

# Combined Holt-Winters and GA trained ANN Approach for Sensor Validation and Reconstruction: Application to Water Demand Flowmeters

Hector Rodriguez<sup>1</sup>, Vicenç Puig<sup>2</sup>, Juan J. Flores<sup>3</sup> and Rodrigo Lopez<sup>4</sup>

**Abstract**—This paper proposes a Double Seasonal Holt-Winters (DSHW) forecasting model with an auxiliary Artificial Neural Network (ANN) trained with a Genetic Algorithm (GA) to model the DSHW residuals. ANN complements and improves the DSHW prediction. The proposed model also includes an on-line validation and reconstruction mechanism useful to detect and correct missing and erroneous data. This mechanism also impacts improving the DSHW prediction accuracy and precision. The proposed model and validation mechanism are applied to predict the time series generated by two monitored flowmeters of two sectors of Barcelona’s drinking water network (DWN). The accuracy and precision improvement of the proposed method with respect to standard DSHW and ARIMA approaches is provided.

## I. INTRODUCTION

Drinking water demand analysis is a relevant activity that impacts the human population wellness and the sustainable and optimal exploitation of this resource. Nowadays the water resource declared by Alma Ata as the most important element for life [1]. However, it faces a serious shortage problem due to the global warming produced by the climate change. These facts lead to make active and strong efforts oriented to optimize this resource. One research direction is the optimization of the management of drinking water delivery in urban areas.

Every year novel studies arise to model the human water demand behavior to forecast the demand accurately with certain safety (confidence). The interest of the scientific community in the research of specialised algorithms to forecast in the short-term the water demand for automatic control has not start growing (e.g., important contributions related to this research topic are found in [2], [3], [4], [5] [6]). These works evidence the importance of the operational and water resources optimization. The proposed forecast models are used to predict the short-term water consumption, and then, the forecast is used as reference by mechanisms of

operational control to dispatch only the necessary resource avoiding the unnecessary emptying of water stores.

The works of Quevedo et al, [7], [8] demonstrate that double seasonal Holt-Winters (DSHW) behaves similar to seasonal ARIMA with patterns (SARIMA pattern). Although SARIMA pattern was slightly better than DSHW, DSHW does not require further analysis of water demand patterns to perform reasonably good.

The most noticeable drawbacks of DSHW is the impossibility of modelling certain nonlinear dynamics of time series and the sensitivity to outliers and erroneous data from sensors producing inaccurate forecasts. This problem is also present in ARIMA models for complex behavior of time series such as the wind speed. In [9], the ARIMA residuals modelling for speed wind forecasting is addressed using also ANN.

This paper contributes with a DSHW based application framework that provides an improvement of the DSHW predictor using an artificial neural network (ANN) (called DSHW+ANN), and a validation - reconstruction algorithm based on confidence intervals. To improve the DSHW forecast performance, the auxiliary ANN is implemented to model the unmodeled nonlinearities by DSHW found in drinking water demand time. The nonlinearities are found by analysing the DSHW residuals. The confidence intervals are computed using the residuals generated by the improved DSHW.

There are time series difficult to model due to the high nonlinearity, as consequence, the residuals generated by any prediction model still preserve dynamic structure that is detected measuring their correlation. In this work, ANN is chosen to model the DSHW residual dynamics due to its capability to model nonlinear time series without trend [10]. The zero trend and a probability distribution with zero mean and certain variance [11], are characteristics present in the residuals. In this work, we assume the residuals present a Gaussian distribution with zero mean in order to estimate the valid range of the confidence prediction.

This article is organized as follows: In Section II, the Double Seasonal Holt-Winters method, the basics of ANN, the GA training method for the residuals and the computation of confidence intervals are described. Section III introduces the ANN+DSHW approach for water demand time series. Section IV shows the experiments and results of the proposed method compared with ARIMA and DSHW. Finally, Section V draws final conclusions and suggested future work regarding the proposed framework.

