARIMA and ANN methods, thus proving to be more effective for predicting water demand. The authors noted that the inclusion of meteorological information in forecasting models increases accuracy. However, even when using only water demand data, DAN2 methods provide excellent adjustments. The results obtained for monthly, weekly and daily forecasts were highly accurate, as well as the hourly models. These results demonstrate excellent efficacy for DAN2 in predicting urban water demand for all time horizons. Firat *et al.* (2009b) estimated the monthly forecast for water demand in the city of Izmir (Turkey) using various Neural networks techniques such as Generalized Regression Neural Networks (GRNN), Feedforward Neural Networks (FFNN) and Radial Basis Neural Networks (RBNN). They used various socioeconomic and climatic factors that affect consumption (average monthly water consumption, population, number of households, gross national product, average monthly temperature, total monthly rainfall, monthly average humidity and inflation). The data set was divided into two subsets (training and testing). The methods that obtained the best adjustments were also compared to Multiple Linear Regression (MLR) method. The obtained results indicated that the GRNN outperforms all other methods in modeling of monthly water consumption.

Subsequently, Firat *et al.* (2010) also estimated the monthly water demand forecast, employing a number of ANN techniques, including Cascade Correlation Neural Network (CCNN), GRNN and FFNN. Six prediction models were constructed. The best adjustment input structure was investigated by comparing the techniques employed. The M5-CCNN model comprises a CCNN network with five-month lags proved to be more efficient than the other models. When comparing the results using the three ANN techniques, the performance of the M5-CCNN model performed slightly better than the others.

Adamowski & Karapataki (2010) analyzed water demand employing weekly peak data and the weather variables weekly maximum temperature and total precipitation for two distinct regions in the city of Nicosia, Cyprus. The authors developed and compared the relative performance of 20 MLR models and 60 neural network Multilayer Perceptron (MLP) models using three different learning algorithms (Levenberg-Marquardt, Resilient back-propagation and Powell Beale conjugate gradient methods). For the two regions analyzed, the method employing the Levenberg-Marquardt algorithm presented the most accurate forecast of weekly peak demand.

Fuzzy and neuro-fuzzy methods

Altunkaynak *et al.* (2005) used the Fuzzy Takagi-Sugeno (FTS) method to forecast monthly water demand in Istanbul (Turkey). The method consists of using the water consumption values for the last three months as the independent variables. That is, present water demand is a function of demand fluctuations during the last three months. The Adaptive Neuro Fuzzy Inference System (ANFIS) method was used to determine the model parameters. The Mean Square Error (MSE) statistic for different method configurations was used to select the most effective method. The authors argue that this model is widely used than the Markov or ARIMA methods, commonly available for stochastic modeling and forecasting. One of the advantages of using the FTS method, compared to ARIMA methods, is that it does not rely on stationary and ergodicity assumptions. Finally, this method also helps to make predictions with less than 10% relative error.

Firat *et al.* (2009a) compared two types of Fuzzy Inference Systems (FIS) to predict the time series of urban water demand. The Fuzzy Inference Systems used include an ANFIS system and a Mandami Fuzzy Inference System (MFIS). The performances of the ANFIS and MFIS methods were analyzed in the training and test stages. To evaluate the best forecasting method, the performances of the two methods, in both the training and testing stages of demand, were compared to the observed values. All levels of threshold statistics employed in the study demonstrated the higher accuracy of the M5-ANFIS model over the M5-MFIS model. The M5 model comprises a Fuzzy Inference Systems with five-month lags. Therefore, the results showed that the M5-ANFIS method is superior to the M5-MFIS method for forecasting monthly demand series, and can be applied successfully for predicting water consumption.

Support vector machines

Peña-Guzmán *et al.* (2016) used the Least-Squares learning method Support Vector Machines (LS-SVM) to predict the monthly water demand for residential, industrial and commercial categories in Bogota, Colombia. To do this, they used the parameters of monthly water demand, number of users and price, for residential, industrial and commercial categories. The city employs a system of socioeconomic stratification, in accordance with the national laws on public services, where residential dwellings are classified into six strata. They used monthly records from January 2004 to December 2014. As proposed by Ghiassi *et al.* (2008), the researchers employed 80% of training data and 20% of test data. The LS-SVM method proved superior to the neural networks method, using the method of error learning back-propagation, for all categories and strata. This proved to be an effective tool for the planning and management of water demand, as it helped to identify the need for administrative decisions to regulate consumption in different strata and for different uses.

Herrera *et al.* (2010) described and compared various methods for predicting water demand in a city in the south of Spain. The methods used were: Support Vector Regression (SVR), FFNN employing the Error Backpropagation learning method, Projection Pursuit Regression (PPR), Multivariate Adaptive Regression Splines (MARS), and Random Forest (RF). In addition to these methods, researchers have proposed a simple method based on the demand profile, using weighted results from exploratory data analysis (WPatt). The results obtained identified the SVR as the most accurate method, followed closely by the MARS, PPR and RF methods.

Statistical methods

The main objective of the study by Fullerton Jr *et al.* (2016) was to analyze the dynamics of the water demand for the city of El Paso (Texas, USA), using several forecasting methodologies, among which the Linear Transfer Function

(LTF), which is an extension of the methodology described by Box and Jenkins (1976). The LTF result was superior when compared to the autoregressive vector (VAR), random walk (RW) and random walk with trend (RWD) methods.

Maria André & Carvalho (2014) estimated the function of residential water demand in the city of Fortaleza, Brazil, considering the potential impact of including spatial effects in the modeling since the exclusion of these effects underestimates the impact of income and the number of toilets on residential water demand marginal prices. First, these authors estimated an econometric method of water demand without taking spatial effects into account. This econometric method was calculated for the following specifications: average price (model AP), marginal price with difference (model MP) and marginal price with difference using the McFadden method (McFadden model). Afterwards, they calculated three models to verify the inclusion of spatial effects on water demand: Spatial Error Model (SEM), Spatial Autoregressive Model (SAR) and Spatial Autoregressive Moving Average Model (SARMA) were estimated using the following explanatory variables: average and marginal income gap, number of male and female residents, and number of bathrooms, under different spatial specifications. Results suggest that the SARMA provides the best results. However, these results contradict findings of Chang *et al.* (2010) and House-Peters *et al.* (2010), who claim that the spatial approach provides more accurate results than the SARMA. After estimating the SARMA (both for the AP model and for the McFadden model), and correcting the direct and indirect effects of estimated parameters, it was concluded that the use of a spatial approach is more advantageous. Not including spatial effects on the variables caused an underestimation of the effect of all the variables in the model. After including these spatial components, the price elasticity in the AP and McFadden models increased by 24.66% and 13.32%, respectively, affecting forecast demand.

Gagliardi *et al.* (2017), proposed a short-term water demand forecasting method based on the Markov Chain (MC) statistical concept, providing estimates for future demands and the probabilities that the forecast demand will fall within the expected variability. Two methods were proposed, one based on Homogeneous Markov Chains (HMC) and one based on non-Homogeneous Markov Chains (NHMC). These methods were applied to three district measurement areas (DMA) located in Yorkshire (UK), in order to predict water demands from 1 to 24 h later. Subsequently, the results were compared with the predictions of the two methods used as benchmarks (ANN, Naive Bayes). The results show that the HMC method provides more accurate short-term predictions than NHMC. Both methods provide probabilistic information on stochastic demand forecasting with reduced computational effort, as compared to most existing methods. This information is not readily available in either the ANN or Naive Bayes benchmark methods. However, it can be obtained through post-processing analysis using Monte Carlo simulations that are computationally more expensive.

Hybrid methods

The main factors that impact urban water demand are often difficult to identify using traditional algorithms due to the many factors that are uncertain and difficult to quantify. Some filters, wrappers and embedded systems may be employed to deal with this problem. Each one has strengths and weaknesses. Table 1 presents an overview of the main strengths and weaknesses of the three types of variable selection methods (Freitas 2007).

