Article

A short-term water demand forecasting model based on a moving window of previously observed data

Elena Pacchin [1], Stefano Alvisi [1,*] and Marco Franchini [1]

Received: date; Accepted: date; Published: date
Academic Editor: name

[1] Department of Engineering - University of Ferrara
* Correspondence: stefano.alvisi@unife.it; Tel.: +39-0532-974849

Abstract: A model for short term water demand forecasting is proposed. The model is structured in order to provide at each hour the water demand forecast for the next 24 hours using coefficients estimated according to a short moving window of previously observed data. More in details, the hourly forecast is performed in two steps: in the first step the average water demand for the next 24 hours (Q24) is forecasted multiplying the average water consumption observed in the last 24 hours by a previously estimated coefficient; in the second step, the water consumption of each of the next 24 hours is forecasted multiplying the forecasted Q24 by hourly coefficients. The coefficients' values (both the one used to forecast the Q24 and those used to forecast the hourly values) are updated at each hour on the basis of the water demands observed in the last n (e.g. n=4) weeks. The model is applied to a real case study; the analysis of the results, and their comparison with those provided by another short term water demand forecasting model already presented in the scientific literature, highlights that the proposed model provides an accurate and robust forecast, resulting in an efficient tool for real time management of water distribution networks requiring a very small effort for its parameterization.

Keywords: water demand, forecast

PACS: J0101

1. Introduction

Water demands are the forcing driver of the Water Distribution Networks (WDNs) and their reliable forecast may represent a useful decision support tool for WDN management. On the basis of the forecasting time horizon considered, water demand forecast can be classified as a) long term forecast, characterized by a forecasting time horizon of a decade or more, b) medium term forecast, characterized by a forecasting time horizon of several months up to one year and finally c) short term water demand forecast, characterized by a forecasting time horizon of several days up to one week. Focusing the attention on the last category, i.e. short term forecast, several models have been recently proposed in the scientific literature (see, for example, [1], [2] and [3], in order to get a review of these models). The short term water demand forecasting models differ in many aspects, mainly in terms of forecasting time-step, input data, and approach used to perform the forecast.

The forecasting time-step varies from one hour (e.g. [4]) up to one day (e.g. [5]). In the scientific literature there are also models with multiple forecasting time-step, in which for example both the daily and hourly water demand are forecasted, as in [6]. As regards the input data, the majority of short term water demand forecasting models is mainly based on the time series of observed data related to the previous weeks/months/years. However, in some cases, climate factors such as humidity and temperature, are also used as input (e.g. [7]). Finally, as far as the approach concerns, many models recently proposed in the scientific literature are based on data-driven techniques such as Artificial Neural Network (e.g. [8]), Fuzzy Logic (e.g. [9]), Project Pursuit Regression, Random

Forests and Multivariate Adaptive Regression Splines (e.g. [4]). Other models are on the contrary based on the representation of periodic behaviors that typically affect water demand, coupled with an analysis of time series, and thus can be in general indicated as pattern based models (e.g. [6]).

It is to be pointed out that the model's parameterization is strictly connected with the structure and the approach of the model itself. The models mentioned above are based on structures that can be more or less simple or complex, but whose parameters have always to be defined before the model application. For example, ANNs models, by their very nature, need a calibration of weights and biases; this calibration can be done using time series of observed data that have to be sufficiently long in order to efficacely train the neural network, taking into account that the water demands can vary significantly through the year. The same consideration applies for the pattern based models. For example the model proposed by [6] needs at least one year of observed data in order to define periodic behaviors, namely seasonal fluctuations, as well as for the characterization of the daily consumptions patterns which are variable through the seasons.

The need of having at least one long time-series of observed data of the water demand of the network to which the forecasting model has to be applied could represent a limiting factor to be taken into account in the choice of the forecasting model itself.

In this article a short term water demand forecasting model, which uses a short moving window of previously observed data as input is presented. In fact, the model is structured in order to use, as input data, the hourly data of consumption observed during some weeks before the forecasting time; as output, it provides the hourly water demands forecast for the next 24 hours. As shown in the next sections, the strengths of the proposed model are the simple structure and the substantial absence of a calibration period. In the next sections we present the structure of the proposed forecasting model (section 2) and its application to a real case study (section 3). The results obtained are analyzed and compared with those provided by another short term water demand forecasting model already proposed in the scientific literature. Conclusive considerations are finally provided in section 4.

