Water Resources Management
https://doi.org/10.1007/s11269-019-02213-y

A Comparison of Short-Term Water Demand Forecasting Models

E. Pacchin[1] · F. Gagliardi[1] · S. Alvisi[1] · M. Franchini[1]

Received: 5 November 2018 / Accepted: 3 February 2019

Abstract
This paper presents a comparison of different short-term water demand forecasting models. The comparison regards six models that differ in terms of: forecasting technique, type of forecast (deterministic or probabilistic) and the amount of data necessary for calibration. Specifically, the following are compared: a neural-network based model (ANN_WDF), a pattern-based model (Patt_WDF), two pattern-based models relying on the moving-window technique (αβ_WDF and Bakk_WDF), a probabilistic Markov chain-based model (HMC_WDF) and a naïve benchmark model. The comparison is made by applying the models to seven real-life cases, making reference to the water demands observed over 2 years in district-metered areas/ water distribution networks of different sizes serving a different number and type of users. The models are applied in order to forecast the hourly water demands over a 24-h time horizon. The comparison shows that a) models based on different techniques provide comparable, medium-high forecasting accuracies, but also that b) short-term water demand forecasting models based on moving-window techniques are generally the most robust and easier to set up and parameterize.

Keywords Water demand · Short-term forecasting · Moving window

1 Introduction

Water demand forecasting provides a valid contribution to the design and optimal management of water distribution systems. For example, both the design and construction of new distribution networks and the expansion or upgrading of existing networks require substantial investments, and there is thus a need to conduct preliminary assessments that take into account the long-term development of the area involved in terms of water demand. Similarly, the management of the installations and facilities serving supply and distribution networks (e.g. water treatment plants, pumping stations, etc.) and control of the networks themselves, as well as of the devices

✉ S. Alvisi
stefano.alvisi@unife.it

1 Department of Engineering, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy

installed in them (e.g. valves), can be optimised based on knowledge of the entity of future water demands over short time horizons.

Depending on the levels of planning at which they are used, water demand forecasting models can be distinguished according to the forecast horizon (i.e. the time interval in which a forecast is made) and forecast frequency (i.e. the time step at which water demand forecasts are generated within the time horizon) (see Donkor et al. (2014) and Ghalehkhondabi et al. (2017) for a general review of water demand forecasting models). Long-term models generally provide demand forecasts on a yearly or monthly basis with a time horizon ranging from 1 to 10 years and are mainly used for design purposes or for allocating resources (Babel and Shinde 2011). Short-term models, by contrast, forecast water demand over more limited time horizons, ranging from 1 month to 1 day, with a time step ranging from daily to sub-hourly and are mainly used for management purposes (Arandia et al. 2016; Bakker et al. 2013; Msiza et al. 2008; Shabani et al. 2018).

In this paper we focus on short-term forecasting models. These models can be classified on the basis of the techniques they use to generate the forecast itself, the type of forecast provided (deterministic or probabilistic), and the information that needs to be gathered in order to develop the model prior to its application. As regards the technique used, it is possible to identify a first category which includes all the models based on data-driven techniques, such as multilinear and nonlinear regression (Adamowski et al. 2012), artificial neural networks (Jain et al. 2001; Anele et al. 2017), support vector machines (Msiza et al. 2008), random forests (Chen et al. 2017), project pursuit regression (Herrera et al. 2010) and genetic expression programming (Shabani et al. 2018). Within this category, the models based on artificial neural networks (ANN) (Romano and Kapelan 2014) take on particular relevance. They have been widely used in the scientific literature to develop water demand forecasting models and compared with other forecasting models, all similarly based on different data-driven techniques (see, for example, Herrera et al. 2010; Adamowski et al. 2012).

The second category includes forecasting models based on the recognition of periodic patterns, in which various techniques of time series analysis are exploited with the aim of simulating the patterns that generally characterise water demands over different periods of time. Zhou et al. (2000, 2002) and Gato et al. (2007) use a water demand forecasting method that distinguishes between a base component and a seasonal component. Alvisi et al. (2007) provide a daily and hourly water demand forecast derived by adding a persistence component, modelled using regression techniques, to seasonal, weekly and daily patterns. Caiado (2010) exploits various techniques such as Holt-Winter, ARIMA and generalized autoregressive conditional heteroskedasticity (GARCH) for pattern recognition. Finally, Bakker et al. (2013) and Pacchin et al. (2017) take into account, at the forecasting stage, the periodicities in demand determined through factors calibrated on the basis of a moving window of observed data.

Considering the type of output provided, it should be noted that although the majority of short-term water demand forecasting models proposed in the literature are of the deterministic type, stochastic models have also been recently proposed. Among them, it is worth mentioning the Bayesian model developed by Hutton and Kapelan (2015), a cascade model aimed at quantifying and reducing the uncertainty of forecasting models, and the model proposed by Cutore et al. (2008), which consists in the application of the SCEM-UA algorithm (Vrugt et al. 2003) to calibrate the parameters of a neural network and estimate the uncertainty associated with the latter in order to estimate, in addition to forecasts of a deterministic type, the uncertainty of the model itself. The model proposed by Alvisi and Franchini (2017), also adopted by Anele et al. (2018), makes use of the model conditional processor (MCP) (Todini 2008), which, by combining the performances of different forecasting models, enables an estimation of predictive uncertainty. Finally, the Markov chain-based model proposed by Gagliardi et al. (2017) provides an estimate of the probabilities that future demands will fall within pre-assigned ranges.

Considering, finally, the observed data that must necessarily be available before the models themselves can be applied, it may be noted that water demand forecasting models need to undergo an initial parameter calibration process, which is carried out using a set of observed data (Bakker et al. 2013). The size of the dataset will vary according to the structure on which the models are based. For example, neural network-based models must be put through an initial training stage in which the network parameters are calibrated (weights and bias). The length of the set of observed data to be used for calibration is not fixed a priori, but must be sufficient to ensure that the variability of the water demands is taken fully into account since an ANN model does not have the capability to extrapolate outside the range of data employed for training (Zubaidi et al. 2018). In the case of models based on pattern reproduction, a calibration needs to be performed on the basis of observed data to enable an estimation of the factors characterising the periodic patterns. In order to provide a complete estimation of these factors, the majority of models must be calibrated using a set containing at least 1 year of observations. Indeed, 1 year is the minimum time window necessary in order to observe both the short term (i.e. daily and weekly) fluctuations of the water consumptions and the long term (i.e. seasonal) oscillations (Zhou et al. 2000; Alvisi et al. 2007).

