Hindawi
Mathematical Problems in Engineering
Volume 2017, Article ID 6343625, 10 pages
https://doi.org/10.1155/2017/6343625

Research Article

Correlation Analysis of Water Demand and Predictive Variables for Short-Term Forecasting Models

B. M. Brentan,[1] G. Meirelles,[2] M. Herrera,[3] E. Luvizotto Jr.,[2] and J. Izquierdo[4]

[1] *Centre de Recherche en Automatique de Nancy, Université de Lorraine, Nancy, France*
[2] *Laboratório de Hidráulica Computacional, Faculty of Civil Engineering, Universidade Estadual de Campinas, Campinas, SP, Brazil*
[3] *EDEn, Department of Architecture and Civil Engineering, University of Bath, Bath, UK*
[4] *FluIng, Institute for Multidisciplinary Mathematics, Universitat Politècnica de València, Valencia, Spain*

Correspondence should be addressed to J. Izquierdo; jizquier@upv.es

Received 28 August 2017; Accepted 29 November 2017; Published 18 December 2017

Academic Editor: Sergey A. Suslov

Operational and economic aspects of water distribution make water demand forecasting paramount for water distribution systems (WDSs) management. However, water demand introduces high levels of uncertainty in WDS hydraulic models. As a result, there is growing interest in developing accurate methodologies for water demand forecasting. Several mathematical models can serve this purpose. One crucial aspect is the use of suitable predictive variables. The most used predictive variables involve weather and social aspects. To improve the interrelation knowledge between water demand and various predictive variables, this study applies three algorithms, namely, classical Principal Component Analysis (PCA) and machine learning powerful algorithms such as Self-Organizing Maps (SOMs) and Random Forest (RF). We show that these last algorithms help corroborate the results found by PCA, while they are able to unveil hidden features for PCA, due to their ability to cope with nonlinearities. This paper presents a correlation study of three district metered areas (DMAs) from Franca, a Brazilian city, exploring weather and social variables to improve the knowledge of residential demand for water. For the three DMAs, temperature, relative humidity, and hour of the day appear to be the most important predictive variables to build an accurate regression model.

1. Introduction

The main objective of water distribution systems (WDSs) is to supply water to consumers with adequate quantity and quality. For water utilities, using accurate water demand estimation has the advantage of allowing better operation and management of their systems. Among the benefits associated with suitable water demand forecasting, leakage identification, optimal operation of pumps and valves, and the possibility of improving planning and design of network expansions must be highlighted. These engineering aspects represent a key step forward in the improvement of WDS operation efficiency, which ultimately will lead to the provision of quality water supply [1].

Water demand forecasting models can be roughly divided into long- and short-term models, whose approaches depend on the time horizon used for scheduling further predictions. Long-term water demand forecast is useful to define rehabilitation and expansion strategies and water source capacity evaluations [2]. In their turn, short-term water demand forecasting models can help to suitable define the operation and management of the water systems, with the aim of supplying water to costumers with maximum efficiency [1, 3]. As a result, working with water predictive models is central to suitably establish the set of variables involved in the model. Donkor et al. [4] present a review of several studies of water demand forecasting including various time horizons. This study is sensitive to the different nature of the used predictive variables. For long-term forecasting models it is proposed to use population density, size of the buildings, water price, and weather variables such as air temperature and relative humidity [5–7]. However, for short-term water demand approaches, several studies in the literature [8–10] include previous demand data, weather variables (such as rain

or wind speed), and calendar variables, such as weekday, hour of the day, and presence of holidays.

Autoregressive integrated moving average- (ARIMA-) based models [11] have been traditionally considered for understanding and modeling urban water demand [12]. With the improvements brought by data mining tools and machine learning techniques, a number of data analysis models have been considered more recently. For instance, several authors [13–15] have applied artificial neural network (ANN) architectures to both long- and short-term demand forecasting. The use of other machine learning tools has also increased during the last years. As an example, [9] has performed a comprehensive comparison of various predictive methods for hourly water demand forecasting, suggesting the use of support vector regression (SVR) as one of the models through which it is possible to reach better results. However, off-line predictive models are likely to develop growing bias, if models are not updated with the arrival of new data. Models can also become rapidly obsolete in the case of abrupt changes occurring in the forecasting framework. These are the models known as intervened [16] and are a consequence of unexpected changes in the scenario in which the demand is computed. For example, opening and closing valves, extreme variation of weather conditions, appearance of new leaks, and celebration of a social events, among others, may change the end-user response regarding water demand.

Despite short-term water demand forecasting models being crucial to improve water system operation and management, there is a lack of recent studies focused on correlation analyses between water demand and the usual (weather, calendar, and hydraulic) predictive variables. This is the main objective of the present paper. In-depth knowledge of those correlations will greatly improve those crucial aspects for the water supply industry.

In mathematical grounds, a general objective for regression methods is to map the input data into a convenient output space to optimally approach predictions through a set of independent variables. As this set might be formed by a large group of inputs, it is often suitable to count on ways of synthesizing the input space with minimum loss of information [17, 18]. For water demand forecasting problems, three main groups of variables integrate the input space: weather, social, and economic variables.

Despite several studies proposing various methodologies to forecast water demand, few investigations present in-depth correlation analyses able to give deeper insight into the causality principle of water demand. Among these studies, weather variables are explored as one of the main components that have influence on water demand. In this case, the impact of air temperature, relative humidity, and amount of rain may be highlighted [19–21]. Depending on the forecasting horizon, the water price and the consumer size (residential houses, businesses, industries, etc.) are used to estimate water demand as well [22]. Furthermore, large-size WDSs are usually divided into DMAs, and the correlation among these DMAs is not usually exploited in the models found in the literature. In this paper, we claim that this aspect can be used to refine water demand forecasting.