<sup>1</sup> Hector Rodriguez works at Instituto Tecnológico de Culiacán, División de Estudios de Posgrado e Investigación, Juan de Dios Bátiz 310 pte. Col Guadalupe C.P. 80220, Culiacán, Mexico. hector120713@gmail.com

<sup>2</sup> Vicenç Puig works at the Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Carrer LLorens Artigas, 4-6, Barcelona, Catalonia, Spain, 08028. vicenc.puig@upc.edu

<sup>3</sup> Juan J. Flores works at División de Estudios de Postgrado, Facultad de Ingeniería Eléctrica, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico juanf@umich.mx

<sup>4</sup> Rodrigo Lpez works at Centro de Cómputo y Procesos de Información Universitaria, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico. rdg1lpz@gmail.com

## II. RELATED METHODS

### A. Holt-Winters

Holt-Winters is an exponential based method developed by C.C. Holt (1957) and P. Winters in (1960) [12].

Exponential smoothing methods are based on the weighted averages of past observations, with the weights decaying exponentially as the observations get older.

The double seasonal Holt-Winters (DSHW) extension with two seasonal components in time series is composed by four equations: level  $\ell_t$ , trend  $b_t$ , and two seasonal equations ( $s_t^1, s_t^2$ ) [13]. DSHW requires smoothing  $\alpha_1, \alpha_2, \alpha_3, \omega$ , and two seasonal parameter  $\tau_1, \tau_2$  with the information about time series periodicity. DSHW mathematical expressions are as follows

$$\ell_t = \alpha_1(y_t / (s_{t-\tau_1}^1 s_{t-\tau_2}^2)) + (1 - \alpha_1)(\ell_{t-1} + b_{t-1}) \quad (1)$$

$$b_t = \alpha_2(\ell_t - \ell_{t-1}) + (1 - \alpha_2)b_{t-1} \quad (2)$$

$$s_t^1 = \alpha_3(y_t / (\ell_t s_{t-\tau_2}^2)) + (1 - \alpha_3)s_{t-\tau_1}^1 \quad (3)$$

$$s_t^2 = \omega(y_t / (\ell_t s_t^1)) + (1 - \omega)s_{t-\tau_2}^2 \quad (4)$$

$$\hat{y}_{t+h} = (\ell_t + hb_t)s_{t-\tau_1+h}^1 s_{t-\tau_2+h}^2 \quad (5)$$

where  $\hat{y}_{t+h}$  is the  $h$ -th step-ahead forecast made from forecast origin  $t$ . The parameters are optimized by minimizing the sum of squared errors of one step-ahead residuals.

### B. Artificial Neural Networks

Artificial neural networks (ANN) are models inspired on the biological neural networks. ANN are widely used to solve engineering problems as function modelling, regression and system identification for control [14] [15]. In this paper, we are interested in modelling a nonlinear time series produced by the residuals of the one-step ahead DSHW.

The basic ANN architecture has an input layer with input vector size, a intermediate or hidden layer that weights the result of applying a basis or activation function to each value of the input vector and the output vector that is the final weighted sum of the values from the hidden layer. The most used basis functions in ANN are Sigmoidal and Radial Basis Functions [16].

The mathematical expression of a three layer ANN that forecast  $\hat{y}_{t+i}$  is defined as:

$$\hat{y}_{t+i} = f_1\left(\sum_i^n w_i x_i\right) \quad (6)$$

$$x_i = f_2\left(\sum_j^h w_{ij} y_{t-j}\right)$$

where  $h$  is number of the neurons in the hidden layer, the output layer corresponds to the forecast ( $\hat{y}_{t+i}$ ), and  $f_1$  and  $f_2$  are sigmoid functions used as the activation functions.  $w$  are the coefficients (or weight) connections.

,  $f_k(x)$  is the  $k$ -th output layer of the neural network and  $\phi$  is the sigmoidal activation function. The weights  $w$  in (6) are optimized to minimise the prediction errors.

The sigmoidal activation function is defined as follows

$$f(v_i) = \tanh(v_i) \quad (7)$$

where  $\tanh$  is the hyperbolic tangent.