Methods for selection of variables are important in the improvement of forecasting models, increasing the efficiency of the process caused by the mitigation of the known curse of dimensionality (Guyon & Elisseeff 2003), reducing the computational cost (Piramuthu 2004) and obtaining a deeper insight into the underlying processes that generated the data (Sayes *et al.* 2007). Efficient search strategies can be planned without necessarily sacrificing

Table 1 | Comparison of methods for selection of variables

Types	Strengths	Weaknesses
Filter	Rapid execution and robustness to data overfit.	Redundant variables selection
Wrapper	High precision	Slow execution and data overfit
Embedded	Combination of filters and wrappers	Preliminary knowledge for variable selection

predictive performance (Guyon & Elisseeff 2003). Several selection strategies are being developed to minimize the computational load caused by the exhaustive search, such as the work of Sorjamaa *et al.* (2007) and Hsu *et al.* (2011). Recently, hybrid approaches have been proposed to combine the advantages of filters and wrappers (Crone & Kourentzese 2010; Stańczyk 2015).

Another aspect of research concern learning using ensemble, where the methods that generate several models are combined to predict a new case. The idea of these methods can be described as the construction of a predictive model through the integration of multiple methods (Dietterich 2000). One of the advantages is the improvement of generalization performance (Hansen & Salamon 1990; Sharkey 1996; Zhang 2003; Melin *et al.* 2012). However, such improvement relies on the quality of its components and the diversity of the error presented by them (Granger 1989; Perrone & Cooper 1993; Krogh & Vedelsby 1995; Sollich & Krogh 1996; Granitto *et al.* 2005; Gashler *et al.* 2008; Al-Zahrani & Abo-Monasar 2015; Ren *et al.* 2016; Wang *et al.* 2018), that is, each of the components in an ensemble should perform well and at the same time should generalize differently. It makes little sense to combine methods that adopt the same procedures and hypotheses for solving a problem (Perrone & Cooper 1993). If components that have the same error pattern are combined, there will be no incremental performance, only increasing the computational cost, without practical performance results.

According to Mendes-Moreira *et al.* (2012), the generation of homogeneous ensembles is the area of learning with better coverage in the literature. In homogeneous ensembles, models are generated using the same algorithm. Heterogeneous ensembles are generated using more than one machine learning algorithm. It is expected that the heterogeneous approach can obtain models with greater diversity due to the different nature of the basic apprentices (Webb & Zheng 2004). Some authors claim that these ensembles perform better when compared to homogeneous ensembles (Granger 1989; Krogh & Vedelsby 1995; Wichard *et al.* 2003). Another approach, usually employed, is the one that combines the use of different induction algorithms merged with different sets of parameters (Rooney *et al.* 2004).

The most widely known ensemble methods are the Bagging (Bootstrap Aggregating) introduced by Breiman (1996), Boosting (Freund & Schapire 1996) and Random Forest (Breiman 2001).

Filter

Xu *et al.* (2018) used the energy spectrum (Oshima & Kosuda 1998) and the largest Lyapunov coefficient (Tsonis 1992) to qualitatively examine the characteristics of the time series of water demand. Results indicate that the water demand time series presents characteristics of chaos, i.e., an upward trend of the time series is observable, but the evolution law and variation characteristics of the data cannot be determined easily. Results are similar to those obtained by Zhao & Zhang (2008) and Bai *et al.* (2014). Thus, the problems of predicting water demand can be translated into problems of forecasting chaotic series. Therefore, pre-processing to improve the accuracy of the results is necessary.

In the filtering procedure, set selection is a data pre-processing step, regardless of the induction algorithm. The filter is usually robust to data overfit, but fails to find the most promising subset of variables (Freitas 2007), on average. The main weakness of this approach is that it ignores the effects of the selected subsets on the performance of the induction algorithm (John *et al.* 1994; Kohavi & John 1997).

In the literature, several types of filters are used: (i) Stepwise Regression used to identify significant delays of the autoregressive component of the dependent variable as inputs to a MLP (Dahl & Hylleberg 2004); (ii) Spectral Analysis used to evaluate the cyclic data patterns, decomposing the time series into the underlying sine and cosine functions of the wavelengths (Kay & Marple 1981); (iii) Singular Spectrum Analysis (Hassani 2007; Du *et al.* 2017); (iv) Empirical Mode Decomposition (EMD) where the time frequency resolution by which the steady and nonlinear time series behavior can be decomposed (Di *et al.* 2014; Shabri & Samsudin 2015); and (v) Kalman filter (Poli & Jones 1994; Nasseri *et al.* 2011; Arandia *et al.* 2016; Karunasingha & Liong 2018) and Hodrick-Prescott filter (Li & Huicheng 2010).

It is worth mentioning that according to Zhang & Qi (2005), machine learning methods, without proper pre-processing, can become unstable and generate suboptimal results.

Wrappers

Wrappers, popularized by John *et al.* (1994) and Kohavi & John (1997), provide a simple and powerful way to address the problem of variable and/or attribute selection, regardless of the learning machine to be employed. According to Freitas (2007), the strength of this method is that it takes into account the bias of induction algorithm and considers the variables in-context. At first, the search is exponential, but it is possible to implement stochastic searches (eg., genetic algorithms (The genetic algorithm is a metaheuristic inspired by the process of natural selection, motivated by an analogy to biological evolution. These algorithms, instead of searching for general and specific hypotheses, from simple to complex, generate successor hypotheses, repeatedly mutating and recombining parts of the best-known hypotheses (Mitchell, 1997).) and simulated annealing) or sequential ones (eg., direct search and backward deletion). Therefore, there are numerous possibilities to be studied empirically.

Embedded

In the embedded model, the task of selecting attributes is dynamically performed by the machine learning algorithm. The attribute selection process is not distinct from the training of the model, and the results are calibrated in relation to a given classifier or specific regressor. For Zanchettin (2008), one of the strengths of this approach is that it makes the best use of the available data, i.e., does not have to divide into training and test data, and is faster because it does not require multiple trainings.

Hybrid models with coupled filters

A hybrid model that combines Extended Kalman Filter (EKF) and Genetic Programming (PG) was proposed by Nasseri *et al.* (2011) to predict monthly water demand. According to Nasseri *et al.* (2011) GP is a symbolic regression method based on a tree-structured approach presented by Koza (1990). This method belongs to a branch of evolutionary method, which mimics the natural process of struggle for existence (Holland 1975). The main advantage of the proposed approach is the possibility of achieving the fewest non-linear and deterministic mathematical formulations for monthly water demand forecasting, via the evolutionary method. The results obtained using the EKF-PG and PG hybrid models showed noticeable effect on the accuracy of the forecast.

Adamowski *et al.* (2012) conducted daily water demand forecasts for the summer months in the Canadian city of Montreal. They employed a hybrid method based on Discrete coupling Wavelet Transforms (W) and Artificial Neural Networks. The W-ANN hybrid models were compared to ANN, MLR, Multiple Non-Linear Regression (MNLR) and ARIMA methods to predict a one-day lead time. The results indicate that the W-ANN were more robust than all other methods, suggesting the promising potential of this method for predicting urban water demand. According to the authors, the accuracy of the W-ANN prediction may be useful in the management, planning and evaluation of existing systems, conservation initiatives, analysis of drought conditions and water pricing policies.

Odan & Reis (2012) aimed to identify a method that best fit the hourly consumption data for a given supply zone in the city of Araraquara (Brazil). Before using the observed values, the procedure of pre-processing missing data resulting from record failures, or the presence of values higher or lower than twice the absolute value of the standard deviation, was used. The pre-processing method used was the Bayesian Principal Component Analysis (BPCA) developed by Oba *et al.* (2003). This method is based on a principal component regression, Bayesian estimation, and a repetitive expectation-maximization algorithm. The method uses an iterative variant Bayesian algorithm to estimate the posterior distribution of the model parameters and the defective data, until their convergence. According to the researchers, the pre-processing technique achieved good results, even when 40% of the data are missing, thus exceeding the performance of the models based on K-Nearest Neighbors (K-NN) and Singular Value Decomposition (SVD). After completion of the preprocessing, correlation analysis was employed to identify the input variables (temperature, relative humidity, time consumption and reservoir level). It was observed that the inclusion of climatic variables (temperature and relative humidity) improved the demand forecasting. The time of consumption was considered as variable in the correlation analysis, due to the cyclical water consumption throughout the day. This study addressed the problem of real-time WSS prediction using ANN, DAN2 and the hybrid models ANN-H and DAN2-H. Both ANN-H and DAN2-H hybrid models use the error produced by the Fourier series prediction as inputs, thus achieving promising results. The DAN2-H hybrid model presented the best results, both for the hourly and the daily forecasts.