In this regard, Coomes et al. [23] reinforce pioneering studies of weather influences on water demand [20] and show the effects of weather variables. Using time-series analysis, specifically autoregressive models, can be highlighted as a classical approach for short-term water demand forecasting [19, 24, 25]. However, posterior machine learning theory developments tackle the main flaw of autoregressive models (constrained to only model linear relationships) by considering nonlinear modeling of water demand. These approaches result in substantial advances for predictive models since, regarding water demand, seasonality, dynamic-featuring, and state-dependent models cannot be built just considering linear relationships [16]. In this line, neural networks and various statistical learning methods have been widely applied to estimate the future demand with the advantage of using nonlinear regression [10, 26–29].

This work presents three methodologies to analyze water demand in close connection with various predictive variables. Firstly, the classical algorithm of correlation evaluation, Principal Component Analysis (PCA), is considered. Then, Self-Organizing Maps (SOMs) and Random Forest (RF) algorithms are also proposed to evaluate data correlations because of their ability to treat nonlinear relationships among the variables, in contrast to PCA. The three approaches are then used to deal with real world data corresponding to three district metered areas (DMAs) of Franca, a Brazilian city, in an attempt to verify potential existing correlations among water demand and some social and weather variables.

The rest of the paper is organized as follows. The following section provides the methodological aspects. The methodology is then applied to the case study, which is first suitably described; this section also includes the main results of the investigation. Finally Conclusions and References close the paper.

2. Correlation Analysis Algorithms

Correlation analysis is important to better understand various interdependences among the variables in a problem. Correlation also allows the construction of mathematical models of many phenomena. By increasing the knowledge of the set of variables that describe a given problem, the capacity to forecast also increases. This opens the possibility of performing better action policies of any related operation linked to the phenomenon under study.

The techniques and algorithms to evaluate the correlation degree among a set of predictive variables and the variable(s) they try to explain are varied. These techniques and algorithms, according to their mathematical nature, can be applied to a wide variety of situations. Among these algorithms, PCA is a classical approach using linear transformations to find data correlation. However, with the increase of data mining techniques, correlation assessment of the analyzed data is frequently not necessary, since neural networks and other machine learning methods are able to treat the data without previous assumptions as those involved in PCA. Among some other significant and powerful machine learning techniques, the SOMs and the RF algorithms can

be highlighted because of their respective ability to process large databases. Next, we concisely present these algorithms and provide the necessary elements for their application in the problem we investigate in this paper.

2.1. Principal Component Analysis (PCA). PCA has the objective to reduce the dimensionality of a dataset and to identify the superposition degree of the variables. In general terms, PCA (orthogonally) linearly transforms the input space by evidencing some correlations between variables.

Given a matrix $X = (x_{ij})$, $i = 1, \ldots, p$; $j = 1, \ldots, m$, which stores p m-dimensional samples it is possible to determine the covariance matrix $\Sigma = (\sigma_{ij})$, $i, j = 1, \ldots, m$, among the m variables whose diagonal elements σ_{ii} are the variances σ_i^2 of variables i, and the other elements are the covariances σ_{ij} between variables i and j.

As the covariance matrix is real and symmetric, using the spectral theory, it is possible to find real eigenvalues and (orthonormal) eigenvectors for this matrix. From the set of sorted eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_m$, and their associated orthonormal eigenvectors $\mathbf{e}_i = [e_{i1}, \ldots, e_{im}]^T$, new variables, named principal components, PC_i, can be written:

$$\mathrm{PC}_i = X \cdot \mathbf{e_i}. \tag{1}$$

The aim behind PCA is to create a component order explaining the variance of the dataset. This is based on the values of the corresponding eigenvalues. Once the eigenvalues are ordered, the principal component PC_i explains the variance of the dataset proportionally to the loading $\lambda_i / \sum \lambda$. When represented on a two-dimensional plane the similarity between the vector directions graphically points to the similarity between variables.

2.2. Self-Organizing Maps (SOMs). SOMs are nonsupervised learning methods based on the brain behavior when excited by external signals [30]. Motivated by this idea, SOMs are processing tools able to find out patterns in a dataset. A widely used application lies on the possibility of visualizing high dimensional datasets in reduced dimensions (usually 2D) while maintaining the topological correlations. This makes SOMs highly useful for correlation analyses among several variables.

SOM represents the input space by a mesh of points, so-called neurons. A neuron is represented by a vector with m components, known as synaptic weights. The synaptic weights change at each iteration to get the best rendering of the input space. The process to adjust the mesh is the training stage. The competitive learning process is responsible to adjust the map.

The process starts by creating a mesh of neurons, where each neuron is described by its synaptic weight vector, $\mathbf{w}_j = [w_{j1}, \ldots, w_{jm}]^T$.

