*Desalination and Water Treatment*
www.deswater.com
doi: 10.1080/19443994.2015.1065447

57 (2016) 11494–11505
May

# Using accurate demand forecasting to improve the efficiency of water supply-distribution chains

G. Arampatzis*, E. Kampragou, P. Scaloubakas, D. Assimacopoulos

*School of Chemical Engineering, National Technical University of Athens, 15780 Zografou, Greece, Tel. +30 2107723135; email: arab@chemeng.ntua.gr (G. Arampatzis), Tel. +30 2107721024; email: ekamp@chemeng.ntua.gr (E. Kampragou), Tel. +30 2106981836; email: pskaloub@otenet.gr (P. Scaloubakas), Tel. +30 2107723218; email: assim@chemeng.ntua.gr (D. Assimacopoulos)*

Received 31 January 2015; Accepted 15 June 2015

ABSTRACT

The efficient management of water supply-distribution systems requires the use of water demand forecasts so as to optimize the use of the resources involved (primarily water and energy) and minimize related costs. Particularly for operational water management, forecasting needs to be: (i) sensitive to rapidly changing factors (usually related to the water system management conditions), (ii) of a high level of accuracy and (iii) of high temporal resolution. Such a demand forecasting methodology, suitable for short-term water supply-distribution and management, has been developed and it is presented in this article, using the "similar days" approach. The method is based on the principle that "similar" days have similar consumption patterns. Similarity of any two days is judged by estimating a "day similarity index," which is a composite measure of the difference of the day's sets of common day attributes (such as the day of week, weather conditions, special events, etc.). The methodology is applied in two case studies, aimed at improving entirely different management procedures: (i) water pump operations at the distribution network of the city of Karlsruhe (Germany) and ii) water resource allocation for the metropolitan area of Barcelona (Spain), respectively. The case study applications confirm that the methodology is easily configurable, it fits well for quite different water management cases, and small forecasting errors can be achieved using readily available data.

*Keywords:* Water demand management; Demand forecasting; Similar days approach; Water demand drivers

## 1. Introduction

The water supply-distribution chain can be aggregated in three wide components [1]: (i) demand, related to the type of users in an area; (ii) water system, referring to the institutional (e.g. priorities in use), technical (e.g. networks) and economic (e.g. cost recovery issues) components that define the quantity and quality of water delivered to the users; and (iii) supply, corresponding to the amount of water abstracted from each

*Corresponding author.

*Presented at the 12th International Conference on Protection and Restoration of the Environment (PRE XII) 29 June–3 July 2014, Skiathos Island, Greece*

source in order to meet water demand. In this context, water demand forecasts are used by water managers for the planning, designing the infrastructure, operating and managing water supply systems, so as to optimize the use of the resources (mainly water and energy) and to minimize the related costs.

Water demand forecasting can be short-term (e.g. a few days), medium-term (e.g. monthly and up to 2–3 y) or long-term (several years), and can be performed at different levels of aggregation (e.g. a river basin, an agricultural system or a city) [2]. The selection of a water demand forecasting methodology is a function of three main criteria [3]: (i) planning objectives, (ii) data requirements and (iii) the availability of resources for the data collection, model set-up, development and calibration. The planning objectives are the most important criteria and define the level of detail needed by the decision-makers, who will utilize the water demand forecast information in their activities. Long-term management procedures (e.g. planning of large-scale investment for water supply enhancement) need forecasting to analyse the impact of changes in the demand drivers (e.g. water user population attributes, technology, price, income and climate) on water demand [4]. On the other hand, short-term management procedures (i.e. operational management) need forecasting to minimize the use of resource inputs and increase the reliability in water use delivery. They typically use different determinants for describing the water demand pattern.

Increasing importance is given to improving the efficiency of short-term management procedures, particularly for operational water management that may provide more immediate, certain and measurable benefits to the water system actors, the environment and the society. Fig. 1 shows the various stages along the water supply-distribution chain, each having more-or-less different short-term demand forecasting requirements for management purposes. Consequently, the development and testing of suitable water demand short-term forecasting methodologies, which can become part of and enhance the efficiency of these procedures, becomes increasingly important.

The methods typically used in short-term forecasting are probabilistic methods, memory-based learning techniques, time series models, neural networks and hybrid methods [5]. An overview of the approaches in-use is presented in Table 1. The desired properties of a short-term forecasting methodology are to be practical, readily applicable to all points/procedures along the water chain, and able to provide real immediate benefits in terms of cost reduction and resource efficiency.

The statistical trends analysis (time series) type of models that are commonly used for short-term forecasts [6–8] might provide good short-term predictions, but these depend on the day-to-day stability of water consumption and the very short-term effect of "hidden" system variables, as explained in [9]. A methodology that is able to encode such variables in a flexible way and to explain day-to-day demand variation is the "similar days" method which in addition can meet the requirements pinpointed above. The method has been mainly used for forecasting energy demand [10,11] and electricity prices [12] but has also been applied in water demand forecasting [13,14]. The method emulates the way that an expert controlling some part of the water supply-distribution chain (e.g. scheduling operation of pumps) does his guesswork about what a day's demand pattern might look like, by looking back on what happened on a past day that seems similar. Typically, it is used to forecast demand for the next 1–2 d. However, it may be used as far ahead as desired, provided that the states of the