According to Liu *et al.* (2013), the main factors that impact urban water demand patterns are often difficult to identify, employing traditional algorithms, due to numerous uncertain factors and difficulties in quantification. In order to get around this problem, researchers have proposed an improved attribute reduction algorithm based on the cumulative weighting mean C-Fuzzy (FCM). This algorithm was used to analyze the main factors that impact the pattern of daytime water demand in the city of Hangzhou (China). The data used in this study included minimum, mean and maximum daily temperature, daily precipitation, day of the week or weekend, and daytime water demand pattern. Later, they used SVR to evaluate the influence of the main factors on the prediction of the diurnal pattern of water demand. The best performing model included minimum and maximum daily temperature, and day of the week. According to the researchers, this algorithm proved to be an effective and feasible method for solving the cluster problem of consecutive curves, as in the daytime water demand pattern.

Recently, Tiwari & Adamowski (2015) performed weekly and monthly forecasts of water demand in the city of Calgary (Canada). The method used in their study was the hybrid Wavelet-Bootstrap Artificial Neural Network (WB-ANN). The use of this method aimed to improve the accuracy and reliability of demand forecasting, incorporating Wavelet processing capacity and Bootstrap (B) analysis using artificial neural networks. This method was then compared to the standard ANN, ANN based on Bootstrap (B-ANN) and based on W-ANN. For prediction of weekly and monthly peaks, the WB-ANN and W-ANN hybrid methods were found to be more accurate when compared to the B-ANN and ANN methods. The results of the forecasts using the WB-ANN hybrid methods and W-ANN were very effective in predicting water demand peaks. This indicates that Wavelet analysis significantly improved the performance of the method, while the Bootstrap technique improved the reliability of the forecasts. Another point highlighted by the researchers is the effectiveness of the methodology in situations with limited data availability.

Due to the difficulty of modeling water demand time series using traditional statistical methods, Shabri & Samsudin (2015) proposed a hybrid model that combines the Empirical Mode Decomposition (EMD), method proposed by Huang *et al.* (1998), and the Least Square Support Vector Machine (LS-SVM) method to predict water consumption. EMD was used to decompose the non-linear and non-stationary series of water demands into various components of intrinsic mode functions (IMF) and a residual component. The LS-SVM algorithm was built to predict these intrinsic and residual components individually, which were later aggregated to produce the expected final value. The empirical results indicate that the proposed method outperforms the single LS-SVM method and the ANN, without EMD pre-processing, and EMD-ANN method.

Tiwari *et al.* (2016) used the newly developed Extreme Learning Machines (ELM) method, either alone or together with wavelet analysis (W) or Bootstrap (B) methods, to forecast daily water demand in the city of Calgary, Canada. Subsequently, they were evaluated and compared to equivalent methods based on traditional ANN (i.e., ANN, W-ANN, B-ANN). The B-ELM and B-ANN hybrid methods provided similar accuracy in the predictions on peak days. However, the W-ANN and W-ELM methods provided higher accuracy, with the W-ELM method surpassing the W-ANN method. The superiority of the W-ELM method over the traditional methods (W-ANN or B-ANN) demonstrates the importance of wavelet transformation in urban water demand forecast modeling. This highlights the ability of wavelet transformation to decompose time series with non-stationary behavior into discrete components, highlighting the cyclical patterns and trends.

Arandia *et al.* (2016) presented a methodology for predicting water demand for the short term through the combination of autoregressive components, moving averages, integration filter and seasonal terms added to the

ARIMA method (SARIMA). According to Caiado (2010), this method has not received much attention for predicting water demand despite its parsimony qualities and ease of interpretation of its parameters in explicit function of mathematical formulations. For Arandia *et al.* (2016), offline mode is best suited for utility operations (such as daily water production sizing), while online mode may be more appropriate for other operations (such as pump scheduling). In offline mode, the method employs the re-estimated models using more recent historical data. In online mode, the method applies the Kalman filter in order to update and optimize the models using real-time 'feed' data. Three qualitatively different sets of data were modeled. Structures and sample size estimation of data used for training were identified. These models were applied to anticipate demands for 24 hours in advance using offline and online modes. Subsequently, results were analyzed, compared and demonstrated the application of the method in the forecast of daily water production using SARIMA methods. Unlike the ANN methods, or other methods known as 'black box', the SARIMA methodology can be shaped in the form of a state-space model, identifying the most appropriate parametric structures for water demands with time resolutions varying from hourly to daily.

Hybrid models with multiple methods employees

Pulido-Calvo & Gutiérrez-Estrada (2009) proposed a hybrid model using Neural Networks, Inference Fuzzy and Genetic Algorithms (GA) to predict the daily water demand in the irrigation district of Andalusia in southern Spain. ANN methods were trained using the Extended-Delta-Bar-Delta (EDBD) algorithm (Pulido-Calvo *et al.* 2003) and, subsequently, recalibrated with a variation of the Error Backpropagation algorithm known as Levenberg-Marquardt (Shepherd 1997). According to Wilamowski & Yu (2010), the Levenberg-Marquardt algorithm is currently one of the most efficient for training artificial neural networks, especially when they involve long time series. The results obtained using the hybrid model indicated that it is superior to autoregressive, univariate and multivariate ANN methods. Therefore, the hybrid model is a very powerful tool for developing appropriate policies on irrigation water consumption.

Caiado (2010) analyzed the performance of demand forecast in Spain using the univariate exponential smoothing methods of Holt-Winters (HW), ARIMA and generalized autoregressive conditional heteroscedasticity (GARCH), for one to seven steps ahead. According to the author, the combination provides more accurate predictions than individual methods. Predictions can be combined using either a simple or a weighted average. All possible combinations of the Holt-Winters, ARIMA and GARCH prediction methods were considered, using the simple average of forecasts for 1 to 7 steps (days) ahead. To calculate the ideal weights, the author considered two approaches. First, the weights were proportional to the inverse MSE of each individual method (Makridakis and Winkler, 1983). Second, the weights were proportional to the inverse of the prediction square errors (PSE) of each individual method. If the performance of individual methods changes during the forecast period, combining predictions with the use of inverse PSE weights may result in more accurate predictions than the method that uses MSE inverse weights. The results indicate that the combination of forecasts can be very useful, especially for short-term forecasts. However, the performance of this approach is not consistent over the seven days of the week. On the other hand, individual predictions of HW and GARCH methods can improve prediction accuracy on specific days of the week.

Azadeh *et al.* (2012) presented a hybrid approach, using ANN and Fuzzy Linear Regression (FLR), to improve the water demand forecasting. According to the researchers, this approach can be easily applied in uncertain or complex environments, given its flexibility. Their proposed hybrid approach was applied to predict daily water demand in Tehran (Iran). The variables used were the daily maximum temperature, the maximum temperature predicted for the following day, the precipitation index and the demand on hot days and cold days. Results indicated that ANN outperformed FLR on hot days due to its ability to deal with complexity and non-linearity. However, both ANN and FLR were ideal on cold days.

According to Al-Zahrani & Abo-Monasar (2015), climatic factors play a fundamental role in predicting short-term water demand since they have a direct influence on water consumption. Their study was conducted in the city of Al-Khobar, Saudi Arabia, and employed the climatic humidity and temperature parameters (minimum, average and maximum), rainfall intensity, occurrence of rainfall and wind speed associated with historical daily water consumption. In this study, the potential of hybrid models, for the daily water demand forecast, was investigated by coupling Time Series to Artificial Neural Networks (TS-GRNN). Results indicate that the TS-GRNN hybrid models provide better predictions when compared to ANN or TS methods alone. According to the authors, the results indicate that temperature is the most important meteorological predictor in the neural network training. Humidity, wind speed and the occurrence of rain also proved to be important, but cannot be used without temperature. On the other hand, rainfall intensity is the parameter that makes the smallest contribution, to the capacity of the model for predicting water demand, during the ANN training process.

