# A short-term, pattern-based model for water-demand forecasting

Stefano Alvisi, Marco Franchini and Alberto Marinelli

## ABSTRACT

The short-term, demand-forecasting model described in this paper forms the third constituent part of the POWADIMA research project which, taken together, address the issue of real-time, near-optimal control of water-distribution networks. Since the intention is to treat water distribution as a feed-forward control system, operational decisions have to be based on the expected future demands for water, rather than just the present known requirements. Accordingly, it was necessary to develop a short-term, demand-forecasting procedure. To that end, monitoring facilities were installed to measure short-term fluctuations in demands for a small experimental network, which enabled a thorough investigation of trends and periodicities that can usually be found in this type of time-series. On the basis of these data, a short-term, demand-forecasting model was formulated. The model reproduces the periodic patterns observed at annual, weekly and daily levels prior to fine-tuning the estimated values of future demands through the inclusion of persistence effects. Having validated the model, the demand forecasts were subjected to an analysis of the sensitivity to possible errors in the various components of the model. Its application to much larger case studies is described in the following two papers.

**Key words** | demand forecasting, short-term, POWADIMA, water distribution

**Stefano Alvisi** (corresponding author)
**Marco Franchini**
Dipartimento di Ingegneria,
Università degli Studi di Ferrara,
Ferrara 44100,
Italy
Tel.: +39 0532 97 4930
Fax: +39 0532 97 4870
E-mail: *salvisi@ing.unife.it*

**Alberto Marinelli**
DISTART,
Università degli Studi di Bologna,
Bologna 40136,
Italy

## INTRODUCTION

### Need for demand forecasting

In the context of operational control, the demand for water comprises not only that required for customers but also leakage from the distribution network, since it is the combined amount which is put into supply. Therefore, demands are normally estimated by means of district metering. For the efficient operation of water-distribution networks, some form of demand forecasting is usually needed in order to programme the pumping arrangements over the next 24 h, to take advantage of the electricity tariff structure. For example, in the case of pump-scheduling (Sterling & Coulbeck 1975; Zessler & Shamir 1989; Jowitt & Germanopoulos 1992; etc.), a short-term demand forecast is required at the beginning of each morning for the following 24 h. More often than not, these are based on averaged demand profiles for the particular day of the week, which may vary according to the season or month of the year. If for any reason, there is a significant difference between the demand profile assumed and that which materializes as the day progresses, it may be necessary to re-run the pump-scheduling software with the revised data but even then, further assumptions need to be made about the future demands for the remaining portion of the 24 h.

### Real-time control requirements

The alternative to using a pre-defined demand profile is to develop an adaptive demand-forecasting process which can be continually updated to account for any deviations between that forecasted and reality. For real-time,

doi: 10.2166/hydro.2006.016

near-optimal control of water-distribution networks, this alternative approach is the more appropriate since the pump and valve control settings need to be re-optimized at short, regular intervals, in response to the highly variable demands. Accordingly, this particular paper describes the demand-forecasting process used in the POWADIMA (Potable Water Distribution Management) research project (Jamieson *et al.* 2007). As such, it forms the third constituent part of the DRAGA-ANN (Dynamic, Real-time, Adaptive, Genetic Algorithm – Artificial Neural Network) control system, the other two being the means of predicting the consequences of different control settings (Rao & Alvarruiz 2007) and the search engine to find the best combination of control settings (Rao & Salomons 2007). The aim was to develop a model for real-time demand forecasting, consistent with the needs of the DRAGA-ANN control process. As such, the spatial variation in demands has been incorporated by having a separate forecast for each of the district metering areas (DMA) that comprise the distribution network.

### Previous examples of demand-forecasting models

Any cursory examination of the scientific literature will confirm that a considerable amount of effort has been expended on water-demand forecasting. In the case of operational control, the interest is restricted to short to medium timescales (hourly, daily and monthly), rather than annual which largely relates to longer-term water-resources planning. Even then, there are still a large number of papers detailing various methodologies for hourly forecasts (for example, Shvartser *et al.* 1993; Zhou *et al.* 2002; etc.) as well as daily/monthly timescales (Maidment & Parzen 1984a,b; Maidment *et al.* 1985; Franklin & Maidment 1986; Smith 1988; Miaou 1990). Without exception, all these papers refer to the recurring patterns and periodicities that exist in water-demand data, at different levels of temporal aggregation. For this reason, work began with an extensive monitoring campaign for a small, urban water-distribution network in order to understand and, later, predict the observed variability in demands, before progressing to the much larger networks of Haifa-A (see Salomons *et al.* 2007) and Valencia (Martinez *et al.* 2007).

## MONITORING CAMPAIGN AND DATA ANALYSES

### Choice of case study

Prior to the outset of the POWADIMA research project, the University of Ferrara was already involved with the Municipality of Castelfranco Emilia in measuring short-term fluctuations of water demands. Therefore, rather than installing a completely new monitoring scheme at a different location, it seemed logical to make use of the existing facilities by extending them to meet the revised requirements. Castelfranco Emilia is situated in Emilia Romagna, Italy, between Modena and Bologna. Its water-distribution network is greatly influenced by the town's development pattern, which extends radially from the centre towards the more sparsely inhabited rural areas. As a result, there is a dense, grid-like network of pipes in the town centre, which is replaced by a more open, ribbon-like structure in the outlying areas (Figure 1). The overall length of pipes forming the network amounts to some 160 km, the network itself being supplied by a local well field via a storage tank near the town centre.

### Monitoring scheme

The distribution network is monitored for both flow and pressure at 14 stations, strategically located so that data could be gathered for different types of customers (residential, commercial, etc.) at different levels of spatial as well as temporal aggregation, thereby enabling an in-depth analysis of how space–time correlation characteristics vary according to the level of aggregation (Alvisi *et al.* 2003). Since the overall aim is to develop a water-demand forecasting model suitable for large, complex distribution networks, as in the case of Haifa-A and Valencia, special attention was focused on data acquisition for Station 9, which is located near to the storage tank serving the whole network. This measured the supply to the entire population of some 23,000 inhabitants, which roughly equates to the size of a DMA found in major cities.