In contrast, the models based on a moving window of data (Bakker et al. 2013; Pacchin et al. 2017), by their very nature, do not require an ample, fixed dataset for calibration purposes. In fact, these models generate a parameter estimate based on the observed data in a moving window, typically a few weeks long, which moves forward together with the forecasting time. Therefore, unlike the models that require a calibration process, in which the parameters are estimated prior to their real-time use, in moving window-based models the parameters are updated at every time step.

This paper presents a comparison of short-term water demand forecasting models whose features vary greatly in terms of the techniques they are based on (data-driven and pattern-based), the type of forecast provided (deterministic or probabilistic) and the information that needs to be gathered for the purpose of fine-tuning the model itself prior to its application. The aim is to highlight the pros and cons of the various approaches and thus provide useful information about the type and structure of model to be used to set up a short term water demand forecasting model. Specifically, the following models are compared: a neural network-based model, a pattern-based model, two models with a moving window structure, a Markov chain-based model and, finally, a benchmark model based on a naïve approach. These models are applied in order to forecast hourly water demands over a 24-h time horizon. The comparison is made by applying the models to seven case studies, making reference to water distribution networks or district-metered areas of different sizes.

A brief description of each of the six models applied to predict hourly demands is given below (section 2). The seven case studies are presented in section 3, along with a description of their main features. The results of the application of the six models in the seven case studies are then analysed and discussed (section 4). The paper concludes with some final considerations (section 5).

2 Models Compared

The short-term water demand forecasting models compared in this study are the following: two deterministic models requiring a preliminary calibration, one based on an artificial neural network (Alvisi and Franchini 2017), hereinafter identified as ANN_WDF, and one based on the reproduction of periodic water demand patterns (Alvisi et al. 2007), hereinafter identified as Patt_WDF; two models, similarly of the deterministic type, but based on the use of a moving window of previously observed data, specifically, the model proposed by Bakker et al. (2013), hereinafter identified as Bakk_WDF, and the model proposed by Pacchin et al. (2017), hereinafter identified as αβ_WDF; finally, a probabilistic model based on the use of the Markov chains (Gagliardi et al. 2017), hereinafter identified as HMC_WDF. It is specified that, in order to compare the latter model with the above-mentioned deterministic models, its ability to provide results of a probabilistic type is not exploited or assessed within the framework of this study; that is, the study does not consider the confidence interval it produces in relation to the forecast provided.

A sixth benchmark model of the naïve type (Gagliardi et al. 2017) was also considered by way of comparison. A brief description of each model is provided below. The reader is referred to the corresponding original publications for further information about each of them.

2.1 ANN_WDF Model

The ANN_WDF model is based on the use of artificial neural networks (Alvisi and Franchini 2017). Such networks draw inspiration from biological neural networks and their ability to receive and analyse incoming signals and produce output signals. The most common types of artificial neural networks include the multilayer perceptron (MLP), in which the neurons are organized in layers: the first layer (*input layer*) receives incoming information and, after appropriately weighting the information received, transfers it to one or more intermediate layers (*hidden layers*) where the information is processed by means of predefined functions before being delivered to the output layer (Romano and Kapelan 2014). More specifically, the neural network model adopted here is aimed at forecasting hourly water demands over a time horizon of $K = 24$ h; it is based on a three-layer feed-forward MLP neural network characterised by a single hidden layer. Every hour the network receives, as input, data related to the observed demands of the previous 24 h and a binary index identifying the type of day (weekday or weekend day) and outputs are the demand forecast for the next 24 h. The number of the neurons making up the hidden layer is set in the model calibration phase; the aim is to identify the smallest number of neurons that can be used without penalizing the forecasting accuracy (Hsu et al. 1995). A log sigmoid transfer function is used in the hidden layer and a pure linear one in the output layer. The network parameters, weights and bias are estimated during network calibration using the Levemberg Marquardt algorithm (Hagan and Menhaj 1994). In particular, in order to prevent overfitting in the calibration phase, the early stopping technique is used and the calibration dataset is divided into two subsets containing 80% and 20% of the data, respectively; the first subset is used for training and the second for testing the network. In order to avoid the risk of signal saturation (Hsu et al. 1995), the data are normalized and scaled in such a way as to belong to the interval [0:1]. The normalization is performed using the mean and standard deviation of demands in the

24 h of the day, calculated using the calibration dataset, with a distinction being made between weekdays and weekend days. The outputs provided by the network then undergo a process of de-normalization and de-scaling.

2.2 Patt_WDF Model

The Patt_WDF model (Alvisi et al. 2007) is structured in such a way as to provide a forecast of water demands for the next $K = 24$ h based on a reproduction of the periodic patterns characterising the water demand time series, namely, (a) a seasonal and weekly cyclical pattern of daily water demands and (b) a daily cyclical pattern of hourly water demands, and on the reproduction of persistence phenomena. In greater detail, the model is divided into two modules, a daily one and an hourly one. In the first module, a forecast is made of the mean daily water demand $Q_m^{d,for}$ of the Julian day (or days) m (with $m =$ 1,2,..,365) in which the 24 h of the forecast fall, taking into account the seasonal and weekly cyclical patterns and short-term persistence, using the following formula:

$$Q_m^{d,for} = Q_m^{d,F} + \Delta_{i,s}^d + \delta_m^d \qquad (1)$$

where $Q_m^{d,F}$ represents the seasonal component modelled by means of a Fourier series,$\Delta_{i,s}^d$ is a correction factor that takes into account the weekly periodicities, i being the day of the week (with $i = 1,2,..,7$, Monday, Tuesday,.., Sunday) and s the season (with $s = 1,2,3,4$, winter, spring, summer, autumn) corresponding to the Julian day m and δ_m^d a correction factor that takes into account the short-term daily persistence represented by means of an autoregressive model AR(1) (Box et al., 1994).

In the hourly module an estimate is made of the average hourly water demand $Q_{t+k}^{h,for}$ for k hours ahead of the current hour t (with $k =$ 1,2,..,K), obtained as the sum of the mean daily water demand estimated in the daily module, $Q_m^{d,for}$, the daily periodicity component, represented by the hourly correction factor $\Delta_{j,i,s}^h$, j being the hour of the day (with $j =$ 1,2,..,24), i the day of the week and s the season corresponding to the forecasted hour $t + k$, and an error ε_{t+k}, which takes into account the short-term hourly persistence modelled by means of a regression process, taking into account the errors observed one and 24 h before the current forecast time t:

$$Q_{t+k}^{h,for} = Q_m^{d,for} + \Delta_{j,i,s}^h + \varepsilon_{t+k} \qquad (2)$$

All parameters of the model (seasonal component $Q_m^{d,F}$, correction factor that takes into account the weekly periodicity $\Delta_{i,s}^d$, hourly correction factor $\Delta_{j,i,s}^h$, coefficients of the AR(1) models and of the regression, which represent the persistence components) are estimated in the calibration phase using a dataset containing the observed demands relating to a period of at least a year, and subsequently applied to the validation set. At least 1 year of observed data is necessary to fully capture the seasonal periodic behaviour of water consumptions modelled by means of a Fourier series (seasonal component $Q_m^{d,F}$ in Eq. 1) (Zhou et al. 2000) and to properly characterize the weekly (see Eq. 1) and hourly factors (see Eq. 2) which, as well, depend on the season s (Alvisi et al. 2007).