Hybrid models with embedded signal processing

Bai *et al.* (2014) proposed a Multi-Scale Relevance Vector Regressor (MS-RVR) approach for predicting daily water demand in the city of Chongqing (China). This method is a hybrid that combines the RVR and the multiscale analysis of the wavelet. The coefficients of the Wavelet, of all scales obtained, were used to train a machine learning model using the Relevance Vector Regression (RVR) method. Subsequently, the estimated coefficients of the RVR were used to generate forecasting results through the inverse wavelet transform. In order to facilitate the prediction of MS-RVR, the chaos characteristics of the daily series of water supply were analyzed using the Adaptive Chaos Particle Swarm Optimization (ACPSO) algorithm to determine the optimal combination of RVR method input variables. Finally, the researchers compared the results of the best MS-RVR method with two recent methods proposed by Firat *et al.* (2010), called GRNN and FFNN, using the same data set and using the same accuracy criteria. Results show the proposed MS-RVR method to be much more precise.

On-line modelling in water distribution systems

According to Zapelan (2002), numerous mathematical methods are used to describe the behavior of WDS. According to Gupta *et al.* (1999), some methods belongs to a group of intrinsically intractable problems commonly referred to as NP-hard. To have meaningful use, WDS mathematical model must first be calibrated (Zapelan 2002). This calibration is defined as a process in which several WDS model parameters are adjusted until the model mimics the actual WDS behavior as closely as possible. WDS hydraulic model calibration is enhanced by the application of appropriate optimization methods. Recently, the application of nature-based stochastic optimization techniques, such as Genetic Algorithm (GA), has expanded. According to Broad *et al.* (2005), the GA demonstrated their applicability in the optimization of WDS operations by minimizing the cost subject to hydraulic constraints. Moreover, according to Broad *et al.* (2005), the focus of optimization has expanded to include issues related to water quality, generating additional complexity and increasing computational overhead. The real-time modeling of WDS, according to Hutton *et al.* (2014), often neglects the multiple sources of system uncertainty, thus affecting the identification of robust operational solutions. In order to minimize such a gap, these authors provide a critical review of various methods applied in the quantification and reduction of uncertainty in each stage of cascade modeling, ranging from calibration through data assimilation to model prediction. This review also includes promising methods to address the uncertainty of the applied model in related scientific fields and considers key issues that govern its application in WDS control.

Odan (2013) implemented the Multi-Algorithm Genetically Adaptive Method (AMALGAM) integrated to the hydraulic simulator (EPANET 2) and the Neural Network Dynamics method (DAN2-H). The study was carried out in the city of Araraquara (Brazil). The method was applied in three different sectors (Eliana, Iguatemi and Martinez) and the resulting operational strategies yielded reductions of 14%, 13% and 30%, respectively, in the cost of electricity consumption. This optimization method has proven to be a robust and efficient tool.

Xu *et al.* (2018) used the energy spectrum (Oshima & Kosuda 1998) and the largest Lyapunov exponent (Tsonis 1992) to qualitatively examine the characteristics of the time series. Their results indicate that the time series of water demand presents chaos characteristics. The results are in agreement with those obtained by Zhao & Zhang

(2008); results of Bai *et al.* (2014) were similar as well. Therefore, the problems of predicting water demand can be translated into problems of forecasting chaotic series.

Romano & Kapelan (2014) presented an innovative methodology to predict water demand for up to 24 hours, aiming to support operational management in near real-time water distribution systems. The methodology is based solely on the analysis of time series of water demand (estimated by mass balance analysis), using Evolutionary Artificial Neural Networks (E-ANN). The Demand Forecasting System (DFS) main features include continuous adaptability to changing water demand patterns, generic applicability and transparency for different signal of demands, drastic reduction in required, human expert effort in the projection of an ANN method, and, feasibility of implementing this methodology in an online environment. The DFS consists of four main components: (i) the data preprocessing module; (ii) the ANN optimization module; (iii) the ANN construction module; and (iv) the Water Demand Forecasting module (WDF). For the specific demand signal being analyzed, the data preprocessing module prepares the raw data to facilitate/improve the process of constructing the E-ANN method and thus obtain a more accurate WDF. The DFS allows the application of two alternative approaches to water distribution systems. The first model (pE-ANN) used many E-ANN models in parallel to predict the demands, separately, for different times of day. The second model (rE-ANN) used a recursively hourly forecast horizon to predict the demands. Both approaches have been used and tested for three District Measures Areas (DMA) and a Water Supply Zone (WSZ) of Yorkshire Water Services (YWS) covering significant parts of two cities in Yorkshire County, in the UK. According to the researchers, this new methodology allows more accurate predictions to be generated, thus demonstrating the potential for providing substantial improvements to the state of the art in the management of intelligent water distribution systems, in near real time. The performances of the predictions were evaluated in terms of Nash-Sutcliffe (NSI), the mean square error (MSE) and mean absolute percentage error (MAPE). Results showed that, regardless of the approach used, the multiple E-ANN models slightly surpass the single E-ANN model in terms of accuracy in predicting water demand.

According to Herrera *et al.* (2010), the use of SVR is one of the best machine learning options for short-term demand forecasting. Therefore, Brentan *et al.* (2016) proposed an online hybrid model applying SVR and Adaptive Fourier Series (AFS) to improve prediction. The study was carried out in the municipality of Franca (Brazil), in which data on water demand by residential consumers, temperature, humidity, precipitation and wind speed were used.

Table 2 summarizes the information about the use of water demand forecasting methods.

DISCUSSION

The literature review shows that in the context of water demand forecasting, several Soft Computing (SC) methods have been studied and applied to deal with problems of precision, stochasticity and non-linearity (Bonissone 1997). This makes predicting the demand for water a very difficult task. In this sense, the SC methods (Fuzzy logic, neural computation, genetic algorithms, evolutionary computation, machine learning and hybrid systems that use such complementarity) contributed a great deal to the methodological advances in urban water demand forecasting. According to Bates & Granger (1969), the possibility of increasing the precision of the prognostics benefits from the complementarity of the information contained in each individual forecast. This results from the proposition that the expected variance of the errors of the combined forecast is less than the smallest of the variances of the individual forecasts. Each of these methods provides additional tools, trying to solve complex real-world problems.

However, a closer look at the revised literature and recent advances in SC methods suggest that there is still room for improvement in predicting water demand.

For neural networks, these can be divided into several types of architecture. However, there are few methods used to predict water demand (Ghalehkhondabi *et al.* 2017). An extensive literature review revealed that only a small number of architectures were employed to predict urban water demand. The most used architecture in predicting urban water demand is the FFNN, also known as MLP. References to the Radial Base Neural Network (RBNN) (Broomhead & Lowe 1988), Probabilistic Neural Network

Uncorrected Proof

Table 2 | Water demand forecasting methods according to the referenced literature

Authors	City/Country	Employed Variable	Forecast horizon	Measures of accuracy	Methods Employed*
Adamowski & Karapataki (2010)	Nicosia/Cyprus	Temperature, precipitation, demand	Weekly	R^2, RMSE, AARE, maxARE, PI	Artificial Neural Networks, Multiple Regression (MLR)
Adamowski *et al.* (2012)	Montreal/Canada	Temperature, precipitation, demand	Daily	R^2, NSI, RMSE, RRMSE	Wavelet Artificial Neural Networks (W-ANN), Autoregressive Integrated Moving Average (ARIMA), Multiple Regression (MLR), Nonlinear Regression (MNLR)
Altunkaynak *et al.* (2005)	Istanbul/Turkey	Demand	Monthly	ARE, RMSE	Inference Fuzzy (ANFIS, MFIS, FTS), Autoregressive Integrated Moving Average (ARIMA)
Al-Zahrani & Abo-Monasar (2015)	Al Kobar/Saudi Arabia	Temperature, precipitation, relative humidity, speed wind, demand	Daily	MAPE, R^2	Artificial Neural Networks (ANN): GRNN, TS-GRNN
Azadeh *et al.* (2012)	Tehran/Iran	Temperature, precipitation, demand	Daily	MAPE	Artificial Neural Networks (ANN), Fuzzy Linear Regression (FLR)
Bai *et al.* (2014)	Chongqing/China	Demand	Daily	MAPE, NRMSE, R	Multi-Scale Relevance Vector Regression (MS-RVR), Artificial Neural Networks (ANN): GRNN, FFNN
Brentan *et al.* (2016)	Franca/Brazil	Temperature, precipitation, relative humidity, wind speed	Daily	RMSE, MAE, R^2	Support Vector Regression (SVR), Support Vector Regression Adaptive Fourier Serie (SVR-AFS)
Caiado (2010)	Spain	Demand	Daily	MSE	Autoregressive Integrated Moving Average (ARIMA), Holt Winters, Generalized Autoregressive Conditional Heteroskedasticity (GARCH)
Firat *et al.* (2009a)	Izmir/Turkey	Demand	Monthly	NRMSE, AARE, TS	Inference Fuzzy: ANFIS, MFIS
Firat *et al.* (2009b)	Izmir/Turkey	Population, number of households, product gross national, temperature, precipitation, humidity, inflation, demand	Monthly	NRMSE, R, NSI	Artificial Neural Networks (ANN): GRNN, FFNN, RBNN
Firat *et al.* (2010)	Izmir/Turkey	Demand	Monthly	NRMSE, R, NSI	Artificial Neural Networks (ANN): GRNN, CCNN, FFNN
Fullerton *et al.* (2016)	El Passo/USA	Water demand per customer, real average revenue price, number de customers, days over 32 °C, total monthly precipitation, non-seasonally adjusted non-farm employment	Monthly	HM *p*-value, U statistic	Linear Transfer Function (LTF) ARIMA, Vector Autoregression (VAR), Random walk (RW), Random Walk with Drift (RWD)