### Water-demand patterns

As expected, an analysis of the observed hourly and daily water-demand time-series revealed the existence of patterns

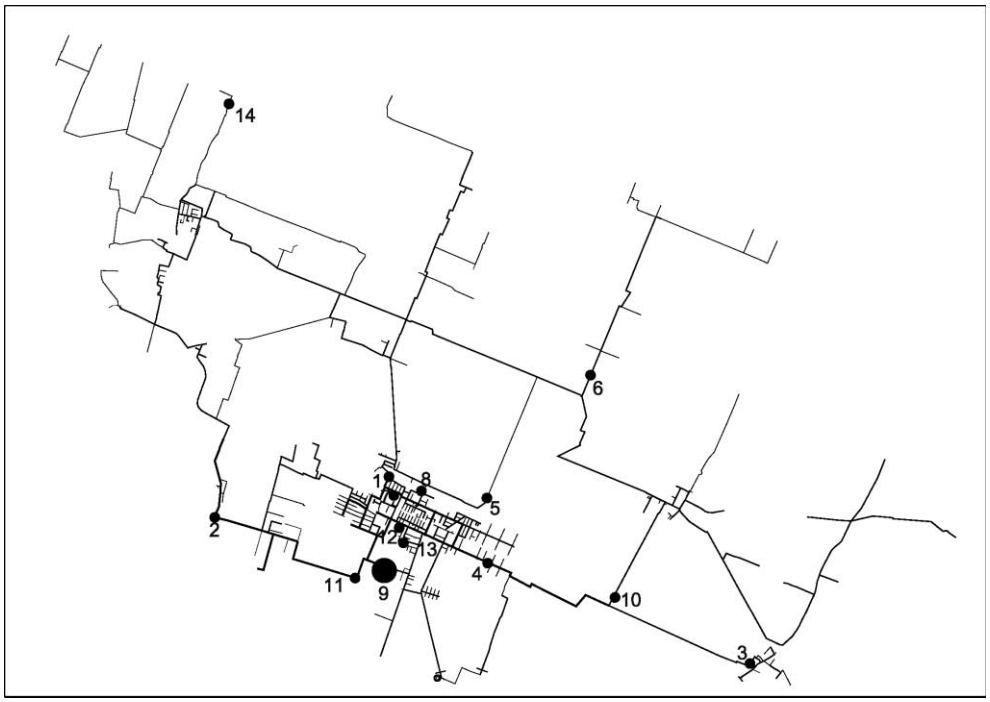


**Figure 1** | The main water distribution network of Castelfranco Emilia (Province of Modena, Italy) and location of the 14 monitoring stations installed. Station 9 was used in this study.

in which it is possible to identify seasonal and weekly periodicities in daily water demands as well as daily periodicities in hourly water demands. Figure 2 shows the pattern in daily water demands for a whole year, with demand rising over the summer period and during the week from Sunday to Saturday. Likewise, the hourly water demands show a variable diurnal behaviour over the day, with different patterns depending whether it is a weekday or weekend and, to a lesser extent, the season (Figure 3). Analysis of the observed daily and hourly time-series data indicated that the demand pattern during public holidays was very similar to that observed during weekends. Therefore, public holidays have been treated as weekends for the purposes of demand forecasting.

## PATTERN-BASED WATER-DEMAND FORECASTING MODEL

### Structure of model

The forecasting model proposed is based on the patterns implicit in the water-demand time-series data and for this reason, it is referred to as the Pattern-based Water Demand Forecasting (Patt_WDF) model. In the first of its two modules, whose purpose is to estimate the average daily water demand on the day (or days) covered by the rolling 24 h forecasting window, the seasonal and weekly patterns are taken into account, together with the medium-term persistence. In the second module, these daily amounts are then combined with the daily demand patterns and short-term persistence to provide hourly forecasts over the following 24 h period (Figure 4).

### The daily water-demand module (DM)

In this approach, the average daily water demand $Q_m^{d,for}$ for the Julian date $m$ is estimated by means of the following relationship:

$$Q_m^{d,for} = \bar{Q}_m^{d,s} + \bar{\Delta}_{i,j}^{d,w} + \delta_m^d \quad (1)$$

where $\bar{Q}_m^{d,s}$ is the long-term average daily water demand representing the seasonal periodic component, $\bar{\Delta}_{i,j}^{d,w}$ is a correction representing the weekly periodic component and