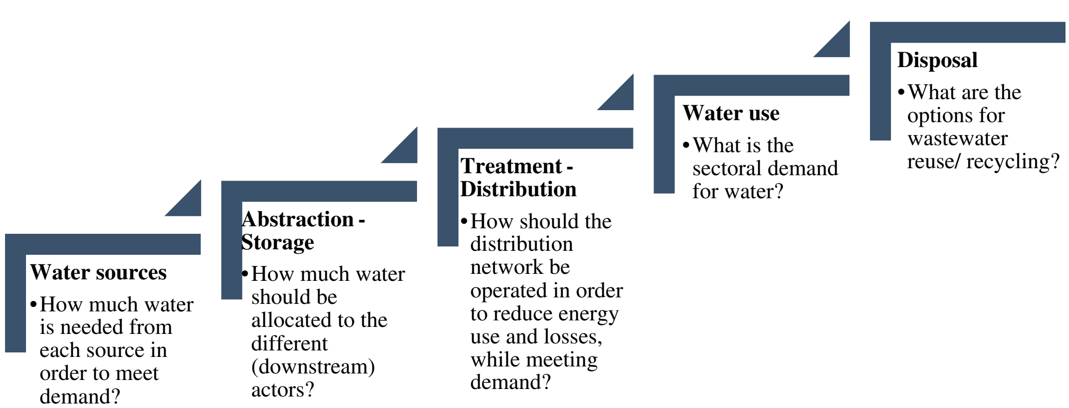


Fig. 1. Guiding questions related to water demand forecasts along the water supply-distribution chain.

Table 1
Methods typically used in short-term forecasting

| Approach | Advantages | Disadvantages |
|---|---|---|
| Time series analysis<br>*(Extracts statistic information about periodicities and other trends in past observations of the forecasted (depended) variable)* | (1) The method is relatively cheap to implement [1]<br>(2) It is based on records of past uses<br>(3) Rather low data requirements<br>(4) Depending on data availability, demand forecasting can be related to different use categories [1] | (1) Forecasting errors, due to changes in trends, do not provide a sound basis for future learning<br>(2) Forecasting accuracy varies depending on the fitting function<br>(3) The method assumes that the historical water use pattern is representative of future water demand<br>(4) Extrapolation techniques typically provide estimates of aggregate demands rather than components of demand [1] |
| Pattern recognition<br>*(Correlates consumption patterns to patterns of daily conditions to select periods of similar pattern)* | (1) Low computational requirements [14]<br>(2) Easy to incorporate new variables or states in the model [14]<br>(3) Low to medium data requirements [14] | (1) Forecasting accuracy depends on the calibration of the variables' values |
| Artificial neural networks<br>*(Emulates in a fuzzy way like other types of mathematical models, by implementing a process for learning from a multitude of input data sets)* | (1) Ability to learn from the facts or input data and the associated output data [15]<br>(2) The use of ANN does not require a priori knowledge of the process [16]<br>(3) ANN models are effective with nonlinear data [16] | (1) ANN models are usually difficult to automate since ANN training may need supervision from experts and models may need more frequent update<br>(2) A large number of high-quality training correlated data sets are required to achieve high forecasting accuracy<br>(3) Rather high data requirements |

system variables can be forecasted. Besides using relatively easy to find data, the methodology is quite extensible, since the forecast time step (day, hour and minutes) can be parametric and the system variables themselves can be also parametric. Thus, system experts using the model may continuously improve its predictive power and fit it to changing conditions and system events.

The aim of this article was to develop a model for short-term water demand forecasting based on the "similar days" methodological approach. Emphasis is given to the methodological steps required to develop, calibrate and apply the model in real situations. A probabilistic view of the model results is also presented in order to support the risk analysis of the estimated demand forecasts. The methodological steps are implemented in two case studies, representing two entirely different water chain management procedures, with quite different goals. The applications help demonstrate that the methodology is easily configurable and it fits well in quite different water management cases and operating environments. Results from both applications show that small forecasting errors can be achieved by utilizing readily available data and the proposed method can be applied in short-term management procedures at all points along the water supply-distribution chain that require or can potentially use demand forecasts.

## 2. Methodology

### 2.1. Model equations

The proposed methodology is based on the "similar days" pattern recognition approach, i.e. on the principle that days having similar conditions (i.e. levels of factors affecting demand) are expected to exhibit also quite similar consumption levels. Future water demand is forecasted by estimating a weighted average of actual historic water consumption data that refer to a set of days in the past, with attributes similar to the forecast day. Thus, the model comprises two stages. At the first stage, a set [S] of $N$ d from the recent past, similar enough to the forecast day, is selected on the basis of a number of criteria (i.e. independent explanatory variables of the model). At the second stage, demand is predicted by estimating a weighted average of the recorded historic consumption values for this particular set of similar days, using the formula:

$$d_{\mathrm{f}} = \sum_{i=1}^{N} w_i \cdot c_i \tag{1}$$

where $d_{\mathrm{f}}$ is the forecasted water demand (daily, hourly or $n$-minute) for the forecast day f, $c_i$ is the observed water consumption of similar day $i$ for the same time step and $w_i$ is the influence (i.e. statistical weight) of similar day $i$ on the demand forecasted.

Past days are included in the similarity set [S] if they are: (i) sufficiently similar to the forecast day and (ii) recent enough, so that the basic hypothesis of the model, i.e. the correlation between day similarity and water consumption, holds true.

The similarity set must be large enough to account for sufficient statistical variation of water consumption and small enough, so that outlier values are not included. Past days are considered similar, if their characteristics (similarity factors) are very close to those of the forecast day. A quantitative expression of this concept of similarity is provided by the similarity index. It is calculated using the formula [14]:

$$r_i = \frac{\sum_{k=1}^{M} v_{ik} v_{\mathrm{f}k}}{\sqrt{\sum_{k=1}^{M} v_{ik}^2 \cdot \sum_{k=1}^{M} v_{\mathrm{f}k}^2}} \tag{2}$$

where $r_i$ is the similarity index for day $i$, $M$ is the number of factors influencing demand, $v_{ik}$ is the mapping value of factor $k$ for day $i$ and $v_{\mathrm{f}k}$ is the mapping value of factor $k$ for forecast day f. Mapping values of similarity factors are explained in the next paragraph (Section 2.2).