Each step of the learning process is made of three different stages: competition, cooperation, and synaptic update. The competition stage is responsible for identifying the most activated region by an input $\mathbf{x}$. This region is defined as a neighborhood of the highest reactive neuron, so-called winning neuron. The winning neuron is identified by the similarity between the input and the neuron itself. This measure is usually computed by the minimal Euclidean distance. So, the winning neuron, $i(\mathbf{x})$, can be written as

$$i(\mathbf{x}) = \operatorname{argmin}\left(\left\|\mathbf{x} - \mathbf{w}_j\right\|\right). \tag{2}$$

Once the winning neuron is identified, the competition stage stops and the cooperation stage begins. In this stage, the activated region is defined and, in particular, how much these neurons are activated is also decided. The cooperation stage determines the influence of the winning neuron in the neighborhood. Finally, in the update stage the activated region has the synaptic weights modified. Following the biological inspiration, the activation decays according to the distance to the winning neuron. The activation power can be written as a monotonic decay, for example, as a Gaussian function,

$$h_{ji}(\mathbf{x}_i) = \frac{\exp\left(-d_{ji}^2\right)}{2\sigma_n^2}, \tag{3}$$

where $h_{ji}(\mathbf{x}_i)$ is the neighborhood topology function, centered in the winning neuron $i(\mathbf{x}_i)$, containing a set of j neurons excited by the winner, and σ_n is the size of the neighborhood in iteration n and is defined by an exponential decay function at each time step. The distance d_{ji} can be written as the Euclidean norm of the difference between two vectors:

$$d_{ij} = \left\|r_j - r_i\right\|^2, \tag{4}$$

where r_j is the position of an excited neuron j and r_i is the position of the winning neuron.

The winning neuron determines a region or neighborhood of influence. The closer a neuron is to the winner, the larger the change of position of this neuron is. The neighborhood defined before is shrunk through a number of iterations attempting to achieve several objectives: improving the process stability, leading the map to the final arrangement of neurons, and making the model better mimic the brain behavior. This reduction process has the disadvantage of reducing the winning neuron power. According to [31], a usual representation of the learning rate is written as

$$\sigma_n = \sigma_o \cdot \exp\left(-\frac{n}{\tau}\right), \tag{5}$$

where σ_o is the size of the initial neighborhood, n is the current iteration, and τ is a time constant, usually defined by a correlation between the maximum number of iterations, $n_{\max}$, and the initial topology size.

The synaptic weights are updated after the neighborhood activation is defined. Each weight (neural position in the topological space of the data) is then updated according to the corresponding increment, $\Delta\mathbf{w}_j$, defined as

$$\Delta\mathbf{w}_j = \eta_o \cdot \exp\left(-\frac{n}{\tau}\right) \cdot h_{j,i}(\mathbf{x}_i) \cdot \left(\mathbf{x} - \mathbf{w}_j\right), \tag{6}$$

where η_o is the initial learning rate.

Random Forest $x_1, \ldots, x_n$;
$X = X_1, \ldots, X_n$: Training Set
$Y = Y_1, \ldots, Y_m$: Response
for $b = 1 : B$ **do**
  Create a bootstrap sample (X_b, Y_b) with k instances of (X, Y)
  Run a regression tree model f^b for the training set (X_b, Y_b)
end
Return the ensemble regression tree model: $\widehat{f}(x) = \sum_{b=1}^{B} f^b(x)$

Algorithm 1: Basic Random Forest algorithm.

Finally, the new neural position is written as

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n + \Delta \mathbf{w}_j. \tag{7}$$

The learning process finishes when the mesh updates are less than a predefined threshold value or when the maximum number of iterations is reached. At this stage, it is expected that this mesh is a good two-dimensional representation of the input space, while preserving the topological relationships of data. Each variable can be represented by its neuron position allowing, by comparison of maps, qualitative inference over the correlation of variables.

2.3. Random Forest (RF). Before introducing Random Forest models, it is necessary to introduce a decision tree (DT) for regression. A DT [32] is based on a recursive partition over the range of the input space (also called instance space). It can be used for either classification or regression depending on whether the response variable is discrete or continuous, respectively. The decision tree consists of a set of nodes containing the status of the dataset partition and edges connecting them in a way in which they form a hierarchical sequence of logical rules.

A DT starts on a special node called "root" with no incoming edges. Following a previously defined sequence of logical rules, the root node iteratively breaks down the instance space into smaller instance subspaces. This is by drawing outgoing edges from the root node to other nodes and from these nodes to further ones. In tree building, all nodes, but the root, have exactly one incoming edge. In the case of a DT for regression analysis, the Sum of Squared Error (SSE) is used to define each split of the tree. The SSE computes the error between the predicted value, considering as predictor the mean per node (instance subspace), and the observed values. Comparing all these errors allows choosing the candidate node to be split at each iteration as the one having the lowest SSE. A stop criterion (tree depth or a certain SSE threshold) provides the final partition. Eventually, a subspace partition is produced and its predictors are in special nodes called leaves or terminal nodes. These are characterized by having incoming but not outgoing edges.

A Random Forest (RF) is an ensemble of tree-based models. RF algorithms can be used for classification when the base models are classification trees, or for regression when the base models are regression trees. The algorithm is based on a bootstrap aggregation (or bagging) of tree models [33].

Given the response, $Y = (Y_1, \ldots, Y_m)$, from the corresponding training set, $X = (X_1, \ldots, X_m)$, a bagging tree is constructed by selecting B samples (sampling with replacement) from (X, Y) and training a DT for each sample. Finally, in the case of regression, the bagging tree is computed by averaging all the resulting single trees. RFs use a variation of the bagging tree method by forcing each split to consider only a subset of the predictors (see Algorithm 1). This makes RFs computationally efficient compared to bagging trees for large datasets. As a general rule, for a k-dimensional problem a subset of $\sqrt{k}$ variables is selected to build single regression trees. These trees are combined in a further ensemble to improve sample tree variability. Other benefits of tree ensembles in RF are to avoid sources of bias in model outcomes and to help reducing overfitting. RFs have proven to be outstanding predictive models in regression (and classification) tasks.

3. Case Study

This work analyzes water demand records along with weather data of a Brazilian city. This case study corresponds to the WDS of Franca, a city with 318,000 inhabitants, one of the most important cities in São Paulo State (Brazil). Franca's WDS is divided into DMAs (see Figure 1). This work uses water demand data of three of its DMAs.