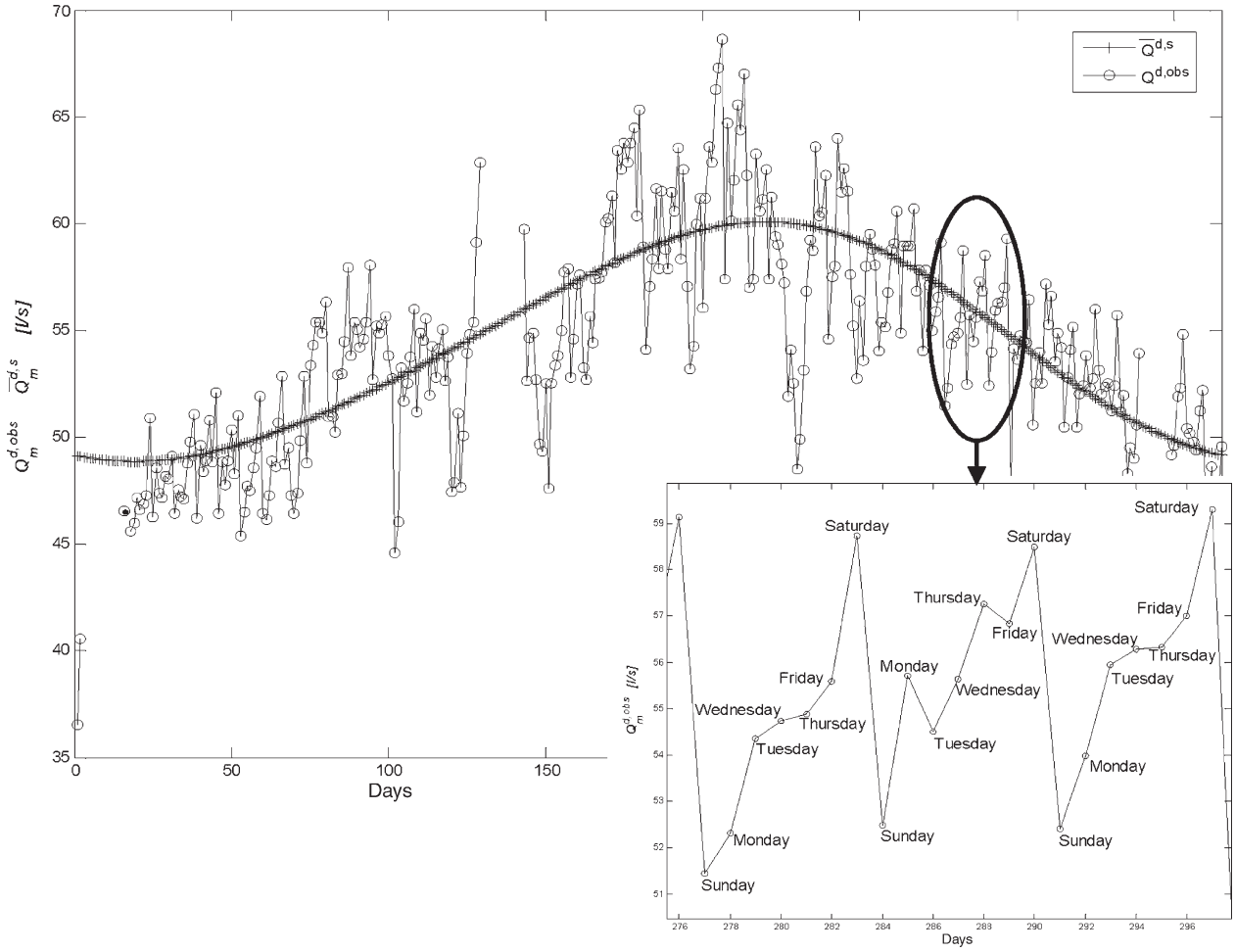


**Figure 2** | Average daily water demand $Q^{d,obs}$ measured at the outlet of the storage tank in Castelfranco Emilia during the year 1998. Seasonal and weekly periodicities are shown.

$\delta_m^d$ is a deviation representing the medium-term persistence component.

The long-term average daily water demand $\bar{Q}_m^{d,s}$ (Figure 2) is modelled using a Fourier series:

$$\bar{Q}_m^{d,s} = a_0 + \sum_{k=1}^{K}\left[a_k \cos\frac{2\pi k}{365}m + b_k \sin\frac{2\pi k}{365}m\right],\ m = 1,2,\ldots,365 \quad (2)$$

where $a_0$ is the mean value of the seasonal cycle, $a_k$ and $b_k$ are Fourier coefficients and $K$ is the number of harmonics considered. The Fourier coefficients $a_k$ and $b_k$ are calculated as:

$$a_k = \frac{2}{365}\sum_{m=1}^{365} Q_m^{d,obs} \cos\frac{2\pi m}{365}k, \quad (3)$$

$$b_k = \frac{2}{365}\sum_{m=1}^{365} Q_m^{d,obs} \sin\frac{2\pi m}{365}k \quad (4)$$

where $Q_m^{d,obs}$ are the observed average daily water demands.

The weekly correction factor $\bar{\Delta}_{i,j}^{d,w}$ is defined as:

$$\bar{\Delta}_{i,j}^{d,w} = \bar{Q}_{i,j}^{d} - \bar{Q}_j^{w} \quad (5)$$

where $\bar{Q}_{i,j}^{d}$ is the mean value of the average daily water demand observed on day $i$ of the week ($i = 1,\ldots,7$, Monday, …, Sunday), $j$ the season ($j = 1,\ldots,4$, winter, spring, summer, autumn) and $\bar{Q}_j^{w}$ is the mean value of the average

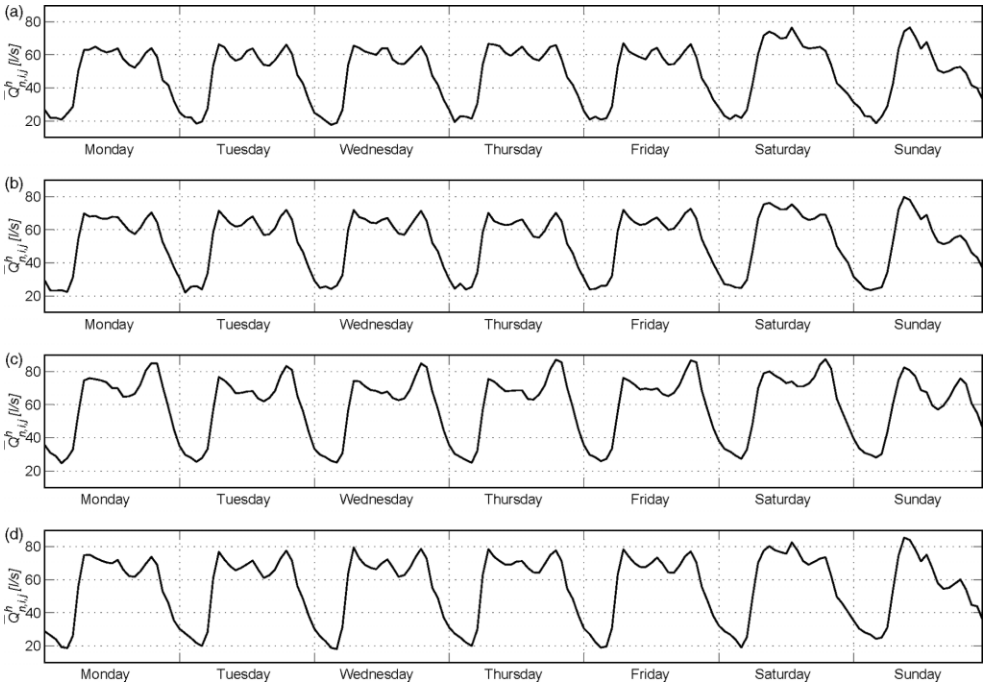


**Figure 3** | Daily patterns in hourly water demand for the seven days of the week: (a) in winter, (b) in spring, (c) in summer and (d) in autumn.

weekly water demand in season $j$. Finally, $\delta_m^d$, which represents the deviation between the average daily water demand $Q_m^d$ and the mean value estimated solely on the basis of the periodic components $\bar{Q}_m^{d,s}$ and $\bar{\Delta}_{i,j}^{d,w}$, is modelled using an autoregressive process $AR(1)$ (Box *et al.* 1994). For forecasting purposes, this may be written as:

$$\delta_m^d = \Phi_1 \cdot \delta_{m-1}^d \tag{6}$$

whose parameter $\Phi_1$ is calibrated on the basis of the observed deviations:

$$\delta_m^{d,obs} = Q_m^{d,obs} - \left(\bar{Q}_m^{d,s} + \bar{\Delta}_{i,j}^{d,w}\right), \tag{7}$$

Figure 5 shows the pattern of the daily residuals observed and modelled by means of Equation (7), with reference to a year of measured data at Station 9.