2. The Proposed Model

The proposed short term water demand forecasting model is based on the observation that water consumptions are typically characterized by periodic behaviors at different time scales. In fact daily average water demands have typical seasonal fluctuations, with increasing consumptions in the warmest months and decreasing consumptions in the coolest ones. Hourly water demands, during the day, have typical fluctuations that denote the users' daily routine. In particular, as regards to residential users, water demands have a daily patter characterized by peaks in the early morning and in the evening, and in general they shows low values during the night and a fluctuating trend during the rest of the day. The patterns of hourly water demands observed during the day may be distinguished on the base of the considered day of the week (the day "type"). For example, on Mondays (or Tuesdays, Wednesdays,..) the consumptions of residential areas or districts show similar patterns; indeed these patterns are quite similar to those of the other weekdays but they differ from those of the weekend.

Another important aspect in the characterization of short term water consumptions is connected to the persistence of consumptions themselves. In fact, analysis of water consumptions time-series having hourly or daily time step, typically shows that the values of the correlation coefficient at lag 1 or 2 are significant [6]; this indicates that increasing trend of the consumption of the current day typically is associated with rather high average daily consumption for the next day.

The proposed model is based on these two aspects. In fact, in order to forecast the hourly water demands for the next K=24 hours, the model uses the information from consumer data observed in the last 24 hours prior to the forecasting time instant. Furthermore the model uses the information related to the same hours of the day of the forecasting time window but observed on the same day "type" in the n weeks earlier. In particular the model is structured in two modules, so that in the first one the average water demand for the next 24 hours is evaluated, and then in the second module, using the prior evaluation, the average hourly water demand is forecasted for each of the 24 hours ahead.

More in details, let us assume the week as made up by 7 different day "type" each one characterized by a particular hourly consumption pattern. With reference to Figure 1, let assume an hourly time-step and let i (with $i = 1,2,\ldots,365$) be the generic day of the year and j the generic hour of the day (with $j = 1,2,\ldots,24$). Furthermore, let us indicate with t the generic hour of the year (starting from the 1st January) in which the forecast is made. Assuming that the forecast is made at the hour j of the day i, t (the forecasting time) is defined as $t = 24\cdot(i-1)+j$. Finally k indicates the forecasting lead-time (with $k = 1,2,\ldots,K=24$) and is counted starting from t.