The index estimated by formula (2) is a measure of the vector distance of the days $i$ and f in the $M$-dimensional similarity factor space. It is analogous to a correlation coefficient and its values are assessed similarly. Past days with sufficiently high similarity index values are considered similar to the forecast day and they are included into the similarity set to be used in the demand forecast estimation formula (1). Sufficiently high are the values above a pre-defined threshold $r_{\mathrm{c}}$ (e.g. $r_i > 0.9 = r_{\mathrm{c}}$). This is a model parameter, representing the similarity condition (i), while a second model parameter, representing condition (ii), is the cut-off point in the consumption timeline (i.e. the number of past days $t_{\mathrm{c}}$ considered for inclusion in [S]) for the similarity set. The values of both parameters are determined during the calibration of the model, by varying in steps $r_{\mathrm{c}}$ and $t_{\mathrm{c}}$ and selecting the values that most consistently minimize the prediction error.

The similarity indices of days included in the similarity set [S] can be used to calculate the relative weights of similar days in formula (1), i.e. weights are linearly analogous to the degree of similarity, according to the following formula:

$$w_i = \frac{r_i}{\sum_{j=1}^{N} r_j} \tag{3}$$

The methodology has the advantage that, due to the way that the similarity set is compiled, error statistics and confidence levels of the resulting forecast can be directly estimated using standard statistical methods [17], taking into account that the most probable values and their probabilities (as estimated by the weights $w_i$

of similar days) have already been calculated. Therefore, risk analysis, on the use of the estimated forecasts in water management procedures, can be easily performed. This can be necessary in some resource optimization cases, where the consequences of a wrong decision could be serious enough.

### 2.2. Methodological steps

The practical application of the forecasting methodology presented in the previous paragraph can be summarized into ten procedural steps:

(1) Identify the factors influencing water demand to be included in the forecasting model.
(2) Establish mapping tables, to obtain the range of values for each similarity factor.
(3) Compile a time series of data for consumption and all related similarity factors.
(4) Obtain/estimate predictions of the mapping values of similarity factors for the forecast day.
(5) Calculate the similarity index for all days in the time series database, which fall within the cut-off time limit $t_c$, using formula (2).
(6) Select the set [S] of similar days to the forecasted day, by applying threshold $r_c$.
(7) Calculate the weights/probabilities of the similar days in [S], using formula (3).
(8) Estimate the demand for the forecast day, using formula (1).
(9) Estimate error statistics and confidence levels.
(10) Perform sensitivity and risk analysis.

Steps 1 to 3 are implemented during the set-up and calibration of the model in a new area, steps 3 to 8 are regular steps performed each time a demand forecast is required and steps 9 and 10 are part of the optional risk analysis phase.

In step 1, a set of similarity factors is selected and incorporated to initialize the mapping table. Such typical factors, expressing the drivers of water demand on a day-by-day basis, are the type of day (e.g. normal day, public holiday, strike, major sports or other special event, etc.), day of the week, month, temperature, precipitation, etc. Selection of factors is done either empirically or through the statistical study of co-variance and they should satisfy two rather obvious criteria of applicability. It should be possible to obtain and update (up to present day) a time series of historic data values, correlated to the corresponding water consumption, and it should be feasible to predict the value of each factor for the forecast day. The main drivers of demand as expressed in the similar days model through the similarity factors and model parameters are depicted in Fig. 2.

Similarity factors can be qualitative or quantitative, taking nominal or numerical values. In order for all factors to be comparable and usable in formula (2), quantitative factors (e.g. temperature) need to be converted into qualitative ones, by subdividing their ranges into numerical intervals and correlating those with discreet day states (nominal values). Every state needs to be mapped to a numerical ranking value in order to be entered in formula (2). This is done in step 2, where the mapping table is filled with numeric values for each possible state (class) of each factor, in such a way as to capture the importance of each factor and provide a basis for the comparison of days. The mapping of the factor states to ranking values is accomplished using the statistical distribution of water consumption values for each state of each factor from available time series data. Fig. 3 illustrates the concept of the mapping database (note that the numeric ranking values are normalized in the [0…100] interval). The final mapping values (weights of similarity) are obtained by multiplying the normalized value of each state with a normalized weighting coefficient for each similarity factor. Thus, the relative importance of each factor on water demand prediction is taken into account. Following that and provided that sufficient data have been made available (step 3) and that the factors' values for the forecast day are predicted (step 4), the actual calculation of the forecasted demand (steps 5–8) using formulas (1)–(3) is quite straightforward.

### 2.3. Implementation issues

The data required to support the forecasting procedure (steps 3 and 4) are composed of three main datasets (water consumption, day characterization and meteorological data). Readily available historical consumption data are used, that should ideally be in a continuous time series, ideally continuous, extending at a minimum for one year from the present day to the past, while a 3-y time series is recommended to ensure that a large enough number similar days will always exist. Note that the time resolution of the forecasts depends only on the resolution of the historical consumption data; thus, the described methodology can as easily forecast hourly or $n$-minute water demand, provided that consumption data of the same resolution (e.g. hourly) are included in the time series.