## The hourly demand-forecasting module (HM)

The hourly module, like the daily module, is composed of two parts: a periodic component and a persistence component. Here, the hourly water demand $Q_{t+k}^{h,for}$ forecasted at the $t$th hour for $k$ hours ahead is given by

$$Q_{t+k}^{h,for} = Q_m^{d,for} + \bar{\Delta}_{n,i,j}^{h} + \varepsilon_{t+k} \tag{8}$$

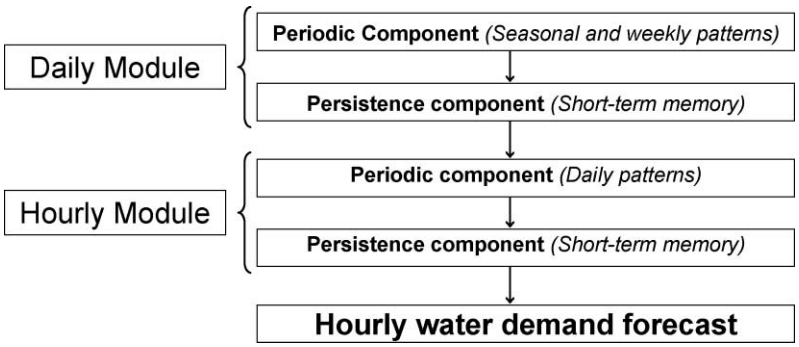


**Figure 4** | Structure of the two modules making up the Patt_WDF model.

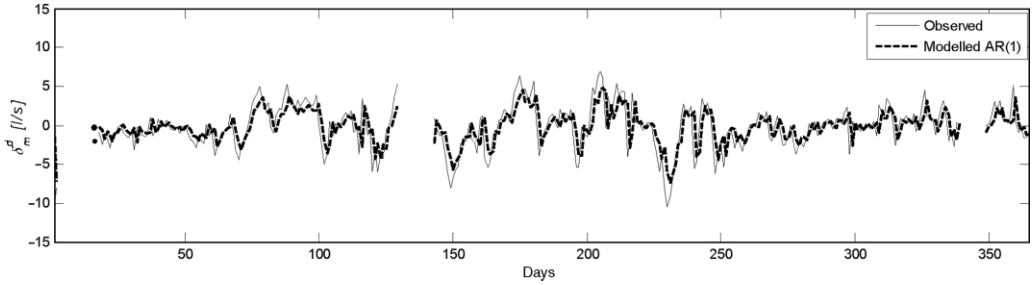


**Figure 5** | Daily residuals $\delta_m^d$ observed (unbroken line) and modelled via the *AR*(1) model (dotted line).

where $Q_m^{d,for}$ is the average daily water demand forecast by means of the *DM* using the information available up to Julian date $m$-1 (see Equations (1) and (6)), which is replaced with $Q_{m+1}^{d,for}$ when $t+k$ leads to the definition of an hour falling in the following day (see Figure 6). $\bar{\Delta}_{n,i,j}^h$ is the hourly deviation representing the daily pattern defined by

$$\bar{\Delta}_{n,i,j}^h = \bar{Q}_{n,i,j}^h - \bar{Q}_{i,j}^d \tag{9}$$

where $\bar{Q}_{n,i,j}^h$ represents the mean value of the average hourly water demands observed in hour $n$ ($n = 1,\ldots,24$, the hour of the day) of day $i$ in season $j$, and $\bar{Q}_{i,j}^d$ is the mean value of the average daily water demands observed on day $i$ in season $j$ (see Figure 7). Finally $\epsilon_{t+k}$ represents the hourly persistence component, which is modelled using a regression on the errors $\epsilon_{t+k\text{-}1}$ and $\epsilon_{t+k\text{-}24}$ as follows:

$$\varepsilon_{t+k} = \psi_1 \cdot \varepsilon_{t+k-1} + \psi_{24} \cdot \varepsilon_{t+k-24}. \tag{10}$$

The coefficients $\psi_1$ and $\psi_{24}$ depend on the hour of the day $t+k \equiv n = 1,2,\ldots,24$ corresponding to the hour $t+k$ considered (it should be noted that the hour $t+k$ is counted starting from the beginning of the year) and are calibrated on the basis of the observed errors $\varepsilon_t^{obs}$ where

$$\varepsilon_t^{obs} = Q_t^{h,obs} - \left(Q_m^{d,obs} + \bar{\Delta}_{n,i,j}^h\right). \tag{11}$$

## ANALYSES OF RESULTS

### Accuracy of forecast

The accuracy of the forecast was initially assessed by separately considering the forecast for 1 h ahead, 2 h ahead and so forth up to 24 h ahead, independently of the time serving as the starting point of the forecast. Then followed an analysis of how the model's performance varied according to the hour in which the forecast was made. These assessments were made in reference to both the year of data on which the calibration was based and the year of data utilised for validation.

Forecasting accuracy was measured by the explained variance (*EV*), the root mean square error (*RMSE*) and the mean absolute percentage error (*MAE%*) as defined by

$$EV = 1 - \frac{\sum_{i=1}^{n}(e_i - \mu_e)^2}{\sum_{i=1}^{n}\left(x_i^{obs} - \mu_{obs}\right)^2} \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n} e_i^2} \tag{13}$$

$$MAE\% = \frac{1}{n}\cdot\sum_{i=1}^{n}\left|\frac{e_i}{\mu_{obs}}\right|\cdot 100 \tag{14}$$

where $n$ is the number of observed data (which in this case is the number of hours in the year), $e = x^{obs} - x^{for}$ the errors, $x^{obs}$ the observed values, $x^{for}$ the forecast values, $\mu_e$ the mean error and $\mu_{obs}$ the mean of the observed values.

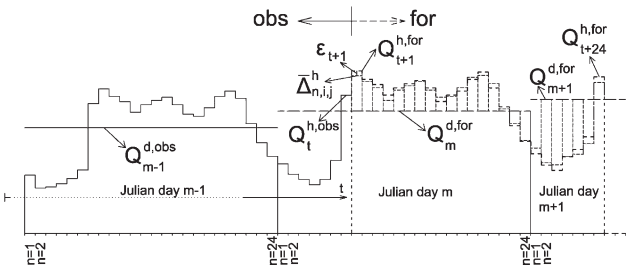


**Figure 6** | Forecast average hourly water demands $Q_{t+k}^{h,for}$ obtained by combining the forecast average daily water demand $Q_m^{d,for}$, the daily pattern $\bar{\Delta}_{n,j,j}^h$ and an estimated error $\epsilon_{t+k}$.