JrnlID 11269_ArtID 2213_Proof# 1 - 12/02/2019

2.3 Bakk_WDF Model

In its original version, the Bakk_WDF model is designed to be used to forecast the average water demand over a time horizon of 48 h with a 15-min time step (Bakker et al. 2013). However, in this study it was decided to use it to forecast hourly water demands over a time horizon of $K = 24$ h, as in the case of the other models, first of all so that a fair comparison could be made and, moreover, because the observed data consisted of historical hourly series.

The model is based on a procedure that can be divided into three steps: in step 1 the average water demand for the next 24 h is determined; in step 2 the average water demand for each step of the forecast horizon is estimated; and in step 3 the entity of the hourly water demand referred to as "extra sprinkle water demand" is estimated, where applicable; the latter relates to a particular use of potable water (i.e. for gardening) in the evening hours of some days of the year.

More specifically, in step 1 the average water demand for the next 24 h after the forecasting time t ($Q_t^{d,for,corr}$) is forecast based on the mean of the hourly demands observed in the previous 48 h, duly corrected:

$$Q_t^{d,for,corr} = C_1 \cdot \left(\sum_{g=t-24}^{t-1} Q_g^{h,obs,corr} \right) + C_2 \cdot \left(\sum_{g=t-48}^{t-25} Q_g^{h,obs,corr} \right) \tag{3}$$

where C_1 and C_2 are two constants and $Q_g^{h,obs,corr}$ are the hourly demands observed in the previous 48 h duly corrected by means of a specific factor typical of the day of the week (see Bakker et al. 2013).

In step 2 the average hourly water demands $Q_{t+k}^{h,for,corr}$ are determined for the generic lead time k (with $k = 1,2,..,K$) based on the daily characterization, given by the coefficient f_i^d, and the hourly characterization, given by the coefficient $f_{i,j+k}^h$ i being the day of the week and j the hour of the day:

$$Q_{t+k}^{h,for,corr} = Q_t^{d,for,corr} \cdot f_i^d \cdot f_{i,j+k}^h \tag{4}$$

In step 3 the extra sprinkle water demand $Q_m^{sprink,for}$ is determined in the hour m in time frame between 18:00 and 0:00 h; once the days for which it is necessary to calculate this supplementary demand have been identified, the procedure is carried out in the same manner as in steps 1-2 for the standard water demand, but in this case using a characteristic coefficient f_m^{sprink}.

The total hourly water demand for every lead time k (with $k = 1,2,..,K$) of the horizon $K = 24$ h, is:

$$Q_{t+k}^{h,for,tot} = Q_{t+k}^{h,for,corr} + Q_m^{sprink,for} \tag{5}$$

In this study, as suggested in the parameter sensitivity analysis conducted by Bakker et al. (2013), it was chosen to adopt a time window of 5 weeks of observed data to determine the coefficient f_i^d and a window of 10 weeks for $f_{i,j}^h$ and f_m^{sprink}.

Similarly, as regards the values of the constants C_1 and C_2 used in the water demand forecasting procedure, it was chosen to use the ones indicated by Bakker et al. (2013), 0.8 and 0.2 respectively.

2.4 αβ_WDF Model

αβ_WDF is a model that provides a water demand forecast for the next $K = 24$ h based exclusively on the observed demands within a narrow interval preceding the time the forecast was made Pacchin et al. (2017). In fact, the model is based on a moving time window of observed data, within which it is possible to identify the characteristic patterns of the days making up the week; the window is characterised by a length of *NW* weeks and moves together with the forecasting time *t*. The forecasting procedure is made up of two steps: in the first step, an estimate is made of the average water demand over the forecast horizon, consisting of $K =$ 24 h; in the second step, based on the forecast made in the first step, the water demand of each of the 24 h of the forecast horizon is estimated by means of suitable hourly coefficients. More precisely, where *t* is the current hour in which the forecast is made and V the vector of *NW* hours corresponding to the same hour *j* of the day and the same type *i* of day of the week as the one in which *t* falls, i.e. $V = \{v_1; v_2; \ldots; v_{NW}\} = \{t-1\cdot 24; t-7\cdot 24\cdot 2; \ldots; t-7\cdot 24\cdot NW\}$, in the first step the average water demand $Q_t^{d,for}$ over the $K = 24$ h after the time *t* is estimated by means of the following relation:

$$Q_t^{d,for} = \alpha_t \cdot \overline{Q}_{t-24}^{d,obs} \tag{6}$$

where $\overline{Q}_{t-24}^{d,obs}$ is the average water demand observed in the 24 h preceding the hour *t* and α_t is a coefficient having a specific value for the 24-h horizon that begins at the time *t*:

$$\alpha_t = \frac{1}{NW} \cdot \sum_{v_{nw}=v_1}^{v_{NW}} \frac{\overline{Q}_{v_{nw}}^{d,obs}}{\overline{Q}_{v_{nw}-24}^{d,obs}} \tag{7}$$

where $\overline{Q}_{v_{nw}}^{d,obs}$ is the average water demand observed in the 24 h following the hour v_{nw} (with $nw = 1, 2, \ldots, NW$) and $\overline{Q}_{v_{nw}-24}^{d,obs}$ is the average water demand observed in the 24 h preceding the hour v_{nw} (i.e. da 24·7·*nw*-24 a 24·7·*nw*).

Once the average water demand $Q_t^{d,for}$ of the $K = 24$ h has been estimated, in the second step the hourly water demand $Q_{t+k}^{h,for}$ of the hour $t + k$ is estimated by means of the following relation:

$$Q_{t+k}^{h,for} = B_{t,k} \cdot Q_t^{d,for} \tag{8}$$

where $\beta_{t,\,k}$ is the coefficient characteristic of the lead time *k* (within the time horizon of $K = 24$ h) that starts at the time *t*:

$$\beta_{t,k} = \frac{1}{NW} \cdot \sum_{v_{nw}=v_1}^{v_{NW}} \frac{Q_{v_{nw}+k}^{h,obs}}{\overline{Q}_{v_{nw}}^{d,obs}} \tag{9}$$

where $Q_{v_{nw}+k}^{h,obs}$ is the hourly water demand in the *k*-th *hour* after the hour v_{nw} (with $nw = 1, 2, \ldots, NW$). It should be noted that at every forecasting time *t*, 24 values of the coefficient $\beta_{t,k}$ are calculated, one for every lead time *k*.