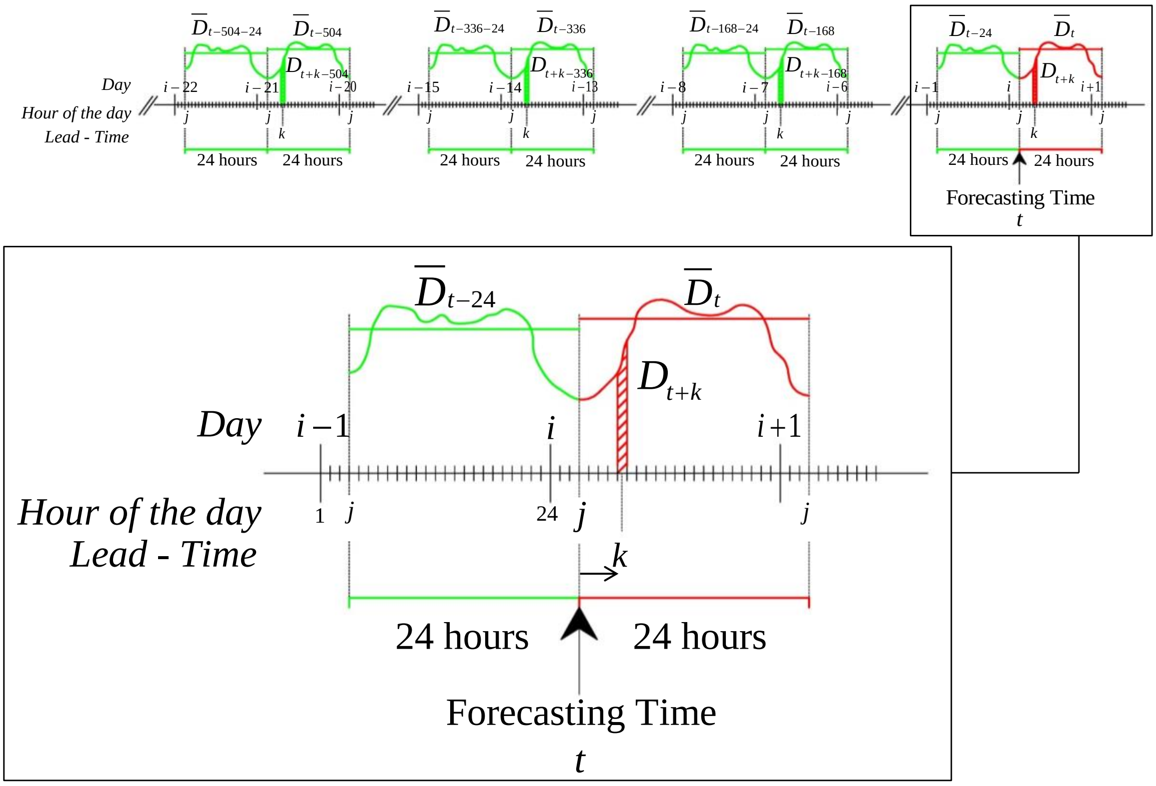

Figure 1. Time structure of the proposed model αβ_WDF.

As noted above, consumptions are characterized by different trends in the 7 days of the week, and thus it can be assumed that the days $\{i, i-7, i-14, i-21,..\}$ belong to the same "type" (for example, all of them are Tuesdays). Therefore, being $t = 24\cdot(i-1)+j$ the forecasting time, it is possible to identify the n hours corresponding to the same hour j of the day for which the forecast is performed but which belongs to the same "type"of the days in the n weeks before the forecasting time t.

The water demand forecast is performed in the following way. At hour t the average water demand $\bar{D}_t$ for the next K=24 hours is evaluated as:

$$\bar{D}_t = \alpha_t \cdot \bar{D}_{t-24} \qquad (1)$$

where $\bar{D}_{t-24}$ is the average consumption observed in the 24 hours prior to the forecasting time t (that is from t-24 up to t) and α_t is a coefficient which has a specific value for the forecast horizon of 24 hours which belong at hour t. This coefficient is fixed as mean value of the ratio of the average water consumptions observed in the 24 hours after each of the n hours corresponding to the forecasting time but belonging to the previous n weeks, and the average water consumptions observed in the 24 hours before each of these n hours (i.e. $\bar{D}_{t-168}$ and $\bar{D}_{t-168-24}$, $\bar{D}_{t-336}$ and $\bar{D}_{t-336-24}$ etc. shown in Figure 1).

Once the average water demand $\bar{D}_t$ of the K=24 hours following the hour t is estimated, the hourly water demand D_{t+k} corresponding to hour t+k is forecast as:

$$D_{t+k} = \beta_{t,k} \cdot \bar{D}_t \tag{2}$$

where $\beta_{t,k}$ is a coefficient related to the specific lead-time k in the horizon of K=24 hours that starts at hour t. As α_t, $\beta_{t,k}$ is evaluated on the basis of the moving window's data. More in details, the n hours having the same characteristics (same hour j of the day and same day "type") of the hour t+k for which the model is forecasting the water demand but which belong to the previous n weeks are identified (i.e. $D_{t+k-168}$, $D_{t+k-336}$,… etc. shown in Figure 1), and $\beta_{t,k}$ is evaluated as the mean value of the ratios between the water consumptions observed in these hours and the average water consumptions observed in the 24 hours including these hours previously identified (i.e. $\bar{D}_{t-168}$, $\bar{D}_{t-336}$ etc. shown in Figure 1).

It must be pointed out that, for every forecasting time-step t, the model calculates 24 values of the coefficient $\beta_{t,k}$, one for each forecasting lead-time k (k = 1, 2,...., K=24).

Finally, it is worth noting that the coefficients α_t and $\beta_{t,k}$ contain the information characterizing the water demand of a certain day "type" and of a certain hour. These coefficients are continuously updated as they are evaluated on a moving window that ends at the forecasting time-step t and moves forward with it. Clearly the moving window has to be wide enough to allow a steady evaluation of the coefficients but in the meantime its length has to be limited up to a maximum of one or two months in order to take into account the consumption's fluctuations connected to the seasonal variation. In fact, using a wider window (e.g. 6 months) if the forecasting time-step occurs in a summer day, the corresponding coefficients would be evaluated on the basis of winter data of consumption too; in this way the seasonal fluctuations could not be properly taken into account. Another advantage given by the use of a narrow moving window is that the model can be applied after only the n weeks required in order to gather the consumption's data. In fact, this model does not require a long calibration data set, as many other models previously mentioned, and operatively it can be used immediately after collecting few weeks of observed water consumptions.