The three studied DMAs are typical residential areas in Brazil, encompassing family customers and small businesses. In Figure 1, Tks are storage tanks and the hatched DMAs are used in this study. Table 1 presents the number of connexions and the mean demand for each DMA.

The available demand data was measured every 20 minutes for the various DMAs. The weather data were obtained from the meteorological station of the University of Franca (Unifran) and the weather database is integrated by air temperature (°C), relative humidity (%), wind speed (m/s), wind direction (°), dew point temperature (°C), and atmospheric pressure values (hPa). All these measurements were taken in an hourly basis. To correspond to the same measurement frequency of water demand, weather data is linearly interpolated. Table 2 presents a brief statistical description of the weather database.

Social behavior is introduced in the model using calendar variables such as the hour of the day, the day of the week, the day of the month, and the month of the year.

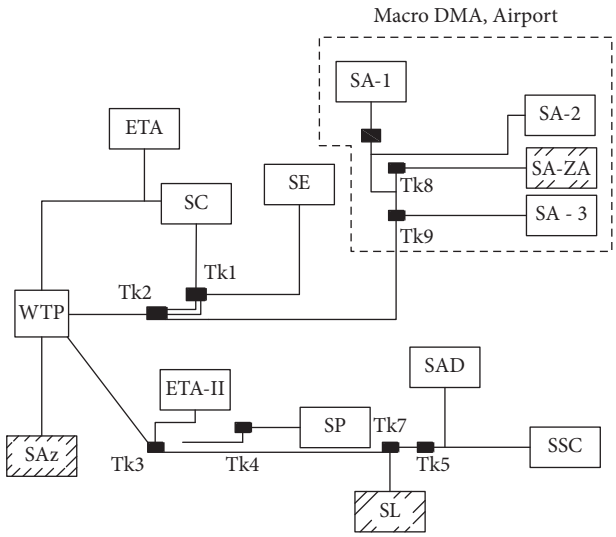

Figure 1: Spatial arrangement of the DMAs in Franca.

Table 1: DMA features.

DMA	Connections	Mean demand (l/s)
Azevedo (SAz)	8,187	96.32
Leporace (SL)	2,506	15.827
Airport (SA)	2,782	14.385

Table 2: Weather statistical features.

Variable	μ	σ	Max.	Min.
Temperature (°C)	24.396	3.399	33.500	9.900
Humidity (%)	50.845	14.109	84.500	20.400
Dew point (°C)	13.025	3.825	21.500	0.200
Wind velocity (m/s)	8.249	3.888	24.100	0
Wind direction (°)	216.873	80.670	357.000	0
Rain (mm)	0.175	1.252	50.80	0
Atm pressure (hPa)	906.323	3.750	895.70	918.00

Applying PCA to the data corresponding to each DMA studied in this work, the loading plots in Figures 2–4 are obtained. It is possible to observe the high positive correlation between water demand and hour of the day. Disregarding the DMA *SL*, temperature also has a positive correlation with water demand. Relative humidity presents a negative correlation with water demand for all DMAs, as observed by the opposite direction of the loading value.

A strong correlation among hour of day, water demand, and temperature can be observed for the three DMAs. A negative correlation between water demand and relative humidity can also be observed. This mainly happens in Azevedo's DMA. The secondary correlations, such as wind direction with month of the year or dew point and rain with atmosphere pressure, may be highlighted. These secondary correlations help validate the results of other methodologies, since they can be observed also in SOM and RF results.

The application of SOMs to identify correlations among variables can be useful, since the SOMs synthesize the topological space of inputs, by their projection onto a two-dimensional space. This projection turns easier data distribution and clustering. For the DMAs previously analyzed by the PCA, Figures 5–7 present the respective maps. The maps are based on the final distribution of neurons and the color of the maps represents the distance between neurons. Light colors represent short distances between neurons, while dark colors represent large distances. The fact that two inputs have similar distribution of neurons, that is to say, similar color distribution in the maps, helps identify qualitative correlations among data.

The distribution trends observed for temperature, hour of the day, and water demand are indicated by the reduction of the number of neurons in the positive diagonal of the maps for Airport DMA. The SOM analysis corroborates the correlations among water demand, hour of the day, and temperature. Also, the correlation among humidity and the dew point can be observed. These variables are also correlated with water demand, even if this correlation is not so clear as the one with the hour of the day.

Some secondary correlations appear in the SOM analysis clearer than in the PCA. That is the case of humidity and dew point temperature. From the physical point of view, this correlation is meaningful and points towards a good correlation analysis by the SOM interpretations. However, despite the correlation results obtained by SOMs having full physical sense, the lack of a quantitative analysis can impair the application of this kind of analysis.

To obtain deeper knowledge of the variables without previous considerations of their relationship, the RF algorithm is applied. When used to evaluate the importance of each variable, the RF algorithm runs as many times as the number of variables, removing a variable by turn and evaluating the improvement of the regression. Using this evaluation, the RF analysis ranks the priority variables. Figure 8 shows the scores of the variables for each DMA standardized in the interval [0, 1].

The hour of the day appears as the most important variable for the short-term water demand forecasting. The second most important variable for the Airport and Leporace DMAs is the temperature, while for the Azevedo DMA it is the month of year. For this DMA, the humidity is the most important weather variable. For the three DMAs, the rain is the lowest important variable, corroborating the previous results obtained by SOM and PCA.

Despite the fact that quantitative analysis of the variables can be easily performed with RFs, the secondary correlations, however, cannot be easily defined by the mathematical approach of RFs. That is to say, once the correlations have been determined by a regression process, the secondary correlations are disregarded.