(*continued*)

Table 2 | continued

Authors	**City/Country**	**Employed Variable**	**Forecast horizon**	**Measures of accuracy**	**Methods Employed***
Gagliardi *et al.* (2017)	Yorkshire County/ United Kingdom	Demand	Horary	NSI	MLP, NB, HMC, NHMC
Ghiassi *et al.* (2008)	San Jose and surrounding cities of Campbell, Cupertino, Los Gatos, Monte Sereno and Saratoga/ USA	Volume pumped	Horary, daily, weekly, monthly	MAPE	Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Network (DAN2)
Herrera *et al.* (2010)	Spain	Temperature, the wind speed, atmospheric pressure, precipitation, demand	Horary	RMSE, MAE, NSI	Artificial Neural Networks: FFNN, PPR, MARS, SVR, RF, WPatt
Liu *et al.* (2013)	Hangzhou/China	Minimum, maximum and average daily temperature, daily index of weather conditions, precipitation, weekday or weekend, demand	Horary	AARE	SVR
Maria André & Carvalho (2014)	Fortaleza/Brazil	Price, marginal price, average price, income, number of bathrooms, garden/land cover area, type of residence, sex	Monthly	R^2, LM, Moran I	Spatial Models (SEM, SAR, SARMA)
Nasseri *et al.* (2011)	Tehran/Iran	Demand	Monthly	R^2, NMSE	Genetic Programming (GP), EKFPG
Odan & Reis (2012)	Araraquara/Brazil	Temperature, relative humidity, demand	Horary	MAE, R	Artificial Neural Networks (ANN): DAN2, DAN2-H, ANN, ANN-H
Peña-Guzmán *et al.* (2016)	Bogotá/Colombia	Demand, number of users, price	Monthly	RMSE, R^2, AARE	LS-SVM, Artificial Neural Networks (ANN): FFNN
Pulido-Calvo & Gutiérrez-Estrada (2009)	Córdoba/Spain	Demand	Daily	R, R^2, RMS, SEP, NSI, PI	Artificial Neural Networks (ANN): FFNN-EDBD, FFNN-LM, FFNN-EDBD FUZZY, FFNN-EDBD FUZZY
Romano & Kapelan (2014)	Yorkshire County/ United Kingdom	Demand	Time, daily	MSE, MAPE, NSI	Artificial Neural Networks (ANN): E-ANN
Shabri & Samsudin (2015)	Batu Pahat /Malaysia	Demand	Monthly	MAE, RMSE, R	LS-SVM, EDM-LS-SVM, Artificial Neural Networks (ANN): ANN, EDM-ANN
Tiwari & Adamowski (2015)	Calgary/Canada	Temperature, precipitation, demand	Daily, monthly	R^2, RMSE, MAE, PI, Pdv	Artificial Neural Networks (ANN): B-ANN, W-ANN, WB-ANN
Tiwari *et al.* (2016)	Calgary/Canada	Temperature, precipitation, demand	Daily	R^2, RMSE, MAE, PI, Pdv	Artificial Neural Networks (B-ANN, W-ANN, WB-ANN), Extreme Learning Machines (ELM, W-ELM, B-ELM)

(PNN) (Specht 1990), and ELM (Huang *et al.* 2006) were also found. However, no studies using the architecture of Recurrent Neural Networks such as Hopfield (Hopfield 1982), Jordan (Jordan 1986), Totally Recurrent Neural Network (FRNN) (Williams & Zipser 1989), Elman Networks (Elman 1990), Local Recurrent Global Feedforward (LRGF) (Tsoi & Back 1994), Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and Neural State Echo (ESNN) (Jaeger 2001) were found to predict water demand.

Recent advances in neural networks, such as the Convolutional Neural Network (CNN or ConvNet), have not been used to predict urban water demand, opening up a promising space for this problem. Recently, Borovykh *et al.* (2017) developed a deep convolutional neural network for predicting multivariate time series, based on the recent WAVENET architecture developed by Oord *et al.* (2016). In addition, according to Qiu *et al.* (2014), deep learning methods have proven to be highly promising in various areas of prediction.

Also, in relation to the neural networks, the number of hidden layers and the algorithms used in the training can affect the ANN performance. Finding the best architecture can be a difficult task, and FFNN may not always be the best method or provide the best results (Herrera *et al.* 2010). Other algorithms such as Dynamic Gaussian Bayesian Network (DGBN) (Froelich 2015) and ELM have been researched and developed to optimize predictions based on neural networks (Tiwari *et al.* 2016).

This literature review shows that the studies are more focused on the operating system of management (short term), according to the classification proposed by Gardiner & Herrington (1990) and used by Donkor *et al.* (2014). There are very few studies that address medium- and long-term forecasting. A possible explanation may be associated with the inadequacy of the basic ANN architecture, such as FFNN, to deal with noisy data (Ghalehkhondabi *et al.* 2017), limiting its application to less complex and linearly inseparable patterns. There are also certain types of patterns in the time series of water demand, which require a great need for pre-processing. In order to improve prediction accuracy, researchers began to develop hybrid methods based on wavelet and bootstrap coupling (Adamowski *et al.* 2012; Tiwari & Adamowski 2015; Tiwari *et al.* 2016; Altunkaynak & Nigussie 2017; Chen *et al.* 2017; Du *et al.* 2017). These methods were more robust in the prediction, compared to the FFNN, RLM, Nonlinear Regression Multiple (NRLM) and ARIMA regression, indicating that the wavelet transform significantly improved the performance of the methods, highlighting their processing capacity in discrete and non-steady state decomposition of time series components. It also highlights cyclical patterns and trends, while the bootstrap technique has improved forecast reliability, suggesting a promising potential of this hybrid method to predict urban water demand.

In relation to the hybrid models, we highlight the Time Delayed Neural Network (TDNN) (Htike & Khalifa 2010), ANN with Fuzzy (Araujo *et al.* 2015), Local Feedback Dynamic Neural Network (LF-DFNN) (Barbounis & Theocharis 2007), Neuro-Fuzzy (ANFIS) (Jang 1991, 1993), ARIMA-ANN (Zhang 2003), ARIMA-SVR (Chen 2011) combination of methods ARIMA-HW-GARCH (Caiado 2010) and Continuous Deep Belief Neural Network (CDBNN) (Xu *et al.* 2018).

The Metaheuristic research line has great potential of application in water demand forecast. However, there are few studies using genetic programming, such as Pulido-Calvo & Gutiérrez-Estrada (2009) and Nasseri *et al.* (2011). Odan (2013) has already used the AMALGAM optimization method and, more recently, Bai *et al.* (2014) used the ACPSO. Other metaheuristics, such as the ant colony optimization (ACO) method (Colorni *et al.* 1991), agile adaptive randomization (GRASP) procedures (Feo & Resende 1989), simulated annealing genetic algorithm (GSAA), can also be employed.