3. Case Study

The proposed model, hereafter named αβ Water Demand Forecasting model (αβ_WDF) is applied to the real case study of Castelfranco Emilia (Italy). The Water Distribution Network of Castelfranco Emilia serves about 23000 inhabitants and it is entirely fed by a tank equipped with a flow measurement device, which provides the hourly data of the total consumption of the town, inclusive of the leakages. More in details, the hourly consumption data of the years 1998 and 2000 were available. The same case study was used in [6] in order to apply and test the model the short term pattern based water demand forecasting model, named Patt_WDF. The results of Patt_WDF are used in this paper as a basis for comparison and to evaluate the accuracy of the proposed model. More in details, the evaluation of the models' performances was done considering separately the different forecasting time horizon and using the MAE% index defined as:

$$MAE\% = \frac{1}{z}\sum_{i=1}^{z}\left|\frac{e_i}{\mu_{obs}}\right| \cdot 100 \tag{3}$$

where z is the number of hours in an year, $e_i = Q^{obs} - Q^{for}$ is the error, being Q^{obs} the value of the observed hourly water demand and Q^{for} the value of the forecasted water demand, and μ_{obs} the average value of the observed consumptions.

It may be reminded that, as observed in the introduction, Patt_WDF is a short term water demand forecasting model based on an hybrid technique and, as αβ_WDF, it is characterized by a 24 hours forecasting horizon and a hourly time-step. However the Patt_WDF model, unlike the αβ_WDF model, requires at least one year of hourly observed water consumption's data in order to perform the calibration phase. In particular, as we have the hourly observed consumption's data of the years 1998 and 2000, the data of 1998 were used to calibrate the Patt_WDF model and those of 2000 are used to validate it. For this reason, even though the αβ_WDF requires only a minimum calibration to be activated, the results are shown considering separately the two years in order to allow a fair comparison with the Patt_WDF's results.

4. Analysis of the Results and Conclusions

Figure 2 shows the values of the MAE% index obtained considering the water demand forecast for 1, 2,…., 24 hours ahead for both the calibration and validation phases and for both the models. Analyzing the results obtained with the αβ_WDF model, it can be observed that the forecasting accuracy is, in general, very good. More in details, the accuracy is greater for short time horizons and it slightly decreases when the forecasting horizon increases, showing limited variations of the value of MAE%. In fact, with reference to year 1998, the value of the MAE% increases from 6% for the 1 hour ahead forecast, up to 7% for the 24 hours ahead forecast. Similarly, with reference to year 2000, the value of the MAE% increases from 5.5% up to 6.5% (respectively for forecasts made for 1 and 24 hours ahead).

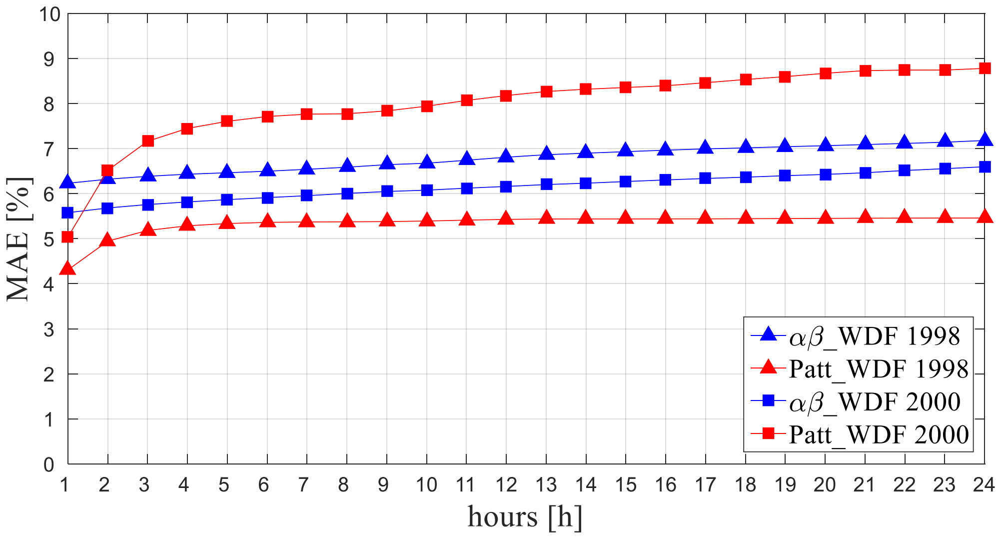

Figure 2. Values of the MAE% obtained for both the calibration (1998) and validation (2000) phases.