SOMs and RFs appear as efficient alternatives to capture complex relationships when compared with PCA. Both provide a clear set of correlations between water demand, social variables, and weather inputs. In this study, RF corroborates the influence of the hour of the day, temperature, and relative

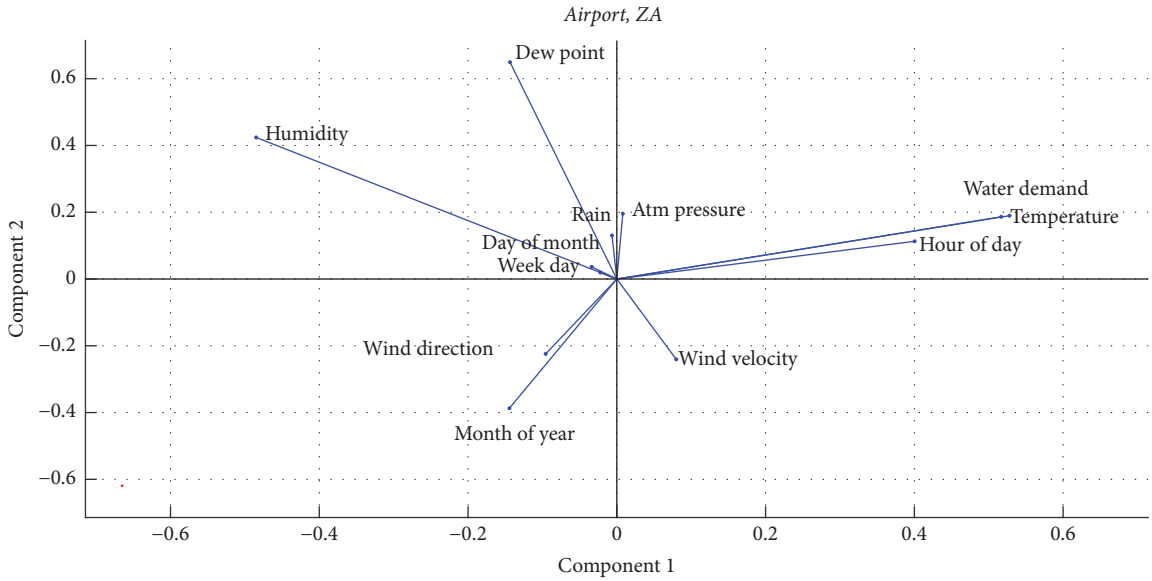

Figure 2: PCA applied to Airport DMA.

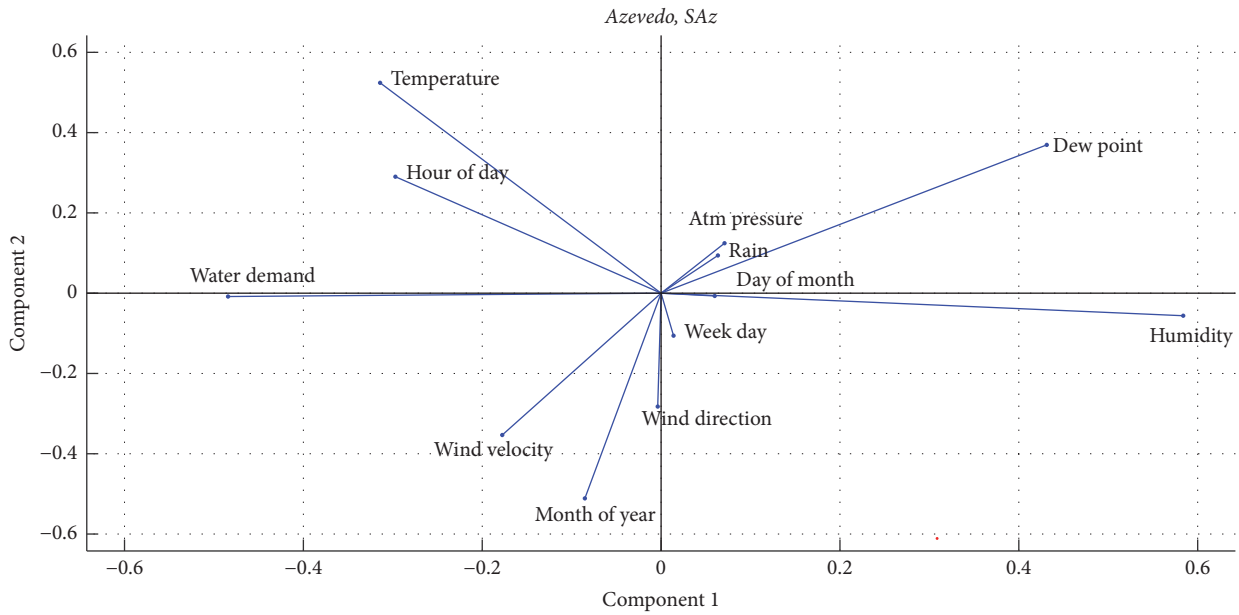

Figure 3: PCA applied to Azevedo DMA.

humidity. Other lower correlations are observed between rain and water demand for all the DMAs in the case study.

All in all, the quantitative analysis performed by SOMs can help identify important correlations, without the assumption of a linear correlation among the variables. However, quantitative analyses require personal interpretation of the maps, and this can lead to faulty correlations. In this sense, a quantitative method like the one provided by a RF is very useful. Quantitative methods can give the magnitude of the correlations and, when combined with quantitative analyses, the correlation identification process can be more powerful.