FINAL CONSIDERATIONS

In this work, an extensive review on urban water demand forecasting employing artificial intelligence was presented in order to provide some guidance regarding methods and models to professionals in the sanitation companies. This article should be used to address short-, medium- and long-term planning decisions, and by researchers that have the goal of improving the models. This review shows that the studies are, mostly, focused on the management of the operating systems. There is, therefore, room for long-term forecasts and support for the development of master plans.

It is worth mentioning that there is no global model, i.e., one soft computing method that outperforms all methods in all cases. Each region should be studied separately, seeking to add the strengths of each model, or combination of models, or choosing a more appropriate model for a given occasion. Another point concerns the robustness of the performance of the models. The results indicate that hybrid and innovative models showed superior results, when compared to the classical models analyzed.

Although major advances in soft computing methods have been made recently, no new method such as deep neural networks, among others, has emerged as the best forecasting approach. Therefore, water demand forecasting still remains a research problem, which leaves room for researchers to develop hybrid or specific methods for specific applications.

The use of statistical applications of Machine Learning and Artificial Intelligence methodologies in water demand forecasting has grown considerably in recent years. Nevertheless, there is still room for improvement with regard to water demand forecasting.

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., Chang, H. F., Prascher, 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* **47**, 1–14. doi:10.1029/2010WR009945.

Altunkaynak, A. & Nigussie, T. A. 2017 Monthly water consumption prediction using season algorithm and wavelet transform-based models. *Journal of Water Resource Planning and Management* **143** (6). doi:10.1061/(ASCE)WR.1943-5452.0000761.

Altunkaynak, A., Özger, M. & Çakmakci, M. 2005 Water consumption prediction of Istanbul city by using fuzzy logic approach. *Water Resources Management* **19**, 641–654. doi:10.1007/s11269-005-7371-1.

Al-Zahrani, M. A. & Abo-Monasar, A. 2015 Urban residential water demand prediction based on artificial neural networks and time series models. *Water Resources Management* **29**, 3651–3662. doi:10.1007/s11269-015-1021-z.

Arandia, E., Ba, A., Eck, B. & Mckenna, S. 2016 Tailoring seasonal time series models to forecast short-term water demand. *Journal of Water Resource Planning and Management* **142** (3). doi:10.1061/(ASCE)WR.1943-5452.0000591.

Araujo, R., Valença, M. & Fernandes, S. 2015 A New Approach of Fuzzy Neural Networks in Monthly Forecast of Water Flow. In: *Advances in Computational Intelligence*, Vol. 9094 (I. Rojas, G. Joya & A. Catala, eds). IWANN 2015. Lecture Notes in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-19258-1_47

Arbués, F., Valiñas, M. A. G. & Espiñeira, R. M. 2003 Estimation of residential water demand: a state-of-the-art review. *The Journal of Socio-Economics* **32**, 81–102. http://www.sciencedirect.com/science/article/B6W5H…3ad935b7298f4feb8cdd

Azadeh, A., Neshat, N. & Hamidipour, H. 2012 Hybrid Fuzzy regression – artificial neural network for improvement of short-term water consumption estimation and forecasting in uncertain and complex environments: case of a large metropolitan city. *Journal of Water Resource Planning and Management* **138** (1), 71–75. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000152

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.

Barbounis, T. G. & Theocharis, J. B. 2007 A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation. *Neurocomputing* **70** (7–9), 1525–1542. https://doi.org/10.1016/j.neucom.2006.01.032.

Bates, J. M. & Granger, C. W. J. 1969 Combination of forecasts. *Operations Research Quarterly* **20** (4), 451–468. doi:10.2307/3008764.

Bonissone, P. P. 1997 Soft computing: convergence of emerging reasoning technologies. *Soft Computing* **1**, 6–18. https://doi.org/10.1007/s005000050002.

Borovykh, A., Bohte, S. & Oosterlee, C. W. 2017 Conditional Time Series Forecasting with Convolutional Neural Networks. *Machine Learning* Cornell University Library. https://arxiv.org/pdf/1703.04691.pdf.

Breiman, L. 1996 Bagging predictors. *Machine Learning* **24** (2), 123–140. https://doi.org/10.1023/A:1018054314350.

Breiman, L. 2001 Random forest. *Machine Learning* **45** (1), 5–32. doi:10.1023/A:1010933404324.

Brentan, B. M., Luziovoto-Jr., E., Herrera, M., Izquierdo, J. & Pérez-Garcia, R. 2016 Hybrid regression model for near real-time urban water demand forecasting. *Journal of Computational and Applied Mathematics*. doi:10.1016/j.cam.2016.02.009.

Broad, D. R., Dandy, G. C. & Maier, H. R. 2005 Water distribution system optimization using metamodels. *Journal of Water Resources Planning Management* **131** (3), 172–180. doi:10.1061/(ASCE)0733-9496(2005)131:3(172).

Broomhead, D. H. & Lowe, D. 1988 Multivariable functional interpolation and adaptive networks. *Complex Systems* **2**, 321–355.

Caiado, J. 2010 Performance of combined double seasonal univariate time series models for forecasting water demand. *Journal of Hydrologic Engineering* **15** (3), 215–222. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000182.

Chang, H., Parandvash, H. & Shandas, V. 2010 Spatial variations of single family residential water consumption in Portland, Oregon. *Urban Geography* **31**, 953–972. doi:10.2747/0272-3638.31.7.953.

Chen, K.-Y. 2011 Combining linear and nonlinear model in forecasting tourism demand. *Expert Systems with Applications* **38** (8), 10368–10376. https://doi.org/10.1016/j.eswa.2011.02.049.

Chen, G., Long, T., Xiong, J. & Bai, Y. 2017 Multiple random forest modelling for urban water consumption forecasting. *Water Resources Management* **31** (5), 4715–4729. https://doi.org/10.1007/s11269-017-1774-7.

Colorni, A., Dorigo, M. & Maniezzo, V. 1991 Distribuited optimization by ant colonies. In *European Conference of Artificial Life*. Elsevier Publishing, Paris, pp. 134–142.

Crone, S. F. & Kourentzse, N. 2010 Feature selection for time series prediction – A combined filter and wrapper approach for neural networks. *Neurocomputing* **73**, 1923–1936. doi:10.1016/j.neucom.2010.01.017.

Dahl, C. M. & Hylleberg, S. 2004 Flexible regression models and relative forecast performance. *International Journal of Forecasting* **20** (2), 201–207. doi:10.1016/j.ijforecast.2003.09.002.

Di, C., Yang, X. & Wang, X. 2014 A four-stage hybrid model for hydrological time series forecasting. *PLoS One* **9** (8), 1–18. https://doi.org/10.1371/journal.pone.0104663.

Dietterich, T. G. 2000 Ensemble methods in machine learning. *Multiple Classifier Systems Lecture Notes in Computer Science* **1857**, 1–15. https://doi.org/10.1007/3-540-45014-9_1.

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. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000314.

Du, K., Zhao, Y. & Lei, J. 2017 The incorrect usage of singular spectral analysis and discrete wavelet transform in hybrid models to predict hydrological time series. *Journal of Hydrology* **552**, 44–51. doi:10.1016/j.jhydrol.2017.06.019.

Elman, J. L. 1990 Finding structure in time. *Cognitive Science* **14** (2), 179–211. https://doi.org/10.1207/s15516709cog1402_1.

Feo, T. A. & Resende, M. G. C. 1989 A probabilistic heurist for a computationally difficult set covering problem. *Operation Research Letters* **8** (2), 67–71. doi:10.1016/0167-6377(89)90002-3.

Firat, M., Turan, M. E. & Yurdusev, M. A. 2009a Comparative analysis of fuzzy inference systems for water consumption time series prediction. *Journal of Hydrology* **374**, 235–241. https://doi.org/10.1002/wrcr.20517.

Firat, M., Yurdusev, M. A. & Turan, M. E. 2009b Evaluation of artificial neural network techniques for municipal water consumption modeling. *Water Resources Management* **23**, 617–632. doi:10.1007/s11269-008-9291-3.

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. http://dx.doi.org/10.1016/j.jhydrol.2010.01.005.

Freitas, A. A. C. 2007 *Previsão de* séries temporais via seleção de variáveis, reconstrução dinâmica, ARMA-GARCH e redes neurais artificias. *Tese (Doutorado)*, Universidade Estadual de Campinas, Campinas. (in Portuguese).