From an operational standpoint, the αβ_WDF model is applied using a moving window with a length of $NW = 4$ weeks so that the seasonal fluctuations in consumption can be characterised (Pacchin et al. 2017).

2.5 HMC_WDF Model

The homogeneous Markov model is based on the application of the statistical concept of homogeneous Markov chains to water demand forecasting (Gagliardi et al. 2017). In this model, the hourly water demand is identified as the variable of a discretized Markov process, in which it is possible to estimate, from a probabilistic viewpoint, the future states of the process once the current state is known and the tendency to transition into different pre-identified states at subsequent points in time. In general, depending on whether its tendency to transition from one state to another is time dependent or not, the process may be identified as non-homogeneous or homogeneous; it is therefore possible to formulate two different types of Markov models (non-homogeneous Markov chain – NHMC and homogeneous Markov chain – HMC model); in particular, it is possible to demonstrate that the HMC_WDF model provides greater forecasting accuracy (Gagliardi et al. 2017) and it was thus decided to apply it in this case. The periodicities that generally influence water demands (i.e. seasonal, weekly and daily) must be removed from the data processed by the HMC_WDF model; this is achieved by subjecting the original data to a de-seasonalization and normalization process. In the first stage, the daily demand of the Julian day m of the year $Q_m^{d,F}$, modelled using a Fourier series, is subtracted from every hourly demand Q_t^h of the original series:

$$Q_t^{h,des} = Q_t^h - Q_m^{d,F} \tag{10}$$

In the second stage, the demands are normalized on the basis of the mean values μ and standard deviation σ of the hourly observed data entered in the calibration phase (Gagliardi et al. 2017). More specifically, the mean values and standard deviations are defined by distinguishing each of the 24 h of the day and distinguishing weekdays (Mon-Fri) from weekend days (Sat-Sun) within the different seasons (since the daily pattern may vary in the different seasons, especially in the case of areas frequented by tourists). Once the data have been normalized, forecasting is performed by identifying a number *NC* of *classes* in the domain of variability of consumption. At this point it is possible to estimate the *NC* probabilities that the water demand in the instant following the current one will belong to each class, contained in the vector p_{t+1}^{for}:

$$p_{t+1}^{for} = p_t^{obs} \times \hat{\prod} \tag{11}$$

where p_t^{obs} is the probability vector representing the probabilities of the water demand belonging to the different classes at the current time, based on real observed data, and $\hat{\prod}$ is the transition matrix (time independent), which contains all the probabilities of demand transitioning from one class to another in consecutive instants, estimated during the calibration phase. It is possible to extend the forecast to lead times k greater than 1 by taking into account, at each instant in time, the forecast obtained at the preceding instant and iteratively applying Eq. 12:

$$p_{t+k}^{for} = p_{t+k-1}^{for} \times \hat{\prod} \text{con}\, k > 1 \tag{12}$$

On the basis of this probabilistic forecast, it is possible to obtain an expected value of the demands Q_{t+k}^{for} by computing a weighted average of the representative values of each class (e.g. the mean value of each class) contained in the vector $\mathbf{u} = [u_1, u_2, \ldots, u_{NC}]$ and using, as weights, the estimated probabilities:

$$Q_{t+k}^{for} = \sum_{nc=1}^{NC} u_{nc} \cdot p_{nc,t+k}^{for} \tag{13}$$

For the purpose of comparing the forecasting models considered in the present study, use was made only of the information obtained from the model, i.e. the information provided by Eq. 13.

Finally, it should be stressed that this model requires a parameterization stage in which the various parameters are estimated: the factors necessary for deseasonalization and normalization of the data and the transition matrix $\hat{\prod}$.

In practice, the HMC_WDF model is applied assuming a number of classes *NC* equal to 4.

2.6 Naïve Model

The naïve model has a decidedly simpler structure than all of the other models analysed and applied in this study. The naïve model adopted as the benchmark is defined in the literature as the 'mean' model (Gelažanskas and Gamage 2015). Indeed, in this case the forecast is based on the mean values $\mu = [\mu_1, \mu_2, \ldots, \mu_{24}]$ of the water demands associated with each of the 24 h of the day calculated on the basis of the calibration dataset. The water demand forecast for a generic hour j of the day is assumed to be equal to the corresponding mean demand μ_j. It may be deduced that the forecasting accuracy of the model is always the same, irrespective of the lead time.

3 Case Studies

The seven real-life cases (CS) considered relate to water distribution networks and district-metered areas in northern Italy varying both in size and in the number of users. Two years of observed data, recorded on an hourly basis, are available for each case considered. The years are identified as y1 and y2. Table 1 shows information regarding the number and type of users and the average water demand in the 2 years of monitoring for each CS.

The first six case studies refer to residential/industrial districts, whereas the seventh refers to a seaside resort characterised by considerable variability in the number of users over the course of the year. Furthermore, it is worth noting that in case studies 1, 2, 3 and 6 (CS1, CS2, CS3

Table 1 Average demand (L/s) and number and type of users for each case study

Case study (CS)	1	2	3	4	5	6	7
Number of users	120,000	20,000	9000	7000	7000	2500	300-3500
Type of users	Res/Ind	Res/Ind	Res/Ind	Res	Res	Res	Res/Tour
Average demand y1 [L/s]	952.4	180.0	101.0	56.0	54.2	24.9	36.2
Average demand y2 [L/s]	966.8	177.5	100.0	67.9	56.8	24.7	29.3

Res residential users, *Ind* industrial users, *Tour* touristic users and CS 6 respectively) the demands did not undergo substantial variations from the year y1 to the year y2. In case study 4 (CS4) the average water demand rose by 21.3% from y1 to y2; in case study 5 (CS5) the average water demand increases of about 5.0% whereas in case study 7 (CS7) a significant decrease in demand, about −19%, was observed between y1 (36.2 L/s) and y2 (29.3 L/s).

In the case of models requiring calibration (ANN_WDF, Patt_WDF, HMC_WDF and naïve), the first year of data (y1) was used for calibration purposes and the second year (y2) for validation, whereas the models Bakk_WDF and αβ_WDF, whose parameters are calculated at every forecasting step over a window of previously observed data, were applied directly to the sequence of the 2 years of observed data.

The performance of the models applied in the seven case studies was assessed for different lengths of forecast time horizon k in terms of mean absolute error (MAE%) and root mean square error (RMSE), defined as:

$$MAE\% = \frac{1}{nd}\sum_{i=1}^{nd}\left|\frac{e_i}{\mu_{obs}}\right|\cdot 100 \qquad (14)$$

$$RMSE = \sqrt{\frac{1}{nd}\sum_{i=1}^{nd} e_i^2} \qquad (15)$$

where nd is the number of data in the period considered (for example a year), $e = Q^{obs} - Q^{for}$ is the error, Q^{obs} is the value of the observed average hourly water demand, Q^{for} is the forecasted average hourly water demand and μ_{obs} is the mean of the observed values. The performances of all the models were assessed considering the results for the year y1 separately from those for the year y2.