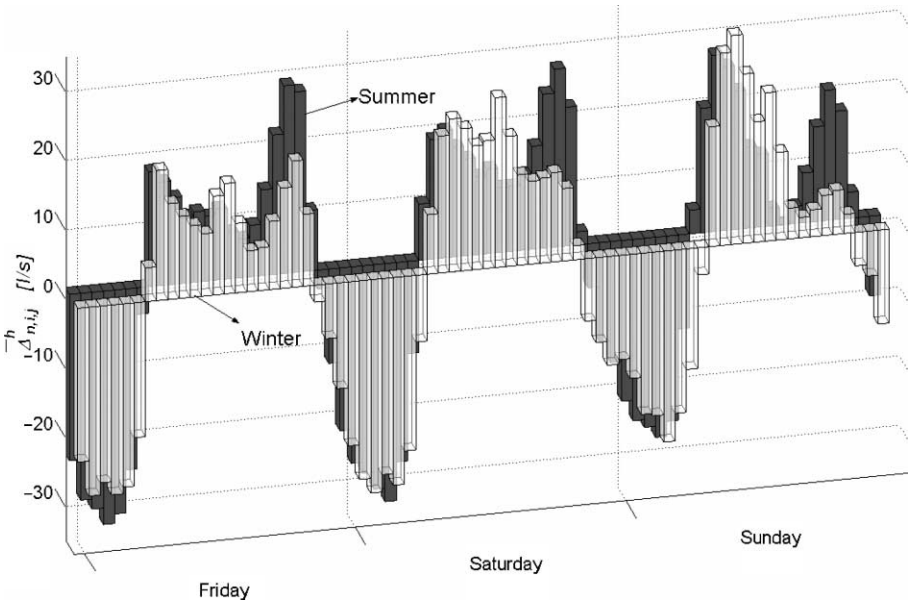


**Figure 7** | The daily patterns $\bar{\Delta}^{h}_{n,ij}$ associated with a weekday (Friday), Saturday and Sunday in the winter and summer seasons. The daily pattern is characterized on the basis of the mean deviations in hourly flow compared to the average daily value.

Figure 8 shows the three indexes obtained considering the water-demand forecast for 1, 2, … , 24 h ahead, independently of the forecast starting point, for both the calibration and validation phases. In the calibration phase, the explained variance for the one-hour-ahead forecast is in excess of 0.97 and close to 0.96 for forecasts over the longer time horizons. The corresponding *EV* values in the validation phase are 0.97 for the one-hour-ahead forecast and 0.94 for forecasts over the longer time horizons. As might have been expected, in both cases the best forecast was obtained for the first hour ahead. This is understandable bearing in mind that the correlation, and hence the persistence, rapidly declines as the number of hours ahead increases. The other observation that can be made relates to the quality of the forecast, which does not vary significantly between the calibration and validation phases.

The *RMSE* and *MAE%* display a similar pattern to the one described for *EV*. In particular, it can be observed that in the calibration phase the *RMSE* is less than 3.5 l/s for the 1-hour-ahead forecast and less than 4 l/s for the 24-hour-ahead forecast, for an average hourly water demand of 54 l/s. This corresponds to a *MAE%* in the range of 4–6%. Similarly, in the validation phase the *RMSE* is less than 4.5 l/s for the 1-hour-ahead forecast and less than 8 l/s for the 24-hour-ahead forecast, for an average hourly water demand of 68 l/s, corresponding to a mean absolute percentage error of between 5–9%.

Again, the quality of the forecast is better for the first hour ahead and does not change significantly between calibration and validation phases. This is demonstrated in Figure 9, where the 1- and 24-hour-ahead water-demand forecasts for the calibration and validation phases are compared to the observed data. In all four cases, the points are near the 45° line, which represents an exact fit, but in the case of the 1-hour-ahead forecast (Figure 9(a) for calibration and Figure 9(c) for validation) the cluster shows less dispersion compared to the 24-hour-ahead forecast (Figure 9(b) for calibration and Figure 9(d) for validation). Moreover, the 1-hour-ahead forecast for the calibration phase (Figure 9(a)) shows a similar degree of dispersion to the 1-hour-ahead forecast for the validation phase (Figure 9(c)).

The statistics mentioned so far, which relate to different lead times independent of the forecasting start point, furnish a good estimate of the model's performance. However, they do not provide an adequate basis for analysing the errors incurred in relation to the different hours of the day to which the forecast refers. Indeed, the water demands observed during the day not only show different average values but also different standard deviations. The latter are largest in the early morning and

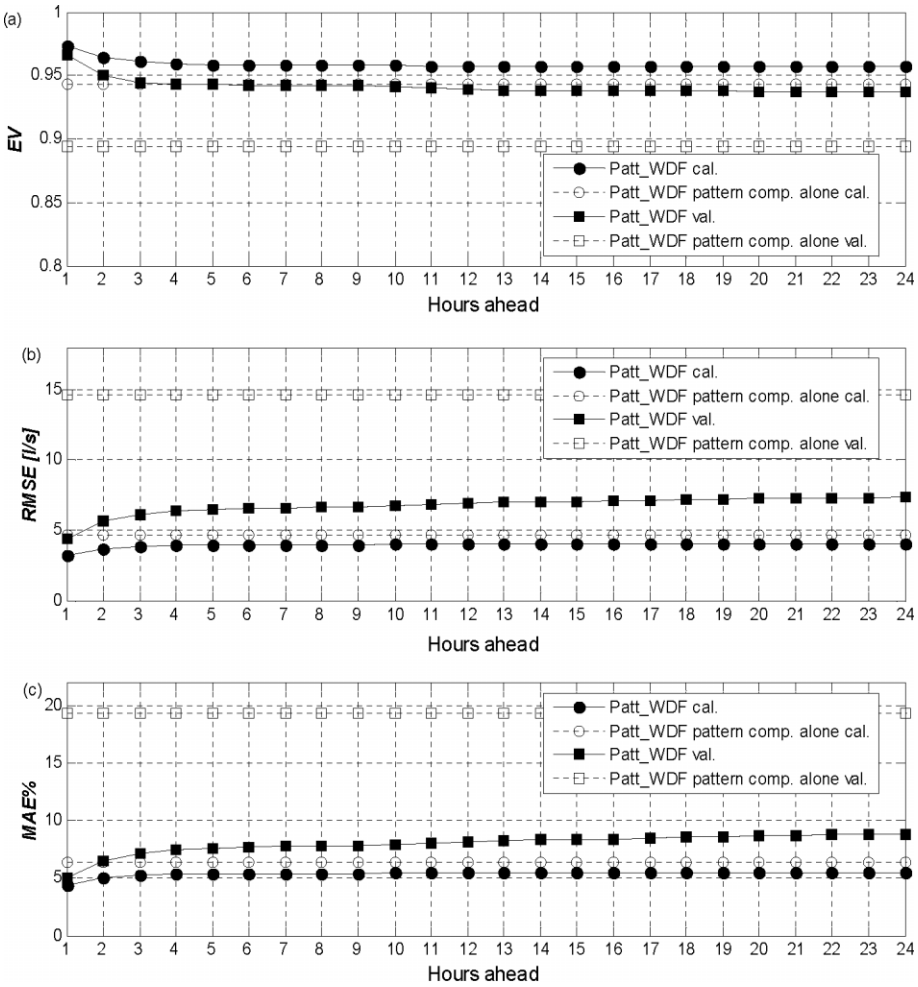


**Figure 8** | *EV*, *RMSE* and *MAE%* between the observed hourly water demands and those forecast with the Patt_WDF model, based on the periodic component alone and on the complete model, 1,2, … ,24 h ahead in the calibration (cal.) and validation (val.) phases.