Since forecasted demand is calculated directly from historical consumption data, the quantity estimated

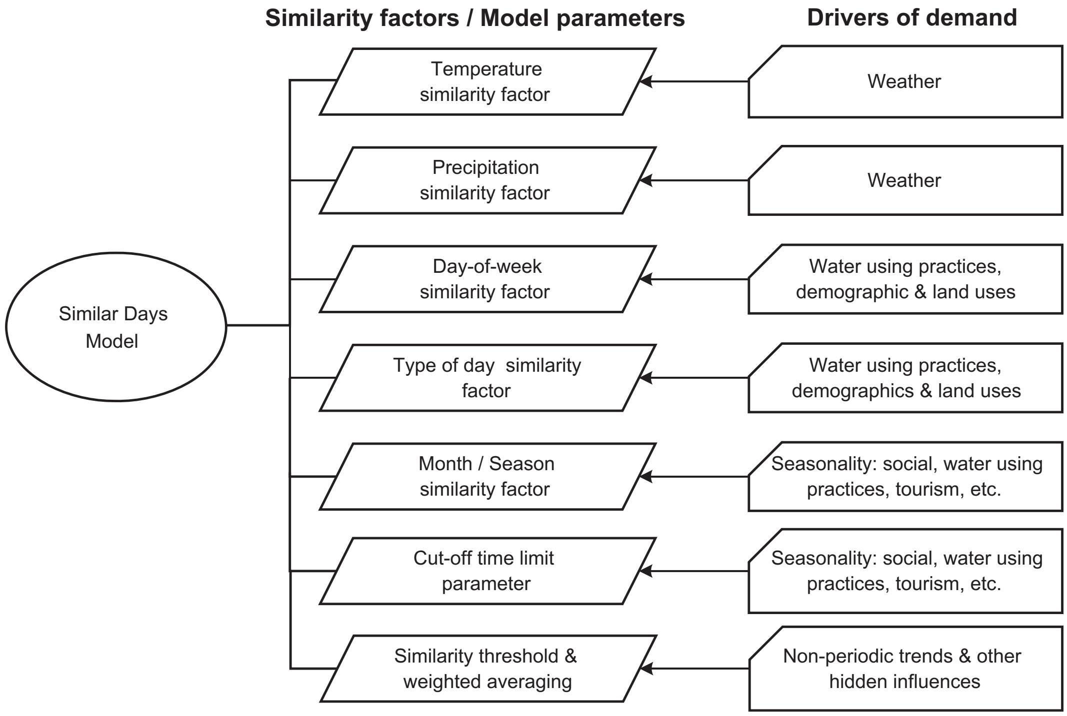


Fig. 2. Similar days model variables and drivers of demand expressed by them.

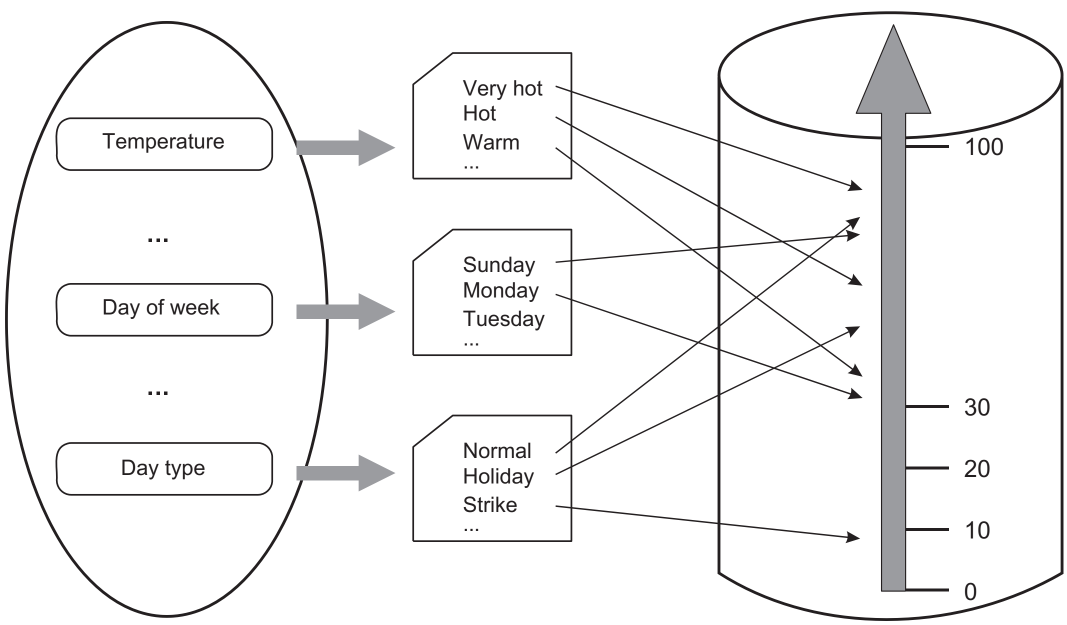


Fig. 3. Mapping similarity factor values.

corresponds exactly to the topological frame of reference of these data. Therefore, the forecast can actually be a prediction of the flow through some node in the water supply-distribution network, where consumption is measured. Forecasts, in addition to regular water use, also incorporate any regular water losses (e.g. leaks and unaccountable water use), since these are incorporated in the historical data. One-off events, such as breaks, unscheduled infrastructure maintenance or supply restriction events, or other abnormal conditions cannot be forecasted (they could be possibly predicted as stochastic events in long-range forecasting, e.g. at a yearly basis). The influence of scheduled events on demand can potentially be predicted by the model, provided that there is sufficient historical data available and that a specific similarity factor is included for this purpose into the model.

The model could be used to forecast as far ahead in time as desired. However, there is a practical limitation, since the time scope is limited by the need to have an accurate forecast of the values of the similarity factors. Since these factors include meteorological conditions, this limits the useful time scope of the model to about 7 d ahead. This is not really a problem, since the model is intended to be used in short-term water management procedure, typically for planning and/or controlling next-day of at most next 48- to 72-h operations.

## 3. Application in case studies

The methodology is applied in two case studies, aimed at improving entirely different management procedures at different points along the supply-distribution chain and at regions exhibiting quite different conditions: (i) water pump operations at the downstream distribution network of the city of Karlsruhe (Germany) and (ii) water resource allocation at the upstream management points of the water supply system for the metropolitan area of Barcelona (Spain), respectively.