4 Analysis and Discussion of Results

Figure 1 shows the trend in the MAE% associated with different lead times, for every model analysed, for the 2 years considered and for every case study. It may be observed, first of all, that all of the models provide better accuracy than the naïve model as regards both y1 and y2. When attention is focused on the differences found in every CS between y1 and y2, it may be observed that in y1 all of the models provide comparable levels of accuracy; in particular the mean percentage values range between 2% and 5% for CS 1-2-3-5-6, between 3.5% and 8% for CS4 and, finally, between 10% and 30% for CS7. It is worth pointing out immediately that in the latter case study the mean percentage errors are distinctly higher than in the other case studies, irrespective of the model. This finding may be explained by the fact that this CS makes reference to a seaside resort, in which the users and water demands are subject to high and sudden variations; therefore, all of the models provide less accurate forecasts. With regard to the year y2, the models tend to show a performance similar to that observed for the year y1, with a few differences. In general, the mean percentage error calculated for the models Patt_WDF, ANN_WDF and HMC_WDF increased; in fact, these are models requiring a calibration based on data observed over a long period and thus tend to provide greater accuracy for the year of calibration (y1) compared to the year of validation (y2). The models based on the moving-window technique (αβ_WDF and Bakk_WDF), by contrast, tend to maintain the same forecasting accuracy in both years.

The difference in behaviour between these two groups of models can also be noted in another respect. The models requiring long-term calibration tend to perform slightly better in the case of short time horizons, while their performance declines slightly and remains stable in

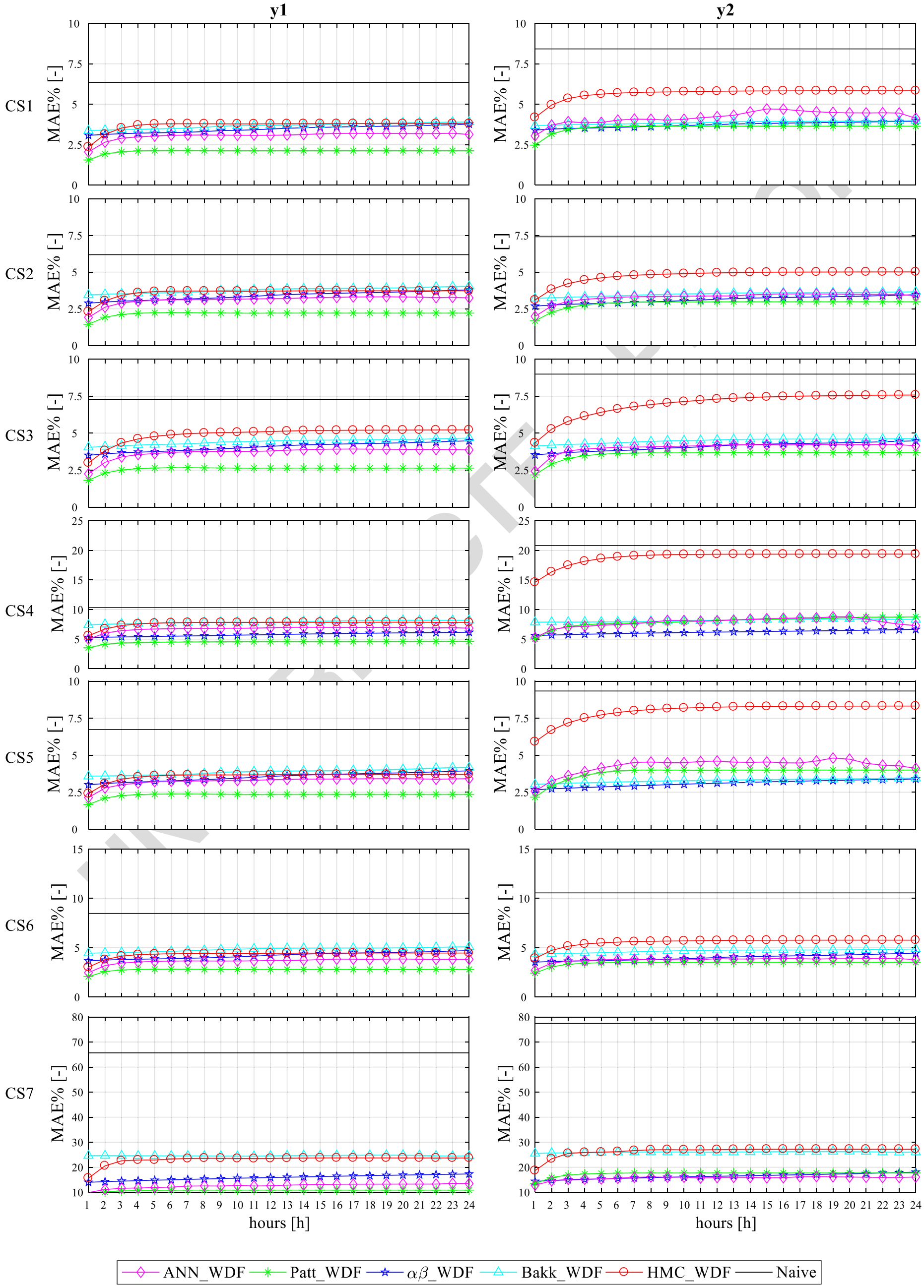

Fig. 1 Values of MAE% for every time horizon (k = 1,2,..24) in the 2 years considered (y1 and y2), for every case study analyzed (CS1,CS2,..CS7)

the case of long time horizons. In the moving-window models, by contrast, the accuracy shows to be more consistent, irrespective of the time horizon.

Again with reference to Fig. 1, a significant increase may be noted in the error associated with the HMC_WDF model in CS4 and CS5 in the year y2; as highlighted previously, these cases are characterised by a considerable difference in average demands in the years y1 and y2; it may be deduced, therefore, that the HMC_WDF model is significantly influenced by the variability in consumption. This problem is reflected to a less marked degree in the errors of the Patt_WDF and ANN_WDF models, which resulted in a decree in accuracy in the year y2.

The considerations set forth thus far are supported by the results shown in Table 2, which shows, for each model and case study, the difference between the MAE%, averaged over the time horizon, for the years y1 and y2. In general, negative values indicate a worsening in performance from y1 to y2, whereas positive values indicate an improvement.