Freund, Y. & Schapire, R. E. 1996 Experiments with a new boosting algorithm. In: *Proceedings of the International Conference on Machine Learning*, pp. 148–156. doi:10.1.1.29.3868.

Froelich, W. 2015 Forecasting Daily Urban Water Demand Using Dynamic Gaussian Bayesian Network. In: *Beyond Databases, Architectures and Structures. BDAS 2015*, Vol. 521 (S. Kozielski, D. Mrozek, P. Kasprowski, B. Małysiak-Mrozek & D. Kostrzewa, eds). Communications in Computer and Information Science. Springer, Cham. https://doi.org/10.1007/978-3-319-18422-7_30

Fullerton Jr, T. M., Ceballos, A. & Walke, A. G. 2016 Short-term forecasting analysis for municipal water demand. *Journal American Water Works Association* **108** (1), 27–38. http://dx.doi.org/10.5942/jawwa.2016.108.0003.

Gagliardi, F., Alvisi, S., Kapelan, Z. & Franchini, M. 2017 A probabilistic short-term water demand forecasting model based on the Markov Chain. *Water* **9**, 507. doi:10.3390/w9070507.

Gardiner, V. & Herrington, P. 1990 *Water Demand Forecasting*. Spon Press, Norwich, U.K.

Gashler, M., Giraud-Carrier, C. & Martinez, T. 2008 Decision Tree Ensemble: Small Heterogeneous Is Better Than Large Homogeneous. In: *7th International Conference on Machine Learning and Applications*, 11–13 Dec., San Diego, USA. doi:10.1109/ICMLA.2008.154.

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

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

Granger, C. W. J. 1989 Combining forecast – twenty years later. *Journal of Forecasting* **8**, 167–173. doi:10.1002/for.3980080303.

Granitto, P. M., Verdes, P. F. & Ceccatto, H. A. 2005 Neural network ensembles: evaluation of aggregation algorithms. *Artificial Intelligence* **163**, 139–162. https://doi.org/10.1016/j.artint.2004.09.006.

Gupta, I., Gupta, A. & Khanna, P. 1999 Genetic algorithm for optimization of water distribution systems. *Environmental*

Modelling & Software **14**, 437–446. https://doi.org/10.1016/S1364-8152(98)00089-9.

Guyon, I. & Elisseeff, A. 2003 An introduction to variable and feature selection. *Journal of Machine Learning Research* **3**, 1157–1182. http://dx.doi.org/10.1162/153244303322753616.

Hansen, L. K. & Salamon, P. 1990 Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12** (10), 993–1001. doi:10.1109/34.58871.

Hassani, H. 2007 Singular spectrum analysis: methodology and comparison. *Journal of Data Science* **5**, 239–257. doi:RePEc:pra:mprapa:4991.

Herrera, M., Torgo, L., Izquierdo, J. & Pérez-García, R. 2010 Predictive models for forecasting hourly urban water demand. *Journal of Hydrology* **387**, 141–150. doi:10.1016/j.jhydrol.2010.04.005.

Holland, J. H. 1975 *Adaptation in Natural and Artificial System*. University of Michigan Press, Ann Arbor, Mich.

Hopfield, J. J. 1982 Neural networks and physical systems with emergent colletive computational abilities. *Proceedings of National Academy of Science* **79** (8), 2554–2558. https://doi.org/10.1073/pnas.79.8.2554.

House-Peters, L., Pratt, B. & Chang, H. 2010 Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in Hillsboro, Oregon. *Journal of the American Water Resource Association* **46** (3), 461–472. doi:10.1111 / j.1752-1688.2009.00415.x.

Hsu, H. H., Hsieh, C. C. H. & Lu, M. D. 2011 Hybrid feature selection by combining filters and wrappers. *Expert Systems with Applications* **38**, 8144–8150. doi:10.1016/j.eswa.2010.12.156.

Htike, K. K. & Khalifa, O. O. 2010 Rainfall Forecasting Models Using Focused Time-Delay Neural Networks. In: *International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, Malaysia. doi:10.1109/ICCCE.2010.5556806.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society* **454**, 903–995. doi:10.1098/rspa.1998.0193.

Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L. & Savić, D. A. 2014 Dealing with uncertainty in water distribution system models: a framework for real-time modeling and data assimilation. *Journal of Water Resources Planning and Management* **140** (2), 169–183. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000325.

Jaeger, H. 2001 *The 'Echo State' Approach Analyzing and Training Recurrent Neural Networks*. GMD Report 148, German National Research Institute for Computer Science.

Jang, J.-S. R. 1991 Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm. In: *Proceedings of the 9th National Conference on Artificial Intelligence*, Anaheim, CA, USA, Vol. 2, pp. 762–767.

Jang, J.-S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics* **23** (3). doi:10.1109/21.256541.

John, H. G., Kohavi, R. & Pfleger, K. 1994 Irrelevant features and the subset selection problem. *Machine Learning: Proceedings of the Eleventh International conference* 121–129. doi:10.1.1.30.3875.

Jordan, M. I. 1986 *Serial Order: A Parallel Distributed Processing Approach*. Technical Report 8604, Institute Cognitive Science. University of California, San Diego.

Karunasingha, D. S. K. & Liong, S. Y. 2018 Enhancement of chaotic hydrological time series prediction with real-time noise reduction using Extended Kalman Filter. *Journal of Hydrology* **565**, 737–746. doi:10.1016/j.jhydrol.2018.08.044.

Kay, S. M. & Marple, S. L. 1981 Spectrum analysis – a modern perspective. *Proceedings of the IEEE* **69** (11), 1380–1419. doi:10.1109/PROC.1981.12184.

Kohavi, R. & John, G. H. 1997 Wrappers for feature subset selection. *Artificial Intelligence* **7**, 273–324. https://doi.org/10.1016/S0004-3702(97)00043-X.

Koza, J. R. 1990 *A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems*. Computer Science Department, Stanford University, Margaret Jacks Hall, Stanford, Cali.

Krogh, A. & Vedelsby, J. 1995 Neural network ensembles, cross validation, and active learning. *Advances in Neural Information Processing Systems* **7**, 231–238.

Li, W. & Huicheng, Z. 2010 Urban water demand forecasting based on HP filter and fuzzy neural network. *Journal of Hydroinformatics* **12** (2), 172–184.

Liu, J. Q., Cheng, W. P. & Zhang, T. Q. 2013 Principal factor analysis for forecasting diurnal water-demand pattern using combined Rough-set and Fuzzy-Clustering technique. *Journal of Water Resource Planning Management* **139** (1), 23–33. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000223.

Maria André, D. & Carvalho, J. R. 2014 Spatial determinants of urban residential water demand in Fortaleza, Brazil. *Water Resources Management* **28**, 2401–2414. doi:10.1007/s11269-014-0551-0.

Melin, P., Soto, J., Castillo, O. & Soria, J. 2012 A new approach for time series prediction using ensembles of ANFIS models. *Expert Systems with Applications* **39**, 3494–3506. https://doi.org/10.1016/j.eswa.2011.09.040.

Mendes-Moreira, J., Soares, C., Jorge, A. L. M. & Sousa, J. F. 2012 Ensemble approaches for regression: a survey. *ACM Computing Surveys* **45** (1), 40. doi:10.1145/2379776.2379786.

Nasseri, M., Moeini, A. & Tabesh, M. 2011 Forecasting monthly urban water demand using extended kalman filter and genetic programming. *Expert Systems with Applications* **38**, 7387–7395. doi:10.1016/j.eswa.2010.12.087.

Nauges, C. & Whittington, D. 2010 Estimation of water demand in developing countries: an overview. *The World Bank Research Observer* **25**, 2.

Oba, S., Sato, M., Takemasa, I., Monden, M., Matsubara, K. & Ishii, S. A. 2003 Bayesian missing value estimation method.

Bioinformatics **19** (16), 2088–2096. doi:10.1093/bioinformatics/btg287.

Odan, F. K. 2013 *Estudo de Confiabilidade Aplicado à Otimização da Operação em Tempo Real de Redes de Abastecimento de água*. *Tese (Doutorado)*, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. (in Portuguese)

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**, 245–256. doi:10.1061/(ASCE)WR.1943-5452.0000177.