The water utility of Karlsruhe (SWKA) serves about 300,000 inhabitants. The supply is ensured by four "Waterworks." The drinking water is obtained through well fields, pumping water from the ground water layer into the treatment sites. The main storage reservoir of the city of Karlsruhe works as a stabilizer of water supply and pressure.

The water supply system of the Barcelona is deployed in a composite river basin area comprising of the basins of the rivers Ter, Daró, Tordera, Besòs, Llobregat and Foix, as well as several small coastal stream basins. The watersheds of Ter and Llobregat typically provide 81% of the water supply, while the remaining 19% is provided by desalination plants, regenerated water from waste water treatment plants and groundwater usage. The main water uses within the river basin include the combined urban–industrial water use (this being the primary use), irrigation water use and environmental use.

Forecasting is incorporated as an integral part of the two entirely different water resource management procedures. In Karlsruhe, hourly forecasts are used to optimize the operation of numerous pumping stations, thus minimizing the use of energy and costs. In Barcelona, daily forecasts are used to estimate the water that should be supplied daily from each one of five main water resources, in order to meet demand and reduce the amount of excess water rejected to the sea.

### 3.1. Model set-up and calibration

#### 3.1.1. City of Karlsruhe

The model is implemented in the city of Karlsruhe using historical data on hourly water consumption for a range of 32 months (May 2010–December 2012). Daily meteorological data (minimum, maximum and average temperatures as well as precipitation height) were also collected for the same time period. These factors are the most significant, it is easy to obtain historical values and feasible to predict their values for the forecast day. Other meteorological parameters (such as relative humidity, wind and cloudiness), that have been shown to influence water consumption [9], may be used as far as a reliable source of information exists. Missing values form the meteorological data set were not filled-in, but the whole days, including the corresponding consumption data, were removed from the database. Moreover, pre-processing of the consumption data set helped in identifying and removing outliers.

The daily aggregated water consumption time series and the correlated average temperature are presented in Fig. 4. The figure reveals a weak positive correlation between the two data sets (correlation coefficient equal to 0.4) with higher consumption during the summer. Unusual high water consumption values are reported during the last week of July 2010 and the first two weeks of August 2010. These values, as well as other unusual individual peaks shown in Fig. 4, are attributed to abnormal conditions (network losses) and to high amounts of water that delivered to other town/villages surrounding Karlsruhe (external distributors). They have been marked as outliers in the database and were not included in the analysis.

Temperature variables were directly used as similarity factors in the forecasting model while the precipitation was used to calculate a "days after rainfall" similarity factor, which has been shown to be correlated with the water consumption. Other similarity factors used are the month, the day of week, the hour of the day and the day type. The latter is used to categorize days into four types: normal day, public holiday, school holiday and public/school holiday. Seasonal variation of water consumption is indirectly considered in the analysis using the month as a similarity factor in the model. In addition, the climate conditions in the city are not so extreme to justify

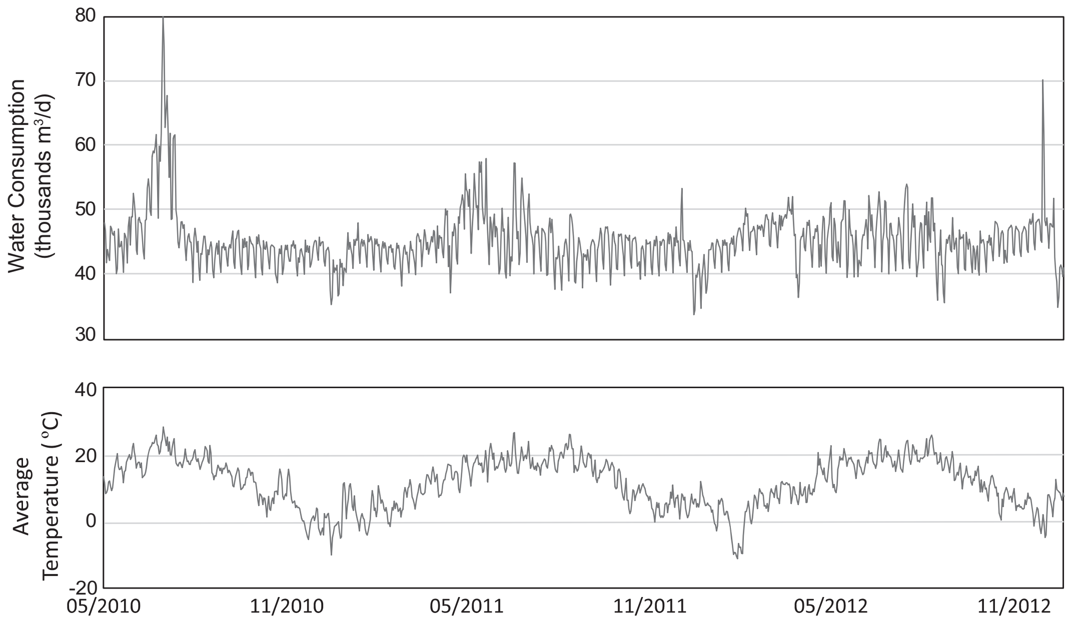


Fig. 4. Daily water consumption and average temperature in the city of Karlsruhe.

significant differences in consumption e.g. during summer.

Fig. 5 presents the variation of daily water consumption by the type of day and the day of week. These distributions have been used to provide the mapping values of the two similarity factors (presented in Table 2), as explained in Section 2.2.

Fig. 5(a) reveals that the water consumed in the city of Karlsruhe during public holidays is almost identical to the water consumed during school holidays. This fact is expressed by the identical mapping values of these two states of the "Day Type" factors, as shown in Table 2. The same applies to "Day of Week" factor, where identical mapping values are used for tuesday and wednesday as well as thursday and friday.