### C. Training Artificial Neural Networks with Genetic Algorithms (GA)

GA is an optimization technique inspired on Darwin's principle of evolution. GA mimics a simplistic version of the process of biological evolution, which consists of creating a population of individuals, where each individual represents a prospective solution of the problem being solved [17]. GA modifies this population using genetic operators: selection, mutation, recombination, etc. This stage, called a generation, repeats until a termination criterion is met. At the end of the process, the best individual (i.e., the fittest one) found during the evolution is returned as the solution of the problem. Determining the best ANN architecture for forecasting is an optimization problem that can be coded to be optimized by GA. In this paper, we use GA as used in the work of Rodriguez et al [18] to find the optimal ANN architecture and its weights. GA defines the architecture of the ANN and the weights of the neurons connections for the residuals modelling problem. It has been proved that using GA to design the net and perform the training process has a better improvement than traditional methods [19], [20], [21]

### D. Confidence Interval

The confidence interval (CI) is an statistical estimate that defines the boundaries whose enclose the most frequent observations of a probability distribution. In order to validate the data, the boundaries act as a threshold defining the valid data. CI is useful to measure the model uncertainty. In this paper, we assume the CI belongs to a normal distribution since the data sample are prediction residuals with zero mean and should shape a Gaussian distribution if the prediction model is correct. The definition of the CI is given by

$$y_t \in [\hat{y}_t, \bar{\hat{y}}_t] \quad (8)$$

where  $\hat{y}_t = \hat{y}_t - \alpha \frac{\sigma}{\sqrt{N}}$  and  $\bar{\hat{y}}_t = \hat{y}_t + \alpha \frac{\sigma}{\sqrt{N}}$ .  $\alpha$  is the confidence degree,  $\sigma$  is the standard deviation, and  $N$  is length of the population.

## III. IMPROVED HOLT-WINTERS FORECASTING WITH ANN RESIDUAL MODELING (DSHW+ANN)

To tune the Holt-Winters model and train the ANN residuals, the required information has a format of time series described by

$$\mathbf{Y} = \{y_1, \dots, y_t, \dots, y_n\} \quad (9)$$

where  $y_1, y_t, y_n$  are the first, current and last elements of the time series. The residuals are stored in the vector  $\mathbf{r}$  defined as follows

$$\mathbf{r} = \{r_1, \dots, r_n\} \quad (10)$$

The residuals in (10) are obtained by using the one-step ahead fitted Holt-Winters model,

$$r_{t+1} = y_{t+1} - \hat{y}'_{t+1} \quad (11)$$

where the prediction  $\hat{y}'_{t+1}$  is given by

$$\hat{y}'_{t+1} = f_{dshw}(\{y_{t-m}, \dots, y_t\}) \quad (12)$$

The residuals  $\mathbf{r}$  are modeled training an ANN with evolutionary computing in order to forecast the residuals with the regressor  $f_{ann}$  defined as

$$\hat{r}_{t+1} = f_{ann}(\{r_{t-m}, \dots, r_t\}) \quad (13)$$

Once we have DSHW and ANN forecasters, they are used to make the one-step ahead prediction using

$$\hat{y}_{t+1} = \hat{y}'_{t+1} + \hat{r}_{t+1} \quad (14)$$

The block diagram describing the mechanism is shown in Figure 1 where the blocks Holt-Winters and ANN contains the forecasters described in Equation 12 and 13.

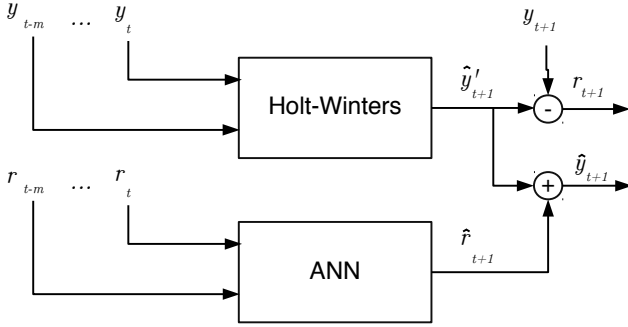


Fig. 1. DSHW+ANN Forecast Model

Once the DSHW+ANN model is available, we proceed to compute the confidence intervals by analyzing the set of prediction errors defined as

$$\mathbf{e} = \{e_1, \dots, e_t, \dots, e_n\}. \quad (15)$$

where the final residuals are given by

$$e_t = \hat{y}_t - y_t \quad (16)$$

In order to validate the incoming data, we consider the two kind of erroneous data: missing data, and outliers. These data are produced by malfunctioning of the sensors, therefore they should be validated and then reconstructed before being recorded and used by the forecasters. The validation process is performed as follows

$$validate(y_t) = \begin{cases} y_t & \text{if } y_t \in [\hat{y}_t, \hat{y}_t] \\ \hat{y}_t & \text{if } y_t = \emptyset \vee y_t \notin [\hat{y}_t, \hat{y}_t] \end{cases} \quad (17)$$

where  $y_t$  is valid only if it exists and is inside of the bounds  $[\hat{y}_t, \hat{y}_t]$ . Otherwise, the measurement is replaced by the model estimation. Figure 2 summarizes in graphical manner using a block diagram the validation/correction process. The block diagram has three modules:

- the Operational Database that stores the valid time series and residuals generated by means of Eqs. 9 and 11,
- the Forecast Model module with the DSHW+ANN (Figure 1) and
- the Validation module that implements the validation function of Equation 2.