evening hours due to the fact that consumption varies significantly at these times, resulting in a wider dispersion. Larger or smaller standard deviations in water demand at different times of the day will result in higher or lower forecasting precision. Figure 10 shows the *RMSE* and corresponding *MAE%* obtained in the calibration phase when the time of forecasting was fixed and the errors incurred in the predictions for the next 24 h were calculated. For example, in Figure 10(a), all the forecasts made at hour 1 for hours 2, 3, … , 24 of the same day and hour 1 of the next day are considered. Looking at the graphs in Figures 10(a), 10(b), 10(c) and 10(d), it can be seen that the *RMSE* pattern is similar but the curve shifts laterally according to the time of forecasting. The patterns only differ for the first or second hour following the time the forecast was made, revealing a slight decrease in the *RMSE* compared to when the water demands of the same hour of the day are forecast many hours in advance. This is due to the effect of short-term persistence, which leads to improvements in the forecasts for the hours that immediately follow the time of forecasting. Beyond this, the *RMSE* behaviour is solely a function of the hour of the day considered which is determined by the representation of the hourly water-demand patterns present in the hourly forecasting module. For example, the forecast of the water demand for hour 7 shows an *RMSE* equivalent to 6 l/s, regardless of whether the forecast was made at hour 1 of the same day (6 h ahead), at hour 12 of the previous day (19 h ahead), or at hour 18 of the previous day (12 h ahead), as indicated in Figures 10(a), 10(c) and 10(d). The *RMSE* is

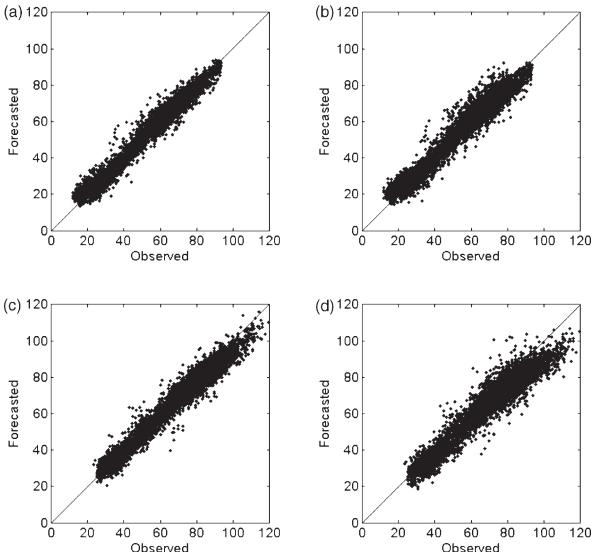


**Figure 9** | Water demand as forecast (a) 1 h ahead in the calibration phase, (b) 24 h ahead in the calibration phase, (c) 1 h ahead in the validation phase and (d) 24 h ahead in the validation phase, compared to the observed data.

slightly lower (5.2 l/s) only when the forecast is made 1 h ahead of time, that is to say at hour 6 (Figure 10(b)). Similar behaviour can be seen in the *MAE%* (Figures 10(e), 10(f), 10(g), 10(h)). In particular, it is evident that there is an increase in the percentage errors during the night-time hours which is due to the low values of the average water demand used to normalize the errors.

The above considerations find further confirmation in Figure 11. In this case, the fixed parameter was the hour of the day pertaining to the forecast. Here, the *RMSE* and *MAE%* relating to the forecasts for a specific hour, made at different times of the day, were calculated during the calibration phase. For example, Figure 11(a) shows the *RMSE* of the forecast for hour 1 made at hours 24, 23, …., 2. Again, the *RMSE* follows a similar pattern to that shown in Fig. 8, except in this case there are 24 different series associated with the 24 forecasted hours of the day. The curves obtained also confirm that the *RMSE* of the water-demand forecasts for a given hour do not depend on when the forecast was made, apart from the two or three hours immediately before the forecasted hour. On the other hand, *RMSE* values vary according to the hour pertaining to the forecast. For example, forecasts for hour 12 are associated with a lower *RMSE* than forecasts for hour 6, since the former are characterized by a lower degree of variability. Similar behaviour may be observed for *MAE%*, though higher values are associated with the night-time hours. This is due to the low values of the average water demand during the night which were used to normalize the errors and hence calculate the *MAE%*.

## Sensitivity analysis

The sensitivity of the forecasting model to its pattern and persistence components has been assessed on the basis of both its daily and hourly water-demand forecasts, taking into

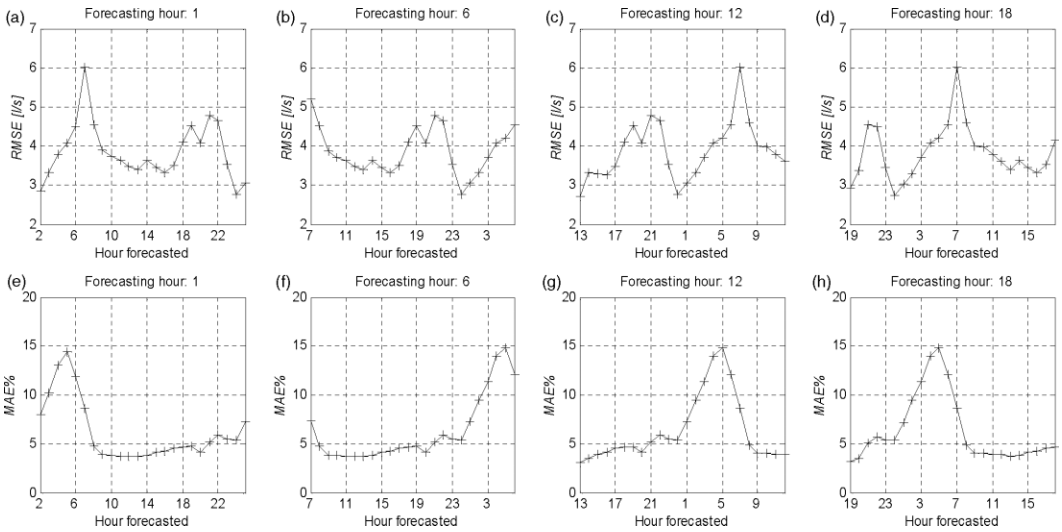


**Figure 10** | Patterns in *RMSE* and *MAE%* obtained by fixing the hour in which the forecast was made and calculating the errors in the forecasts for each of the next 24 h.