Comparing the results of the αβ_WDF model with those of the Patt_WDF model, it can be noticed that in 1998 the Patt_WDF proves to be slightly more accurate than the αβ_WDF model for all the forecasting time horizons: the MAE% varies from 4.5% for the forecasts made for 1 hour ahead up to 5.5% for the 24 hours ahead. On the other hand, it can be observed that in year 2000 the αβ_WDF model is more accurate than the Patt_WDF model: αβ_WDF's MAE% values are similar to those obtained by Patt_WDF for forecasts with horizon equal to 1 hour (5% for Patt_WDF and 5.5% for αβ_WDF) but for all the other forecasting time horizons the αβ_WDF model is much more accurate. In fact, considering the 24 hours horizon, the MAE% of the Patt_WDF model value is equal to 9% whereas the MAE% of the αβ_WDF model is equal to 6.5%.

In general, the comparison of the statistical indexes obtained by the two models, for year 1998 and 2000, shows that the performances of the αβ_WDF model are steady through both the years. On contrary the performance of the Patt_WDF model worsens from 1998 to 2000. This is understandable, recalling that the Patt_WDF model requires a long calibration dataset (1 year) and in fact, the Patt_WDF model show a high forecasting accuracy when applied to the year of calibration (1998), which is used to characterize the periodic behaviors that are used as the forecasting basis of the model. On the other hand, considering a different year, and a rather different water consumption time series, its performances tend to worsen. On the contrary, the performances of the αβ_WDF model, which requires a very short time period for setting up the initial parameterization and is based on coefficients evaluated on a moving window of observed consumptions, are steady even considering several different years.

In brief, the proposed αβ_WDF model has a good forecasting accuracy through the entire considered forecasting time horizon (24 hours). Furthermore it does not require any calibration period, as it is based on the water demands observed in the few weeks prior to the forecasting time. This peculiarity makes it usable even in those cases in which there are not long time series of observed values; in fact it can be used right after one or two months of collected data. Finally, still due to the absence of calibration, the forecasting accuracy remains constant and high even if the model is applied to several years, making it a robust and effective instrument for managing water distribution networks.

Acknowledgments: This study was carried out as part of the PRIN 2012 project "Tools and procedures for an advanced and sustainable management of water distribution systems", n. 20127PKJ4X, funded by MIUR, and under the framework of Terra&Acqua Tech Laboratory.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:
WDN: Water Distribution Network
MAE%: Mean Absolute Percentage Error

References

1. Billings, B., and Jones, C. (2008). Forecasting Urban Water Demand (2nd ed.). AWWA.
2. Adamowski, J., Chan, H.F., Prasher, S.O., Ozga-Zielinski, B. & Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resour Res* 48(1): W01528.
3. Donkor, E., Mazzuchi, T., Soyer, R., and Alan Roberson, J. (2014). Urban Water Demand Forecasting: Review of Methods and Models. *J Water Resour Plan Manag*, 140(2), 146-159.
4. Herrera, M., Torgo, L., Izquierdo, J., and Perez-Garcia, R. (2010). Predictive models for forecasting hourly urban water demand. J Hydrol, 387(1-2), 141–150. doi:10.1016/j.jhydrol.2010.04.005
5. Gato, S., Jayasuriya, N., and Roberts, P. (2007). Forecasting residential water demand: Case study. *J Water Resour Plan Manag -ASCE*, 133(4), 309–319.
6. Alvisi, S., Franchini, M., and Marinelli, A. (2007). A short-term, pattern-based model for water-demand forecasting, J Hydroinf, 9 (1), 39-50. *J Hydroinf* 9: 39–50.
7. Dos Santos, C. C. and Pereira Filho, A. J. (2014). Water Demand Forecasting Model for the Metropolitan Area of Sao Paulo, Brazil, *Water Resour Manage*, 28, 4401–4414.
8. Romano, M. and Kapelan, Z. (2014). Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environ Modell Software* 60, 265–276.
9. Lertpalangsunti, N., Chan, C.W., Mason, R. & Tontiwachwuthikul, P. (1999). Toolset for construction of hybrid intelligent forecasting systems: Application for water demand prediction. *Artif Intell Eng*, 13(1), 21–42.