The correlation analysis of weather and social variables to improve water demand forecasting models are exploited in this work with different computational tools. Applying the methodologies to various DMAs, it is possible to evaluate the correlation properties by different spatial levels in DMAs with different size. Furthermore, the hour of the day, a predictive variable highlighted as very important for the three methodologies, corroborates the typical approach used to

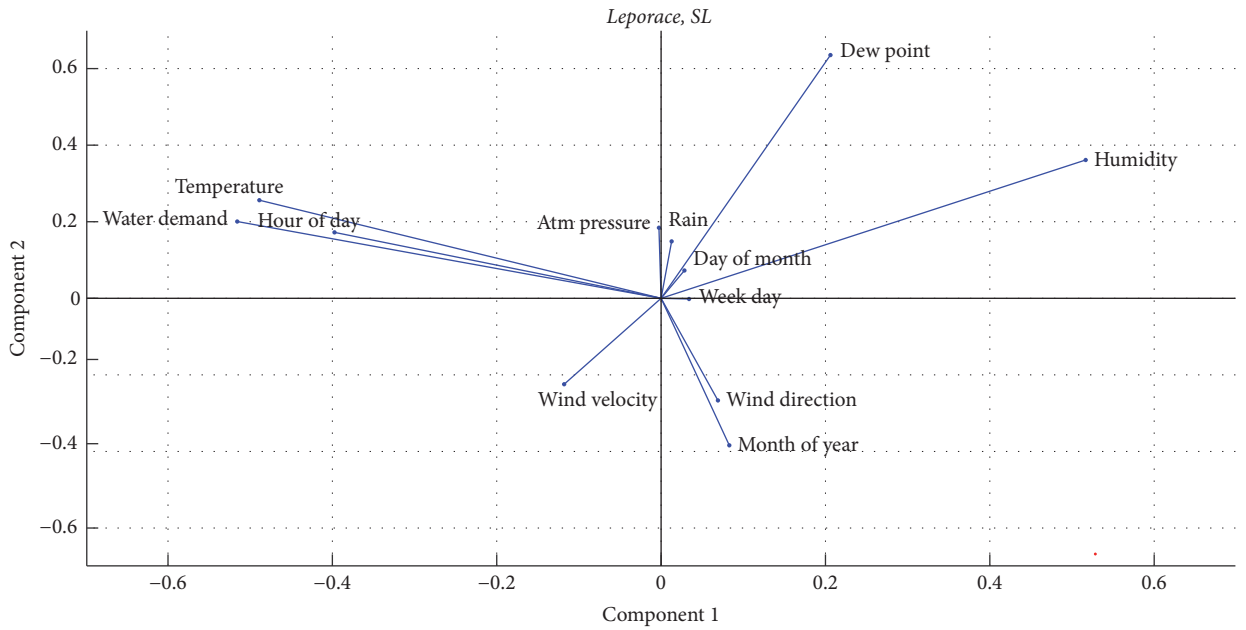

Figure 4: PCA applied to Leporace DMA.

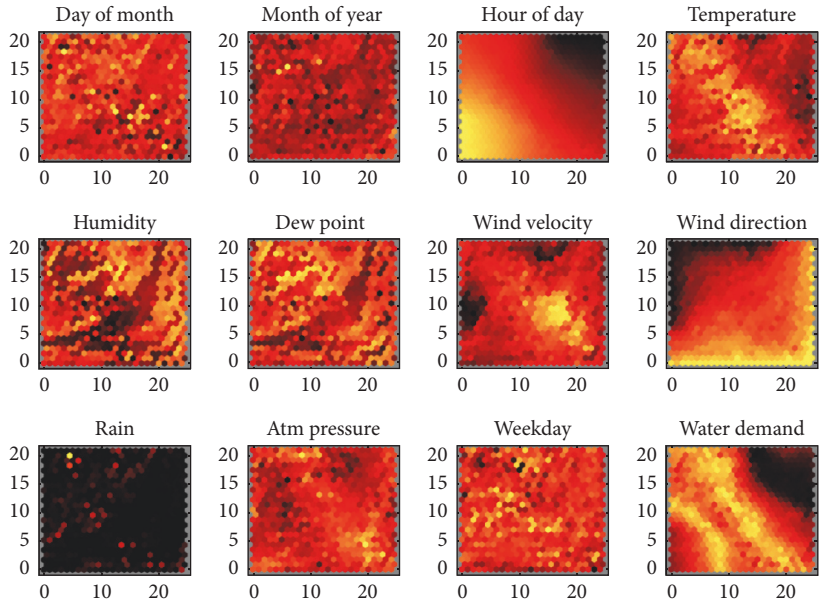

Figure 5: SOM analysis for Airport DMA.

consider temporal trends of water demand in forecasting processes.

4. Conclusions

Water demand forecasting models help decision-making processes dealing with various issues in water resources planning and management. Several studies propose using soft computing techniques to model water demand for different time horizons. However, these approaches are not exploited enough regarding the understanding of correlations between predictive variables and water demand. In this sense, assessing social and weather variables is a useful approach to improve the accuracy of regression models on water demand. This work presents three techniques to evaluate the correlation among water demand, social variables, and weather information.