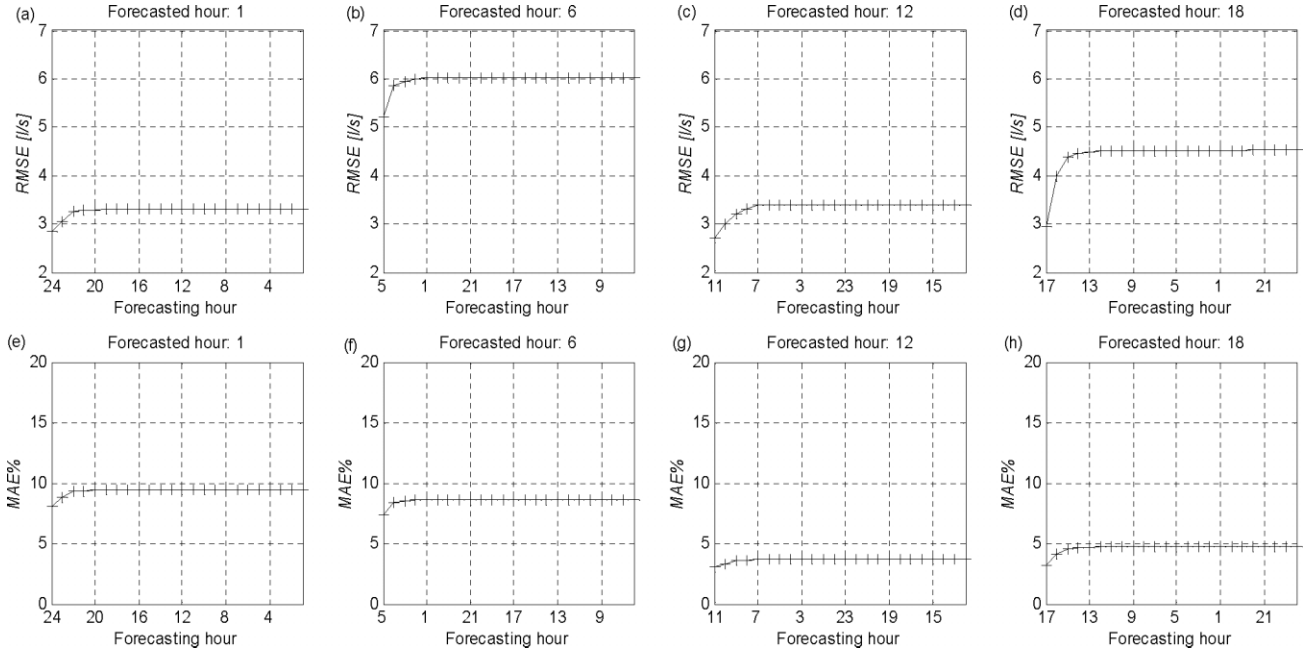


**Figure 11** | Patterns in *RMSE* and *MAE%* obtained by fixing the hour the forecast was for and calculating the errors in the forecasts made starting from each of the previous 24 h.

account, firstly, the periodic components alone and, secondly, the periodic and persistence components together. To that end, the forecasts were compared with the actual water demands observed in order to analyse the errors incurred for each case. More specifically, the *EV* and *RMSE* were calculated for the daily water-demand forecasts based on the patterns alone (the seasonal pattern $\bar{Q}_m^{d,s}$ and weekly pattern $\bar{\Delta}_{i,j}^{d,w}$), in addition to the forecasts based on both the periodic and persistence components (i.e. also taking account of the residual $\delta_m^d$, modelled using the autoregressive process $AR(1)$). Similarly, the *EV* and *RMSE* were calculated for the hourly water demand forecasts based on the patterns alone ($\bar{Q}_m^{d,s}$, $\bar{\Delta}_{i,j}^{d,w}$ and $\bar{\Delta}_{n,i,j}^{h}$) and the forecasts based on both the periodic and persistence components (i.e. also taking into account the daily residual $\delta_m^d$ as well as the hourly residual $\varepsilon_{t+k}$, modelled using a regression equation). Table 1 shows the values obtained in the four cases described for both the calibration and validation phases.

**Table 1** | Explained variance *EV* and *RMSE* of daily and hourly water demand as forecast on the basis of patterns alone and on the basis of both the periodic and persistence components of the Patt_WDF model in the calibration (average observed water demand 54 l/s) and validation (average observed water demand 68 l/s) phases

| | Calibration | | Validation | |
|---|---|---|---|---|
| | Daily | Hourly | Daily | Hourly |
| *EV* periodic comp. | 0.69 | 0.94 | 0.36 | 0.89 |
| *EV* periodic + persistence comp. | 0.84 | 0.97 | 0.76 | 0.97 |
| *RMSE* periodic comp. | 2.83 | 4.63 | 13.52 | 14.64 |
| *RMSE* periodic + persistence comp. | 2.01 | 3.17 | 4.61 | 4.42 |

In both instances, inclusion of the persistence component has the effect of fine-tuning the forecast. This can also be seen from Figure 8, where the *EV*, *RMSE* and *MAE%* relating to the hourly forecasts using the complete Patt_WDF model, are compared to the values obtained when only the periodic components of the model were considered. When the persistence components are absent, the *EV* takes on a constant value of about 0.94, regardless of lead time, in comparison with an *EV* of 0.97 for a lead time of 1 and an *EV* of 0.96 for all other lead times, using the complete model. In terms of mean absolute percentage error, omitting the persistence component would cause the *MAE%* to increase from 4.3% for a lead time of 1 and 5.4 for all other lead times, to 6.4%. Whilst the contribution of the persistence component for the calibration period is modest, it increases

markedly when the focus shifts to the validation period (see Table 1). In this latter case, the variance explained by the periodic component alone decreases significantly but this is compensated by the increased contribution of the persistence component. For example, when using the complete model at the hourly level, both the calibration and validation phases have an *EV* of 0.97 whereas for the periodic component alone, the *EV* deteriorates from 0.94 in the calibration phase to 0.89 in the validation phase. Similar trends can be seen in Figure 8 where, in the case of the complete model, the *MAE%* is 5.0% for a lead time of 1 and 7.7% for lead times greater than 1, whilst that for the pattern component alone is 19.3%.