Oord, A. V. D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrener, N., Senior, A. & Kavukcuoglu, K. 2016 WaveNet: A Generative Model for Raw Audio. *Sound* Cornell University Library.arXiv:1609.03499.

Oshima, N. & Kosuda, T. 1998 Distribution reservoir control with demand prediction using deterministic chaos method. *Water Science Technology* **37** (12), 389–395. https://doi.org/10.1016/S0273-1223(98)00378-3.

Peña-Guzmán, C., Melgarejo, J. & Prats, D. 2016 Forecasting water demand in residential, commercial, and industrial zones in Bogotá, Colombia, using Least-Squares Support Vector Machines. *Mathematical Problems in Engineering*. http://dx.doi.org/10.1155/2016/5712347.

Perrone, M. P. & Cooper, L. N. 1993 When networks disagree: Ensemble methods for hybrid neural networks. In: *Neural Networks for Speech and Image Processing Chapter 10* (R. J. Mammone, ed.). Chapman-Hall.

Piramuthu, S. 2004 Evaluating feature selection methods for learning in data mining applications. *European Journal of Operational Research* **156** (2), 483–494. https://doi.org/10.1016/S0377-2217(02)00911-6.

Poli, I. & Jones, R. D. 1994 A neural net model for prediction. *Journal of American Statistical Association* **89** (425), 117–121. doi:10.1080/01621459.1994.10476451.

Pulido-Calvo, I. & Gutiérrez-Estrada, J. C. 2009 Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems Engineering* **102**, 202–218. https://doi.org/10.1016/j.biosystemseng.2008.09.032.

Qiu, X., Zhang, L., Ren, Y., Suganthan, P. N. & Amaratunga, G. 2014 Ensemble deep learning for regression and time series forecasting. In: *Paper Presented at the Symposium on Computational Intelligence in Ensemble Learning (CIEL)*. doi:10.1109/CIEL.2014.7015739.

Ren, Y., Zhang, L. & Suganthan, P. M. 2016 Ensemble classification and regression-recent developments, applications and future directions. *IEEE Computational Intelligence Magazine* **11**, 41–53. doi:10.1109/MCI.2015.2471235.

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

Rooney, N., Patterson, D., Anand, S. & Tsymbal, A. 2004 Dynamic integration of regression models. *Proceedings of the International Workshop on Multiple Classifier Systems. Lecture Notes in Computer Science* **3181**, 164–173. https://doi.org/10.1007/11871842_82.

Sayes, Y., Inza, I. & Larrañga, P. 2007 A review of feature selection techniques in bioinformatics. *Bioinformatics* **23** (19), 2507–2517. https://doi.org/10.1093/bioinformatics/btm344.

Scheich, J. & Hillenbrand, T. 2009 Determinants of residential water demand in Germany. *Ecological Economics* **68**, 1756–1769. https://doi.org/10.1016/j.ecolecon.2008.11.012.

Shabri, A. & Samsudin, R. 2015 Empirical mode decomposition–least squares support vector machine based for water demand forecasting. *International Journal of Advances in Soft Computing and its Application*. **7** (2). ISSN 2074-2827

Sharkey, A. J. C. 1996 On combining artificial neural nets. *Journal Connection Science* **8** (3–4), 299–314. https://doi.org/10.1080/095400996116785.

Shepherd, A. J. 1997 *Second-order Methods for Neural Network: Fast and Reliable Methods for Multi-Layer Perceptrons*. Springer, New York.

Sollich, P. & Krogh, A. 1996 Learning with ensembles: How overfitting can be useful. In: *Advances in Neural Information Processing Systems 8* (D. Touretzky, M. Mozer & M. Hasselmo, eds). The MIT Press, pp. 190–196.

Sorjamaa, A., Hao, J., Reyhani, N., Ji, Y. & Lendasse, A. 2007 Methodology for long-term prediction of time series. *Neurocomputing* **70**, 2861–2869. https://doi.org/10.1016/j.neucom.2006.06.015.

Specht, D. F. 1990 Probabilistic neural networks. *Neural Network* **3** (1), 109–118. doi:10.1016/0893-6080(90)90049-Q.

Stańczyk, U. 2015 Feature evaluation by filter, wrapper, and embedded approaches. In: *Feature Selection for Data and Pattern Recognition. Studies in Computacional Intelligence*, Vol. 584 (U. Stańczyk & L. C. Jain, eds). Springer, pp. 29–44.

Tiwari, M. & Adamowski, J. 2015 Medium term water demand forecasting with limited data using an ensemble Wavelet-Bootstrap-Machine-Learning approach. *Journal of Water Resources Planning and Management* **141** (2). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000454.

Tiwari, M., Adamowski, J. & Adamowski, K. 2016 Water demand forecasting using extreme learning machines. *Journal of Water and Land Development* **28**, 37–52. doi:10.1515/jwld-2016-0004.

Tsoi, A. C. & Back, A. D. 1994 Locally recurrent globally feedforward networks: a critical review of architectures. *IEEE Transactions Neural Network* **5** (4), 229–239. doi:10.1109/72.279187.

Tsonis, A. A. 1992 *Chaos: From Theory to Applications*. Springer, New York.

Tsutiya, M. T. 2006 *Water Supply*, 3rd edn. Department of Hydraulic and Sanitary Engineering of the Polytechnic School of the University of São Paulo, São Paulo. (in Portuguese)

Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. & Lang, K. J. 1989 Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37** (3), 328–339. doi:10.1109/29.21701.

Wang, L., Wang, Z., Qu, H. & Liu, S. 2018 Optimal forecast combination based on neural network for time series forecasting. *Applied Soft Computing* **66**, 1–17. https://doi.org/10.1016/j.asoc.2018.02.004.

Webb, G. I. & Zheng, Z. 2004 Multi-strategy ensemble learning: reducing error combining ensemble learning techniques. *IEEE Transactions on Knowledge and Data Engineering* **16** (8), 980–991. doi:10.1109/TKDE.2004.29.

Wentz, E. A. & Gober, P. 2007 Determinants of small area water consumption for the city of Phoenix, Arizona. *Water Resources Management* **21**, 1849–1863. doi:10.1007/s11269-006-9133-0.

Wichard, J. D., Merkwirth, C. & Ogorzałek, M. 2003 *Building Ensembles with Heterogeneous Models*. Available from: www.j-wichard.de/publications/salermo_Incs_2003.pdf (accessed 22 August 2018).

Wilamowski, B. M. & Yu, H. 2010 Neural network learning without backpropagation. *IEEE Transactions on Neural Networks* **21** (11), 1793–1803. doi:10.1109/TNN.2010.2073482.

Williams, R. J. & Zipser, D. 1989 A learning algorithm for continually running fully recurrent neural networks. *Neural Computational* **1** (2), 270–280. https://doi.org/10.1162/neco.1989.1.2.270.

Xu, Y., Zhang, J., Long, Z. & Lv, M. 2018 Daily urban water demand forecasting based on chaotic theory and Continuous Deep Belief Neural Network. *Neural Processing Letters*. https://doi.org/10.1007/s11063-018-9914-5.

Zanchettin, C. 2008 *Otimização Global de Redes Neurais Artificiais. Tese (Doutorado)*, Universidade Federal de Pernambuco, Recife. (In Portuguese).

Zapelan, R. 2002 *Calibration of Water Distribution System Hydraulic Models*. PhD (Thesis), University of Exeter, UK.

Zhang, G. P. 2001 An investigation of neural networks for linear time-series forecasting. *Computers & Operations Research* **28** (12), 1183–1202. https://doi.org/10.1016/S0305-0548(00)00033-2.

Zhang, G. P. 2003 Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0.

Zhang, G. P. & Qi, M. 2005 Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research* **160** (2), 501–514. doi:10.1016/j.ejor.2003.08.037.

Zhang, G. P., Patuwo, B. D. & Hu, M. Y. 1998 Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting* **14**, 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7.

Zhao, P. & Zhang, H. W. 2008 Chaotic characters and forecasting of urban water consumption. *China Water Wastewater* **24** (5), 90–94. https://doi.org/10.1007/s11063-018-9914-5.

First received 3 March 2019; accepted in revised form 12 August 2019. Available online 29 August 2019