A classical tool, such as PCA, can find the main correlations: temperature, hour of the day, and water demand. However, PCA is unable to find deeper correlations as they

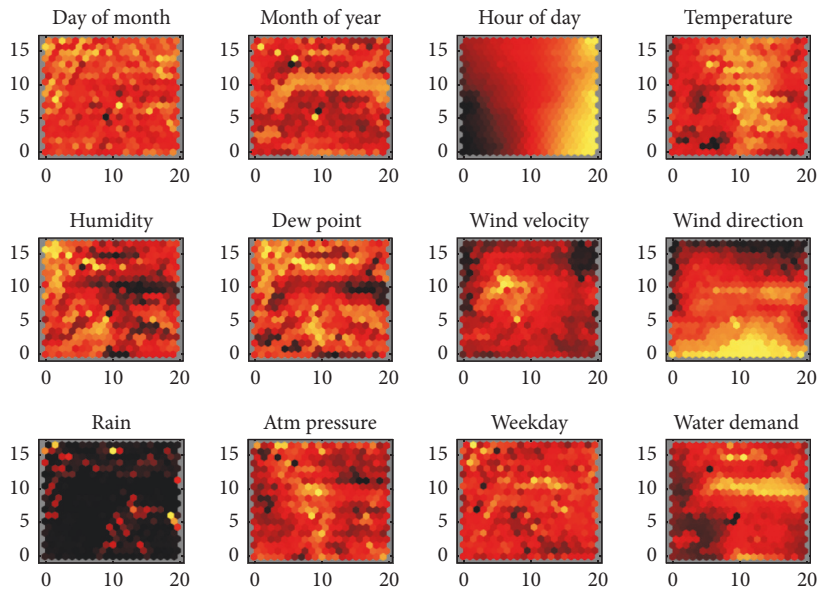

Figure 6: SOM analysis for Azevedo DMA.

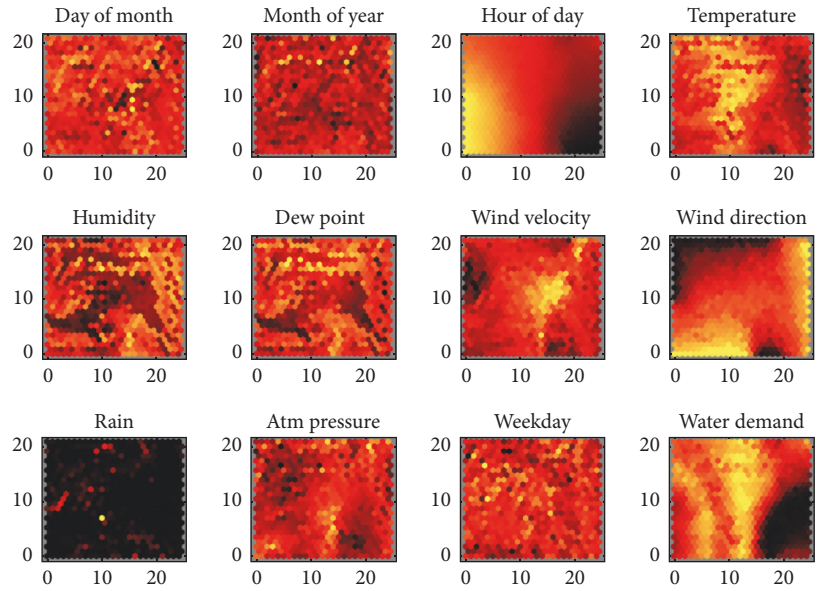

Figure 7: SOM analysis for Leporace DMA.

are of nonlinear nature. SOM analyses process the input space and turn further visual analytics easier. However, the qualitative nature of these analyses can affect the final results. RF algorithms are able to evaluate the influence of a predictive variable through comparisons of regression models. Using this approach, RF algorithms quantify the influence of each variable into the quality of the regression model.

This paper also presents a water demand analysis for various related DMAs with different consumers' features. A bullet point of this work is the space-time analysis of the correlation among water demand and the studied input variables. The obtained results allow concluding the good generalization capacity of the presented tools based on SOMs and RF algorithms.

Last but not least, it is worth mentioning that accurate water demand models help improve urban water system operation, as the degree of uncertainty in water demand is reduced. In this regard, the operation of pumps and valves then might be approached under better hydraulic conditions. Consequently, better knowledge of short-term future water demands may directly translate into several improvements on water, energy, and economic resources.

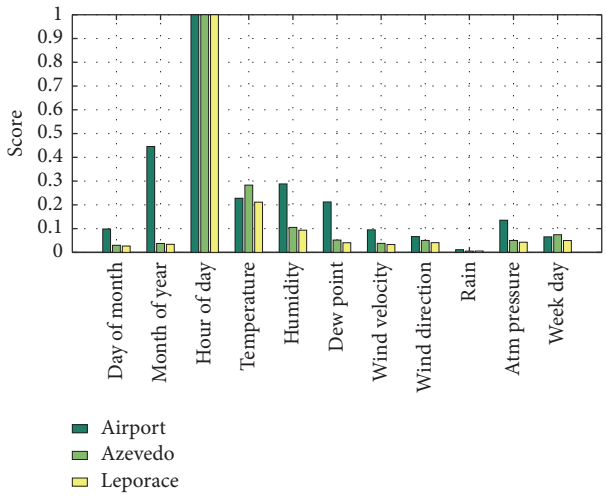

Figure 8: Scores of RF variables evaluation for the three studied DMAs.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] F. K. Odan and L. F. R. Reis, "Hybrid water demand forecasting model associating artificial neural network with fourier series," *Journal of Water Resources Planning and Management*, vol. 138, no. 3, pp. 245–256, 2012.

[2] S. Behboudian, M. Tabesh, M. Falahnezhad, and F. A. Ghavanini, "A long-term prediction of domestic water demand using preprocessing in artificial neural network," *Journal of Water Supply: Research and Technology-AQUA* , vol. 63, no. 1, pp. 31–42, 2014.

[3] R. Klempous, J. Kotowski, J. Nikodem, and J. Uasiewicz, "Optimization algorithms of operative control in water distribution systems," *Journal of Computational and Applied Mathematics*, vol. 84, no. 1, pp. 81–99, 1997.