*3.1.2. Metropolitan area of Barcelona*

The application of the model to the Metropolitan area of Barcelona uses historical data on daily water consumption for a range of six years (January 2006–December 2012). Meteorological data (minimum, maximum and average temperature, precipitation height) and a day characterization data set were also provided.

Fig. 6 presents the water consumption time series and the correlated average temperature values. A positive correlation between the two data sets is again obvious (correlation coefficient equal to 0.45) with higher consumption during the summer. Some unusual low water consumption values are identified, that correspond to unscheduled infrastructure maintenance events, due to damage of the main tube of water.

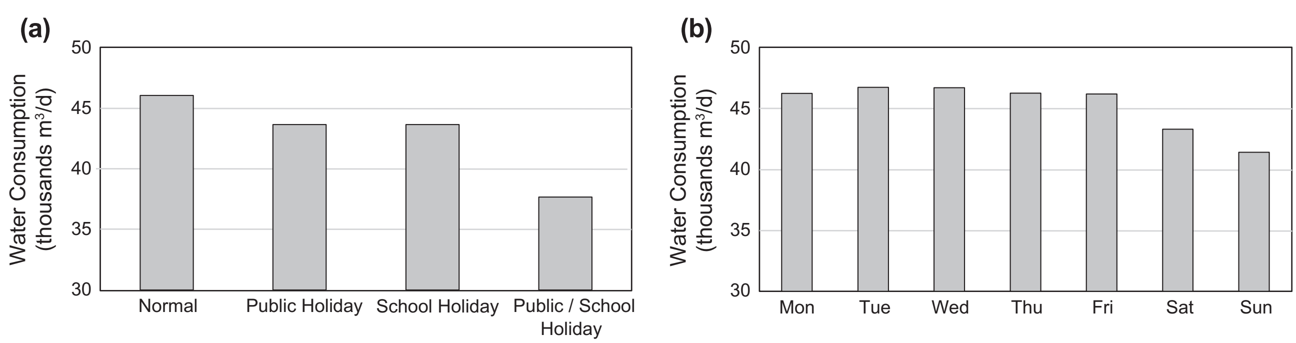


Fig. 5. Variation of daily water consumption in the city of Karlsruhe (a) by type of day and (b) by day of week.

Table 2
States and corresponding mapping values of "Day Type" and "Day of Week" similarity factors for the city of Karlsruhe

| Day type | Mapping value | Day of week | Mapping value |
|---|---|---|---|
| Normal | 100 | Monday | 90 |
| Public holiday | 60 | Tuesday | 100 |
| School holiday | 60 | Wednesday | 100 |
| Public/school holiday | 0 | Thursday | 90 |
| | | Friday | 90 |
| | | Saturday | 50 |
| | | Sunday | 0 |

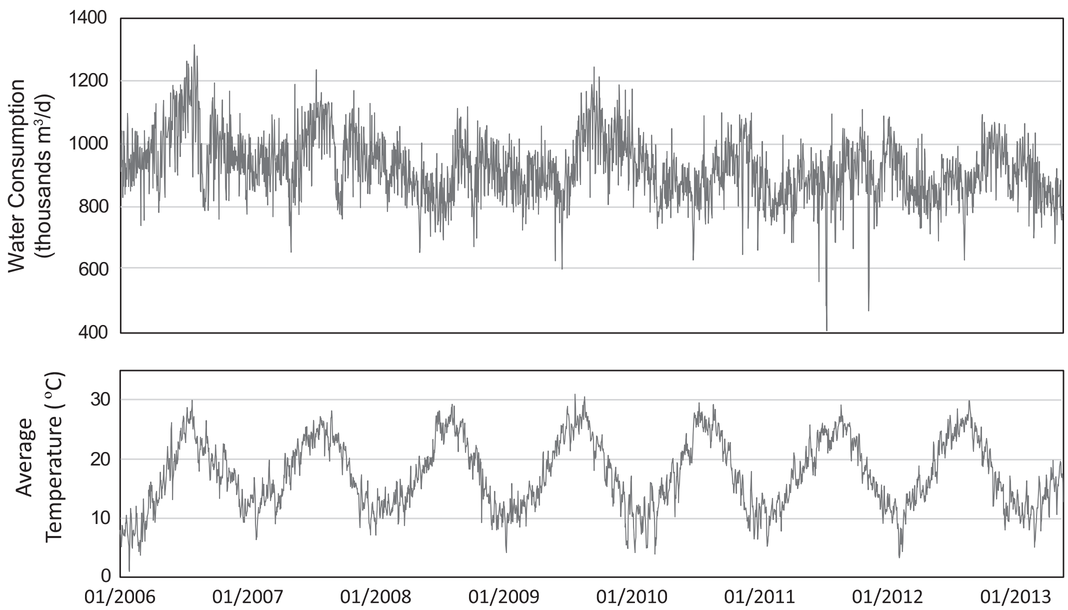


Fig. 6. Daily water consumption and average temperatures in the area of Barcelona.

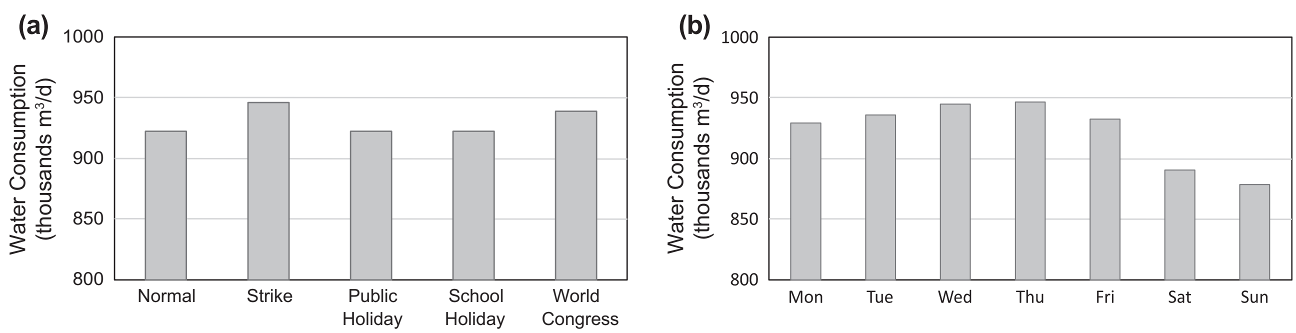


Fig. 7. Variation of daily water consumption in the area of Barcelona (a) by type of day and (b) by day of week.