It is possible to note that all models requiring long-term calibration (Patt_WDF, ANN_WDF, HMC_WDF and naïve) showed a negative difference in the MAE% in every CS, whereas in the case of the models based on the moving-window technique (αβ_WDF and Bakk_WDF), the difference in the MAE% is negative for CS1-3-4-7 and positive for the remaining case studies. Looking at the values contained in Table 2, considered in absolute terms, it may be noted that the largest difference in performance between y1 and y2 corresponds to the naïve and HMC_WDF models, with values equal to 11.73% (CS7) and 11.18% (CS4), respectively; the difference is lower for the Patt_WDF and ANN_WDF models, which show maximum differences (in absolute terms) equal to 6.68% (CS7) and 3.09% (CS7), respectively. Finally, the maximum differences in the MAE% shown for the 2 years by the models Bakk_WDF and αβ_WDF are smaller (in absolute terms), equal to 1.4% (CS7) and 0.5% (CS7), respectively.

Summarising, it may be affirmed that, on average, the model that delivered the best forecasting performance for the year y1 is Patt_WDF, though the differences compared to all the other models were minimal. In the year y2, a greater variability in forecasting accuracy was observed: in CS 1-2-3-6, the αβ_WDF, Bakk_WDF, Patt_WDF and ANN_WDF models provided excellent demand forecasts, in CS4 the model that performed best was αβ_WDF, in CS5 the αβ_WDF and Bakk_WDF models provided the highest accuracy and, finally, in CS7 the αβ_WDF, Patt_WDF and ANN_WDF models showed the best performance. Thus the same forecasting accuracy can be achieved using both data-driven and pattern-based techniques. On the other hand it is worth observing that the naïve model is undoubtedly the least refined of the models considered and represents the simplest method for making a forecast. Not coincidentally, compared to this model all of the other models produce an improvement in forecasting for both years, y1 and y2. The performance of the naïve model decreases

Table 2 Difference between the MAE%, averaged over the time horizon, of the years y1 and y2 for every CS and every model applied

$\overline{MAE\%}_{y1}-\overline{MAE\%}_{y2}$	CS1	CS2	CS3	CS4	CS5	CS6	CS7
ANN_WDF	−1.15	−0.18	−0.33	−1.09	−1.04	−0.11	−3.09
Patt_WDF	−1.47	−0.69	−0.96	−3.47	−1.51	−0.68	−6.68
Bakk_WDF	−0.19	0.28	−0.06	−0.25	0.59	0.16	−1.42
αβ_WDF	−0.27	0.26	−0.02	−0.42	0.44	0.22	−0.52
HMC_WDF	−1.97	−1.15	−2.05	−11.18	−4.39	−1.23	−3.34
Naïve_WDF	−2.07	−1.22	−1.72	−10.50	−2.60	−2.08	−11.73

drastically in the event of a strong variability in demand during the year, whereas the decrease is attenuated in the case of greater uniformity; this finding is consistent with the fact that the average value is more closely representative in relation to the range of possible values.

The same conclusions can be drawn from an analysis of the coefficient RMSE, illustrated in Fig. 2, bearing in mind that it is influenced by network size and the number of users in the CS considered; thus, the modest values of the RMSE associated with CS5 and CS 6 are also tied to the smaller size of the corresponding networks.

The results shown in Fig. 3 confirm what has been said thus far and enable some additional considerations to be made regarding the accuracy and precision provided by the models in the different case studies. The figure shows the cumulative sampling distributions of the errors e for a fixed time horizon (in this case 1 h). In particular, where the error is defined as the difference between the observed and forecasted demands, a positive error corresponds to an underestimate of the demand forecast by the model, whereas a negative error corresponds to an overestimate.

It may be observed that a lower variability between the minimum and maximum errors, associated with a steep slope of the cumulative distribution curve, indicates a good precision of the forecast, whereas a greater symmetry of the cumulative probability curve relative to the point $e = 0$ (that is, when the curve tends to pass and become symmetrically distributed relative to the point ($e = 0$, $F = 0.5$)) indicates that the model is accurate, that is, it tends neither to overestimate nor to underestimate consumption. In particular, it may be observed, for example from the graphs corresponding to CS4, that with respect to the year y1 all models are characterised by a similar accuracy ($F(0) \approx 0.5$) and precision, with the exception of the naïve model, which tends to underestimate demand and is characterised by a greater scattering of errors. Indeed, the t-test (Benjamin and Cornell 1970) highlights that for the naïve model the hypothesis of mean of the error equal to 0 has to be rejected at the 5% significant level, whereas it is accepted for all the other model. On the other hand, it may be noted from the graph representing y2 that the αβ_WDF and Bakk_WDF models maintain a high degree of accuracy, whereas the Patt_WDF and ANN_WDF models show less accuracy and a tendency to underestimate demand ($F(0) < 0.25$); finally, the HMC_WDF and naïve models greatly underestimate the demand for the year y2, and thus forecast with less precision and accuracy. Indeed, for the year y2 the hypothesis of mean of the error equal to 0 has to be rejected at the 5% significant level for the Patt_WDF, ANN_WDF, HMC_WDF and naïve models, whereas it is accepted only for the αβ_WDF and Bakk_WDF models.

Analogous considerations also apply for the remaining case studies, as all the models show similar performances for the year y1, whereas if the focus is shifted to the year y2, it may be observed that the accuracy and precision of the αβ_WDF and Bakk_WDF models remains substantially unchanged, whereas Patt_WDF, ANN_WDF and HMC_WDF model show a decrease in accuracy. Thus, summing up, models based on the moving-window technique show to deliver a high, more stable accuracy with respect to the 2 years of application whereas the models requiring calibration on the basis of a long series of data undergo a decrease in accuracy from the year of calibration to the year of validation. This decrease is more or less marked depending on the difference between the 2 years in terms of average yearly water demand and results in an under/overestimation of demands in the year y2 depending on whether the average observed demand in the year y2 is higher/lower than the observed demand in the year y1.