[4] E. A. Donkor, T. A. Mazzuchi, R. Soyer, and A. Roberson, "Urban water demand forecasting: review of methods and models," *Journal of Water Resources Planning and Management*, vol. 140, no. 2, pp. 146–159, 2014.

[5] E. R. Levin, W. O. Maddaus, N. M. Sandkulla, and H. Pohl, "Forecasting wholesale demand and conservation savings," *Journal - American Water Works Association*, vol. 98, no. 2, pp. 12–111, 2006.

[6] P. Cutore, A. Campisano, Z. Kapelan, C. Modica, and D. Savic, "Probabilistic prediction of urban water consumption using the SCEM-UA algorithm," *Urban Water Journal*, vol. 5, no. 2, pp. 125–132, 2008.

[7] W. Li and Z. Huicheng, "Urban water demand forecasting based on HP filter and fuzzy neural network," *Journal of Hydroinformatics*, vol. 12, no. 2, pp. 172–184, 2010.

[8] M. Firat, M. A. Yurdusev, and M. E. Turan, "Evaluation of artificial neural network techniques for municipal water consumption modeling," *Water Resources Management*, vol. 23, no. 4, pp. 617–632, 2009.

[9] M. Herrera, L. Torgo, J. Izquierdo, and R. Pérez-García, "Predictive models for forecasting hourly urban water demand," *Journal of Hydrology*, vol. 387, no. 1-2, pp. 141–150, 2010.

[10] B. M. Brentan, J. Luvizotto, M. Herrera, J. n. Izquierdo, and R. Pérez-Garca, "Hybrid regression model for near real-time urban water demand forecasting," *Journal of Computational and Applied Mathematics*, vol. 309, pp. 532–541, 2017.

[11] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, Calif, USA, 1976.

[12] S. Alvisi, M. Franchini, and A. Marinelli, "A short-term, pattern-based model for water-demand forecasting," *Journal of Hydroinformatics*, vol. 9, no. 1, pp. 39–50, 2007.

[13] M. K. Tiwari and J. F. Adamowski, "Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach," *Journal of Water Resources Planning and Management*, vol. 141, no. 2, Article ID 04014053, 2015.

[14] C. Bennett, R. A. Stewart, and C. D. Beal, "ANN-based residential water end-use demand forecasting model," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1014–1023, 2013.

[15] S. L. Zhou, T. A. McMahon, A. Walton, and J. Lewis, "Forecasting operational demand for an urban water supply zone," *Journal of Hydrology*, vol. 259, no. 1-4, pp. 189–202, 2002.

[16] M. Herrera, J. C. Garca-Daz, J. Izquierdo, and R. Pérez-Garca, "Municipal water demand forecasting: tools for intervention time series," *Stochastic Analysis and Applications*, vol. 29, no. 6, pp. 998–1007, 2011.

[17] C. M. Andersen and R. Bro, "Variable selection in regression-a tutorial," *Journal of Chemometrics*, vol. 24, no. 11-12, pp. 728–737, 2010.

[18] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[19] D. R. Maidment, S. Miaou, and M. M. Crawford, "Transfer Function Models of Daily Urban Water Use," *Water Resources Research*, vol. 21, no. 4, pp. 425–432, 1985.

[20] D. R. Maidment and S. Miaou, "Daily Water Use in Nine Cities," *Water Resources Research*, vol. 22, no. 6, pp. 845–851, 1986.

[21] S. L. Zhou, T. A. McMahon, A. Walton, and J. Lewis, "Forecasting daily urban water demand: A case study of Melbourne," *Journal of Hydrology*, vol. 236, no. 3-4, pp. 153–164, 2000.

[22] F. K. Odan, L. F. Ribeiro Reis, and Z. Kapelan, "Real-time multiobjective optimization of operation of water supply systems," *Journal of Water Resources Planning and Management*, vol. 141, no. 9, Article ID 04015011, 2015.

[23] P. A. Coomes, B. D. Kornstein, T. D. Rockaway, and J. A. Rivard, "North America Residential Water Usage Trends," *Proceedings of the Water Environment Federation*, vol. 2010, no. 9, pp. 6488–6500, 2010.

[24] A. Jain, A. K. Varshney, and U. C. Joshi, "Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks," *Water Resources Management*, vol. 15, no. 5, pp. 299–321, 2001.

[25] H. A. Mombeni, S. Rezaei, S. Nadarajah, and M. Emami, "Estimation of Water Demand in Iran Based on SARIMA Models," *Environmental Modeling & Assessment*, vol. 18, no. 5, pp. 559–565, 2013.

[26] J. Bougadis, K. Adamowski, and R. Diduch, "Short-term municipal water demand forecasting," *Hydrological Processes*, vol. 19, no. 1, pp. 137–148, 2005.

[27] J. Adamowski and C. Karapataki, "Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: Evaluation of different ANN learning algorithms," *Journal of Hydrologic Engineering*, vol. 15, no. 10, Article ID 005010QHE, pp. 729–743, 2010.

[28] M. Herrera, S. Canu, A. Karatzoglou, R. Pérez-García, and J. Izquierdo, "An approach to water supply clusters by semi-supervised learning," in *International Congress on Environmental Modelling and Software*, A. David, Y. Wanhong, A. Voinov, A. Rizzoli, and T. Filatova, Eds., pp. 1925–1932, 2010.

[29] Y.-S. Xu, Y.-D. Mei, and T. Yong, "Combined forecasting model of urban water demand under changing environment," in *Proceedings of the 2011 International Conference on Electric Technology and Civil Engineering, ICETCE 2011*, pp. 1103–1107, China, April 2011.

[30] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[31] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2004.

[32] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[33] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.