These days have been marked in the database as outliers and not included in the analysis.

The day characterization data set uses five discrete day types: normal day, strike, public holiday, school holiday and world congress. Fig. 7 presents the variation of water consumption by the type of day and day of week while Table 3 presents the corresponding mapping tables. As in the case of Karlsruhe, seasonal variation of

Table 3
States and corresponding mapping values of "Day Type" and "Day of Week" similarity factors for the area of Barcelona

| Day type | Mapping value | Day of week | Mapping value |
|---|---|---|---|
| Normal | 50 | Monday | 60 |
| Strike | 100 | Tuesday | 75 |
| Public holiday | 10 | Wednesday | 90 |
| School holiday | 0 | Thursday | 100 |
| World Mobile congress | 80 | Friday | 60 |
| | | Saturday | 10 |
| | | Sunday | 0 |

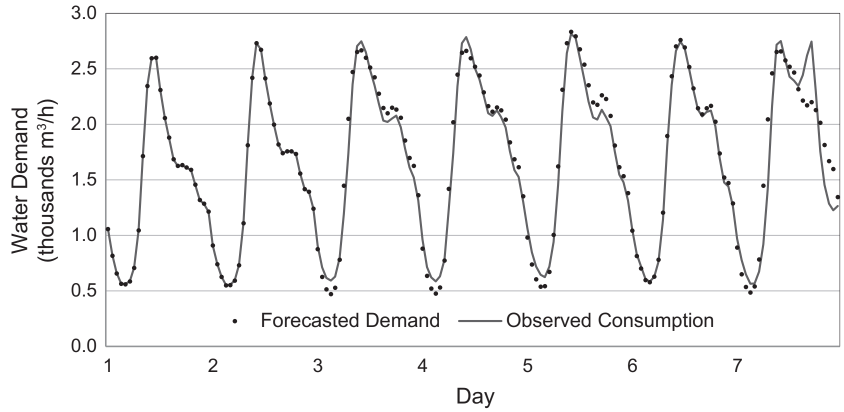


Fig. 8. Hourly forecasted water demand and observed consumption in the city of Karlsruhe.

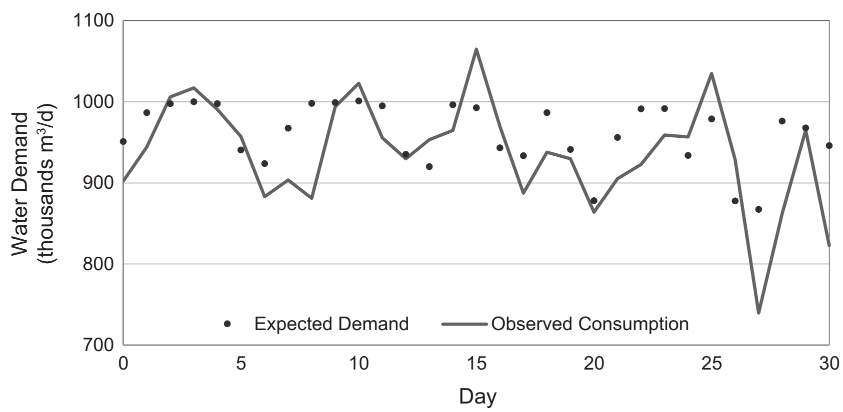


Fig. 9. Daily forecasted water demand and observed consumption in the area of Barcelona.

water consumption is indirectly considered in the analysis using the month as a similarity factor in the model.

### *3.2 Model implementation and validation*

#### *3.2.1. City of Karlsruhe*

The hourly demand forecasting model is tested against a data subset of 168 h (corresponding to the last week of December 2012). These data were excluded from the data set used in the mapping database.The forecasted demand is compared to the actual consumption, and the results are presented in Fig. 8. The model performs satisfactorily with a mean absolute percentage error (MAE) equal to 5.4% and a root mean square error (RMSE) of 0.13 thousands $m^3$/h.

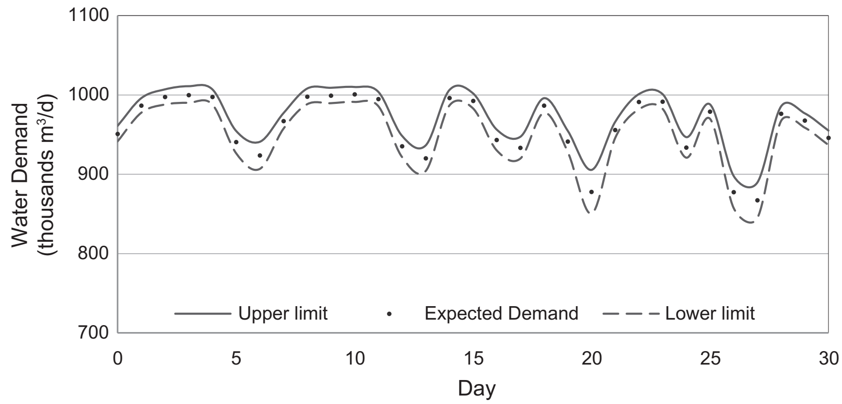


Fig. 10. 95% confidence interval of forecasted demand in the metropolitan area of Barcelona.