The explanation of why the contribution of the persistence component increased in the validation phase is likely to be a result of the fact that the patterns used for validation were the same as those defined in the calibration year. Whilst using such patterns makes it possible to represent periodicities in water demand effectively, some changes may well occur from one year to the next. For example, although the observed water-demand patterns in Castelfranco Emilia have been very similar over the past few years, average consumption has undergone a slight increase, thereby increasing the size of the deviations. However, it would seem from this analysis that the persistence component is able to compensate for these systematic discrepancies.

## CONCLUSIONS

### Methodology developed

Prior to selecting this modular, pattern-based approach to short-term water-demand forecasting, consideration was given to using an artificial neural network (Stark *et al.* 1999) to serve the same end, which perhaps was more in keeping with the nature of the overall research project. Whilst the latter provided equally good results for a particular day, difficulties were encountered with the transition from one day to the next, especially between weekdays and weekends. By adopting a pattern-based approach, which contains seasonal, weekly and daily cycles, these difficulties could be avoided. However, reliance on the cyclic patterns alone also has its own shortcomings since amongst other things, no attention is paid to short-term exogenous variables such as the prevailing weather which can significantly influence demands on a day to day basis. Nevertheless, by including a persistence component as well as a periodic component, these limitations can, to a large extent, be overcome since such factors are implicitly taken into account by the hourly updating of the persistence component. Therefore, the resulting demand-forecasting model developed for the POWADIMA research project comprised two modules, each containing a periodic component and a persistence component. Whereas the aim of the first module was to forecast the daily water demands, that for the second module was to superimpose the hourly fluctuations. By this means, the observed seasonal, weekly and daily cycles can be preserved as well as the persistence effects.

### Attributes of the model developed

An analysis of the results obtained shows that the model is capable of delivering an accurate and robust forecast of future water demands on an hourly basis. Whilst the periodic component represents the mainstay of the model, the persistence components enables the forecast to be fine-tuned. This is of particular relevance to real-time control where the forecast is updated at short, regular intervals by 'grounding' any discrepancies between the actual demand at the next update and that forecast at the previous time-step. In this way, rather than just scaling the existing 24 h forecast in accordance with the actual demand, a new forecast can be generated at the next time-step, taking the internal structure of the demand patterns into account. However, it is important to stress that, in the light of the results obtained from the calibration and validation phases of the Castelfranco Emilia's water-demand forecasting model, consideration should be given to re-estimating the periodic and persistence parameters whenever there is a distinct increase in the number of customers served or a noticeable change in the per capita consumption.

## ACKNOWLEDGEMENTS

This water-demand forecasting model was developed as a constituent part of the POWADIMA research project,

which was funded by the European Commission under its Vth Framework thematic programme on Energy, Environment and Sustainable Development (Contract Number EVK1-CT-2000-00084). Members of the POWADIMA consortium would like to take this opportunity to thank the Commission and its project officers for their support. We also wish to acknowledge the assistance of Castelfranco Emilia Municipality and Meta Spa which acted as subcontractors to the Università degli Studi di Ferrara for this particular aspect of the project.

## REFERENCES

Alvisi, S., Franchini, M. & Marinelli, A. 2003 A stochastic model for representing drinking water demand at residential level. *Wat. Res. Mngmnt.* **17** (3), 197–222.

Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. 1994 *Time Series Analysis Forecasting and Control*, 3rd edn. Prentice-Hall Englewood Cliffs, NJ.

Franklin, S. L. & Maidment, D. R. 1986 An evaluation of weekly and monthly time series forecasts of municipal water use. *Wat. Res. Bull.* **22** (4), 611–621.

Jamieson, D. G., Shamir, U., Martinez, F. & Franchini, M. 2007 Conceptual design of a generic, real-time, near-optimal control system for water-distribution networks. *J. Hydroinformatics* **9** (1), 3–14.

Jowitt, P. W. & Germanopoulos, G. 1992 Optimal pump scheduling in water-supply networks. *J. Wat. Res. Plann. Mngmnt., ASCE* **118** (4), 406–422.

Maidment, D. R. & Parzen, E. 1984a Time patterns of water uses in six Texas cities. *J. Wat. Res. Plann. Mngmnt., ASCE* **110** (1), 90–106.

Maidment, D. R. & Parzen, E. 1984b Cascade model of monthly municipal water use. *Wat. Res. Res.* **20** (1), 15–23.

Maidment, D. R., Miaou, S. P. & Crawford, M. M. 1985 Transfer function models of daily urban water use. *Wat. Res. Res.* **21** (4), 425–432.

Martinez, F., Hernandez, V., Alonso, J. M., Rao, Z. & Alvisi, S. 2007 Optimizing the operation of the Valencia water-distribution network. *J. Hydroinformatics* **9** (1), 65–78.

Miaou, S. P. 1990 A class of time series urban water demand models with non-linear climatic effects. *Wat. Res. Res.* **26** (2), 169–178.

Rao, Z. & Alvarruiz, F. 2007 Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model. *J. Hydroinformatics* **9** (1), 15–24.

Rao, Z. & Salomons, E. 2007 Development of a real-time, near-optimal control process for water-distribution networks. *J. Hydroinformatics* **9** (1), 25–38.

Salomons, E., Goryashko, A., Shamir, U., Rao, Z. & Alvisi, A. 2007 Optimizing the operation of the Haifa-A water-distribution network. *J. Hydroinformatics* **9** (1), 51–64.

Shvartser, L., Shamir, U. & Feldman, M. 1993 Forecasting hourly water demands by pattern recognition approach. *J. Wat. Res. Plann. Mngmnt., ASCE* **119** (6), 611–627.

Smith, J. A. 1988 A model of daily municipal water use for short-term forecasting. *Wat. Res. Res.* **24** (2), 201–206.

Stark, H.L., Stanley, S.J. and Buchanan, I.L. 1999. Water demand forecasting using artificial neural network., *OWWA Conference paper, Ottawa*.

Sterling, M. J. H. & Coulbeck, B. 1975 A dynamic programming solution to the optimisation of pumping costs. *Proc. ICE* **59** (2), 813–818.

Zessler, U. & Shamir, U. 1989 Optimal operation of water distribution systems. *J. Wat. Res. Plann. Mngmnt., ASCE* **115** (6), 735–752.

Zhou, S. L., McMahon, T. A., Walton, A. & Lewis, J. 2002 Forecasting operational demand for an urban water supply zone. *J. Hydrol.* **259**, 189–202.