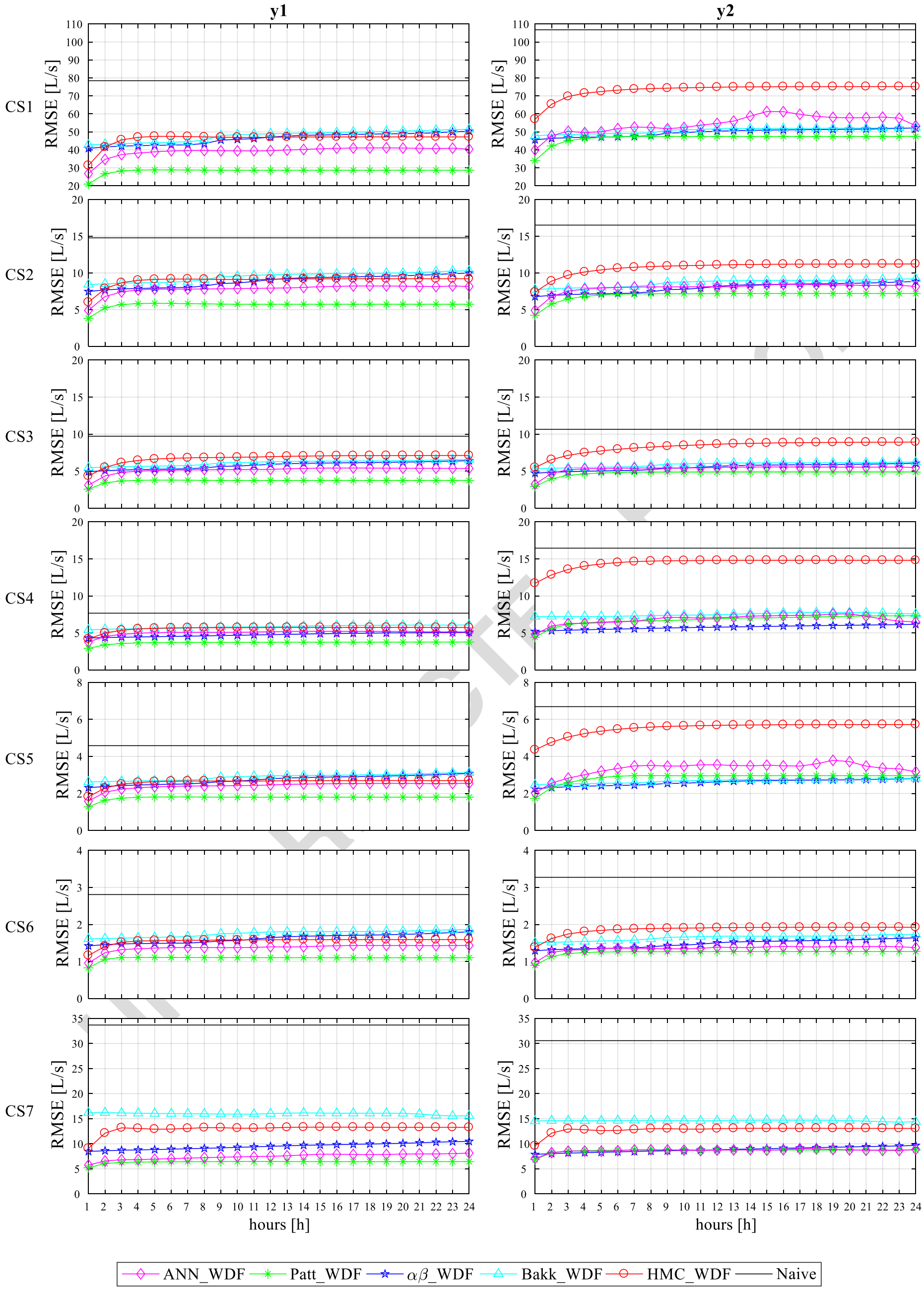

Fig. 2 Values of RMSE for every time horizon (k = 1,2,..24) in the 2 years considered (y1 and y2), for every case study analyzed (CS1,CS2,..CS7)

5 Conclusions 491

This paper presents a comparison between different hourly water demand forecasting models 492
for a 24-h time horizon, already present in the literature providing useful information about the 493

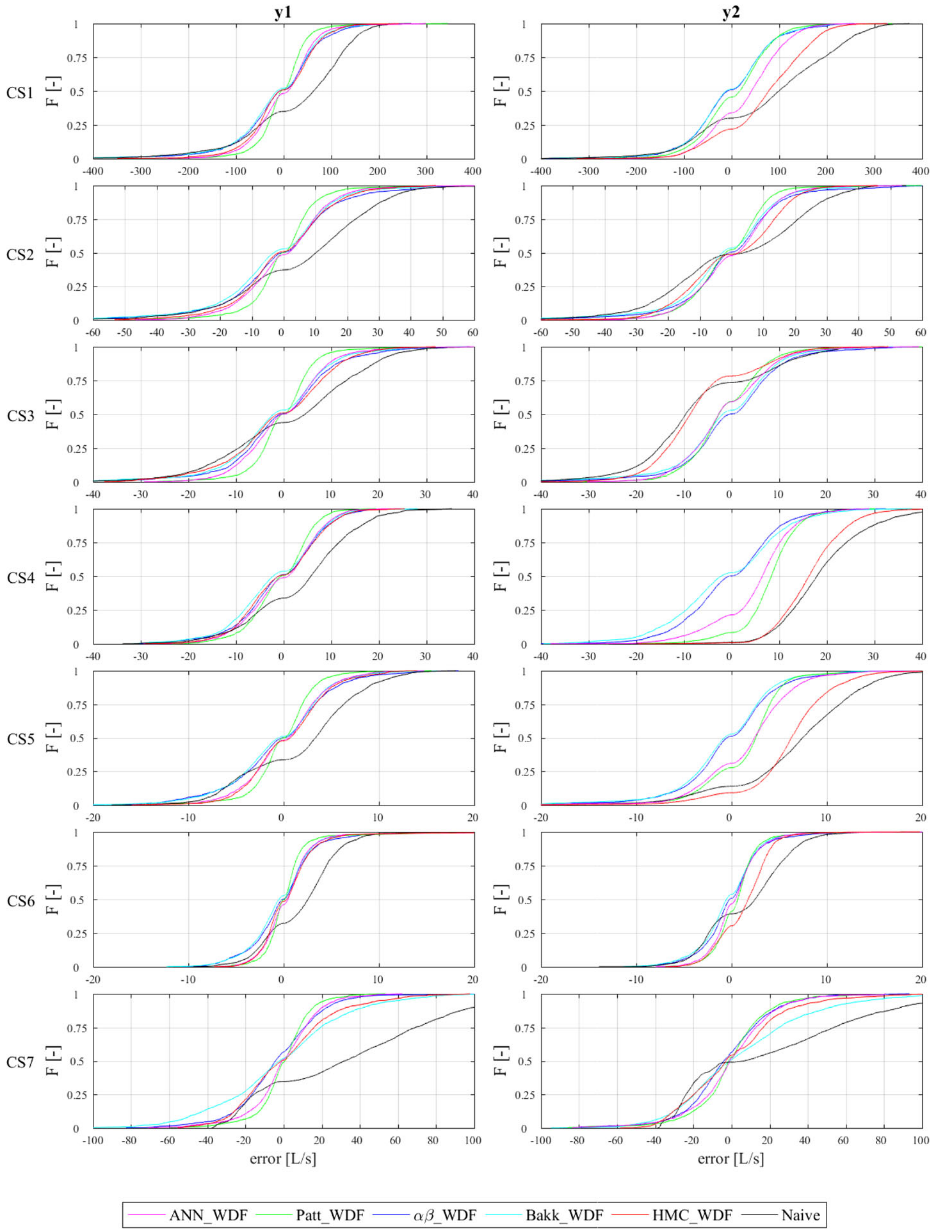

Fig. 3 Cumulative sampling distribution of errors in the 2 years considered (y1 and y2) for each case study analyzed (CS1, CS2,..CS7), with a fixed time horizon of 1 h

pro and cons of the different type and structure of the models. The comparison regarded seven real-life cases of water distribution networks and district-metered areas of different sizes and with a different number and type of users. Data regarding the average hourly water demands in two different years were used.