### 3.2.2. *Metropolitan area of Barcelona*

The model is applied to provide daily water demand forecasts for 30 test days (December 2012). The results presented in Fig. 9 show that the model performs satisfactorily, with a MAE equal to 6% and a RMSE equal to 56,000 $m^3$/d.

A unique feature of the model is that, besides the forecast value, it also produces a number of similar days with their consumptions. These data can be analysed, using standard statistical methods to provide ranges of demand or even confidence levels. As an example, Fig. 10 presents the confidence interval for the forecasted water demand at a 95% level.

## 4. Conclusions

The possibility, even need, for the improvement of the efficiency of short-term management procedures along the water supply-distribution chain, through accurate and practical water demand forecasting, has been stressed. A methodology for short-term water demand forecasting, based on the "similar days" approach, has been presented, as a candidate for this role. It has many advantages, most importantly that of requiring easily obtainable data and of providing accurate demand forecasts at almost any resolution (hourly or *n*-minute) up to daily. It is transparent, i.e. easy to comprehend, improve and calibrate and, due of that, has the side benefit of being able to provide, for its users, new knowledge about the water system they manage. It can also produce measures of confidence in and variation/distribution of forecast values, thus enabling the use of probabilistic (risk) analysis, necessary for resource optimization.

The methodology has been applied, for demonstration and testing, to two case studies, representing two entirely different management procedures along the water supply-distribution chain and at regions that have quite different attributes. Results obtained from these applications help demonstrate that the methodology fits well in different water supply-distribution management procedures, which are already in use, while no particular data collection and data processing. Sufficiently small forecasting errors (e.g. 4.5% in the case of Karlsruhe and 6% in the case of Barcelona) can be achieved using readily available data, without excessive calibration effort. It is concluded that this methodology can potentially provide accurate results and it could be of high value for practical use, in cases where very short-term forecasting with high accuracy and temporal resolution is a requirement. Applying the methodology to other nodes and processes along the water supply-distribution chain or even to other water chains is straightforward.

## Acknowledgements

This work has been financed by the EU-funded FP7 WatERP project (Grant agreement no: 318603). The authors would like to thank partners from Water Catalan Agency (ACA), Deutsche Vereinigung des Gas-und Wasserfaches (DVGW) and Stadtwerke Karlsruhe GmbH (SWKA) for the provision of the data used in the analysis.

## References

[1] D. Lumbroso, Handbook for the Assessment of Catchment Water Demand and Use, HR Wallingford & Department for International Development, Wallingford, Oxon, 2003.

[2] M.L. Froukh, Decision-support system for domestic water demand forecasting and management, Water Resour. Manage 15 (2001) 363–382.

[3] W.Y. Davis, Water Demand Forecast Methodology for California Water Planning Areas—Work Plan and Model Review, California Bay-Delta Authority, Corbondale, Illinois, IL, 2003.
[4] S. Renzetti, Water demand forecasting, in: A. Dinar and D. Zilberman (Eds.), The Economics of Water Demands, Natural Resource Management and Policy, vol. 22, Springer, Boston, MA, 2002, pp. 145–156.
[5] J. Bougadis, K. Adamowsk, R. Diduch, J. Bougadis, Short-term municipal water demand forecasting, Hydrol. Processes 19 (2005) 137–148.
[6] S. Alvisi, M. Franchini, A. Marinelli, A short-term, pattern-based model for water demand forecasting, J. Hydroinf. 9 (2007) 39–50.
[7] S. Gato, N. Jayasuriya, P. Roberts, Temperature and rainfall thresholds for base use urban water demand modelling, J. Hydrol. 337 (2004) 364–376.
[8] S. Zhou, T. McMahon, A. Walton, J. Lewis, Forecasting daily urban water demand: A case study of Melbourne, J. Hydrol. 236 (2000) 153–164.
[9] V. Kanakoudis, K. Gonelas, Forecasting the residential water demand, balancing the full water cost pricing and the non-revenue water reduction policies, Procedia Eng., 89 (2014) 958–966.
[10] E. Srinivas, A. Jain, A methodology for short term load forecasting using fuzzy logic and similarity, in: The National Conference on Advances in Computational Intelligence Applications in Power, Control, Signal, Processing and Telecommunications (NCACI), Bhibaneswar, 2009.
[11] J. Campillo, F. Wallin, D. Torstensson, I. Vassileva, Energy demand model design for forecasting electricity consumption and simulating demand response scenarios in Sweden, in: International Conference on Applied Energy ICAE 2012, Suzhou (2012).
[12] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, A.K. Srivastava, A novel approach to forecast electricity price for PJM using neural network and similar days method, IEEE Trans. Power Syst. 22 (2007) 2058–2065.
[13] Q. Mu, Y. Wu, X. Pan, L. Huang, X. Li, Short-term load forecasting using improved similar day method, in: Power and Energy Engineering Conference, Chengdu, 2010.
[14] G. Arampatzis, N. Perdikeas, E. Kampragou, P. Scaloubakas, D. Assimacopoulos, A water demand forecasting methodology for supporting day-to-day management of water distribution systems, in: A. Liakopoulos, A. Kungolos, C. Christodoulatos, A. Koutsospyros (Eds.), 12th International Conference "Protection & Restoration of the Environment", Skiathos, Thessaloniki, Greece, 2014.
[15] A. Jain, A.K. Varshney, U.C. Joshi, Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks, Water Resour. Manage 15 (2001) 299–321.
[16] J. Adamowski, H. Fung Chan, S.O. Prasher, B. Ozga-Zielinski, A. Sliusarieva, 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 (2012) W01528, doi: 10.1029/2010WR009945.
[17] D.F. Gatz, L. Smith, The standard error of a weighted mean concentration—I. Bootstrapping vs other methods, Atmos. Environ. 29 (1995) 1185–1193.