The models applied differ from one another in terms of their characteristics, including the type of structure, whether they are data-driven or pattern based, use a deterministic or probabilistic approach and require or do not require the use of a long dataset for their calibration.

The analysis of the results has shown that models based on different forecasting techniques deliver high accuracies, and their performances are comparable, when the year of calibration is considered. Indeed the same forecasting accuracy can be achieved using both data-driven and pattern-based techniques.

A more marked difference may be noted between the models requiring calibration on the basis of a long series of data and those based on the moving-window technique. Indeed, it may be observed that, in every case study analysed, the former undergo a decrease in accuracy from the year of calibration to the year of validation. In contrast, models based on the moving-window technique show to deliver a high, more stable accuracy irrespective to the year considered by virtue of their structure, which provides for parameters to be set in a very short moving window. The variability of water demands during the year also impacts all of the other models, though to a lesser extent. In fact, the case study regarding a distribution network characterised by high variability in the number of users over the course of the year showed a general decrease in forecasting reliability, though this was attenuated in the case of models based on the moving-window technique, since their parameters are continuously updated and they can thus better capture variations in demand, albeit with a slight time lag.

References

Adamowski J, Fung Chan H, Prasher SO, 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 Resour Res 48(1):W01528

Alvisi S, Franchini M (2017) Assessment of predictive uncertainty within the framework of water demand forecasting using the model conditional processor (MCP). Urban Water J 14(1):1–10

Alvisi S, Franchini M, Marinelli A (2007) A short-term, pattern-based model for water-demand forecasting. J Hydroinf 9(1):39–50

Anele AO, Hamam Y, Abu-Mahfouz AM, Todini E (2017) Overview, comparative assessment and recommendations of forecasting models for short-term water demand prediction. Water 9(11):887

Anele AO, Todini E, Hamam Y, Abu-Mahfouz AM (2018) Predictive uncertainty estimation in water demand forecasting using the model conditional processor. Water 10(4):475

Arandia E, Ba A, Eck B, McKenna S (2016) Tailoring seasonal time series models to forecast short-term water demand. J Water Resour Plan Manag 142(3):1–10

Babel MS, Shinde VR (2011) Identifying prominent explanatory variables for water demand prediction using artificial neural networks: a case study of Bangkok. Water Resour Manag 25(6):1653–1676

Bakker M, Vreeburg JHG, van Schagen KM, Rietveld LC (2013) A fully adaptive forecasting model for short-term drinking water demand. Environ Model Softw 48:141–151

Benjamin JR, Cornell CA (1970) Probability, statistics and decision for civil engineers. McGraw-Hill, New York

Caiado J (2010) Performance of combined double seasonal univariate time series models for forecasting water demand. J Hydrol Eng 15(3):215–222

Chen G, Long T, Xiong J, Bai Y (2017) Multiple random forests modelling for urban water consumption forecasting. Water Resour Manag 31:4715. https://doi.org/10.1007/s11269-017-1774-7

Cutore P, Campisano A, Kapelan Z, Modica C, Savic D (2008) Probabilistic prediction of urban water consumption using the SCEM-UA algorithm. Urban Water J 5:125–132

Donkor EA, Mazzucchi TA, Soyer R, Roberson AJ (2014) Urban water demand forecasting: a review of methods and models. J Water Resour Plan Manag 140(2):146–159

Gagliardi F, Alvisi S, Kapelan Z, Franchini MA (2017) Probabilistic short-term water demand forecasting model based on the Markov chain. Water 9(11):507

Gato S, Jayasuriya N, Roberts P (2007) Forecasting residential water demand: case study. J Water Resour Plan Manag 133(3):218–229

Gelažanskas L, Gamage K (2015) Forecasting hot water consumption in residential houses. Energies 8(11): 12702–12717

Ghalehkhondabi I, Ardjmand E, Young WA, Weckman GR (2017) Water demand forecasting: review of soft computing methods. Environ Monit Assess 189:313. https://doi.org/10.1007/s10661-017-6030-3

Hagan MT, Menhaj M (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5(6):989–993. https://doi.org/10.1109/72.329697

Herrera M, Torgo L, Izquierdo J, Pérez-Garcìa R (2010) Predictive models for forecasting hourly urban water demand. J Hydrol 387(1–2):141–150

Hsu KL, Gupta HV, Sorooshian S (1995) Artificial neural network modelling of the rainfall-runoff process. Water Resour Res 31(10):2517–2530. https://doi.org/10.1029/95WR01955

Hutton CJ, Kapelan Z (2015) A probabilistic methodology for quantifying, diagnosing and reducing model structural and predictive errors in short term water demand forecasting. Environ Model Softw 66:87–97

Jain A, Varshney AK, Joshi UC (2001) Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. Water Resour Manag 15(5):299–321

Msiza IS, Nelwamondo FV, Marwala T (2008) Water demand prediction using artificial neural networks and support vector regression. J Comput 3(11):1–8

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

Romano M, Kapelan Z (2014) Adaptive water demand forecasting for near real-time management of smart water distribution systems. Environ Model Softw 60:265–276

Shabani S, Candelieri A, Archetti F, Nase G (2018) Gene expression programming coupled with unsupervised learning: a two-stage learning process in multi-scale, short-term water demand forecasts. Water 10:142.2

Todini E (2008) A model conditional processor to assess predictive uncertainty in flood forecasting. Int J River Basin Manag 6(2):123–137

Vrugt JA, Gupta HV, Bouten W, Sorooshian S (2003) A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 39(8):1201

Zhou SL, McMahon TA, Walton A, Lewis J (2000) Forecasting daily urban water demand: a case study of Melbourne. J Hydrol 236(3–4):153–164

Zhou SL, McMahon TA, Walton A, Lewis J (2002) Forecasting operational demand for an urban water supply zone. J Hydrol 259(1–4):189–202

Zubaidi SL, Gharghan SK, Dooley J, Alkhaddar RM, Abdellatif M (2018) Short-term urban water demand prediction considering weather factors. Water Resour Manag 32:4527