The Operational Database block provides DSHW residuals and validated time series information to the Forecast Model. The Forecast Model also feeds the Validation module deciding whether the estimated forecast  $\hat{y}_{t+1}$  or the raw measurement  $y_{t+1}$  should be stored in the Operational Database.

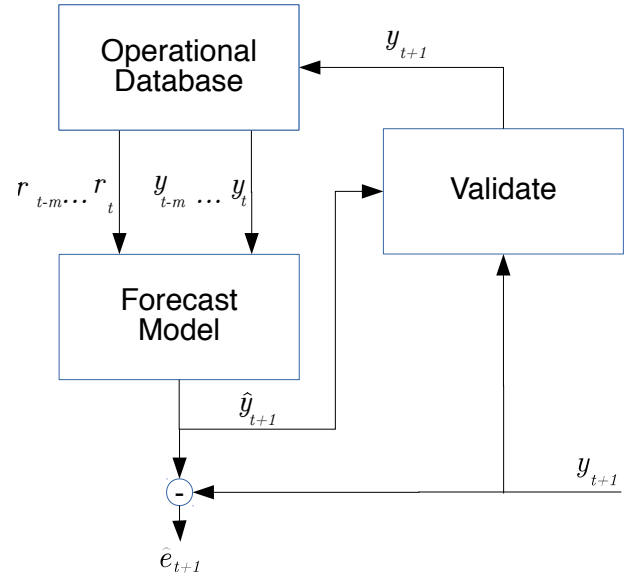


Fig. 2. Validation process

#### IV. RESULTS

The proposed approach was tested with flow meter time series from the database of Barcelona water transport network provided by the company Aguas de Barcelona (AGBAR) [22]. This company supplies water to Barcelona city, and also to the metropolitan area. It covers 23 municipalities included in an area of 424 km<sup>2</sup>, with 4,645 km of pipes in order to satisfy the water demand of about 3 million people. The complete transport network contains: 63 storage tanks, 3 surface sources and 7 underground sources, 79 pumps, 50 valves, 18 nodes, and 88 demand sectors. The network is controlled through a SCADA system with sampling periods of 1 hour. In this work, we considered the data from water demand flow meter sensors that present typical water consumption behavior of urban areas. The experiments are performed with the time series produced by the flowmeters identified by labels  $p10017$  and  $p10013$  in the Barcelona

water distribution network. The time series length considered to perform the experiments is one month in hourly scale (720 data), splitting the data in a set with 70% of the data (504 data points) and another set with 30% for the validation (216 data points).

The experiments report the effectiveness of the proposed DSHW+ANN with the validation model by means of accuracy and precision compared with DSHW and ARIMA.

To use the validation model and reconstruct the false data, first we need to model the behavior of the water demand time series as best as possible. Then, we define a valid threshold to detect incorrect data using interval estimate provided by the Confidence Interval (8).

The experiments demonstrate the improvement of the DSHW+ANN using the validation model with respect to DSHW by reporting the forecast accuracy and precision.

The same training set is used to tune the DSHW and train the ANN. Once the DSHW is trained, the residuals are used to train the ANN using GA as described in Section II. DSHW period parameters are  $\tau_1 = 24$  and  $\tau_2 = \tau_1 \times 7$  (daily and weekly seasonal parameters), and the model is obtained by using the function *dshw* of the forecast package found in the *R* software [23]. ANN is trained by using GA with the evolutionary parameters reported in Table I.

The prediction of the DSHW model is depicted in Figure 3, which shows the water demand data using the solid black line and the one-step ahead prediction performed by the DSHW model using the dotted gray line. Figure 4 plots the DSHW residuals obtained as difference of the previous signals presented in Figure 3.

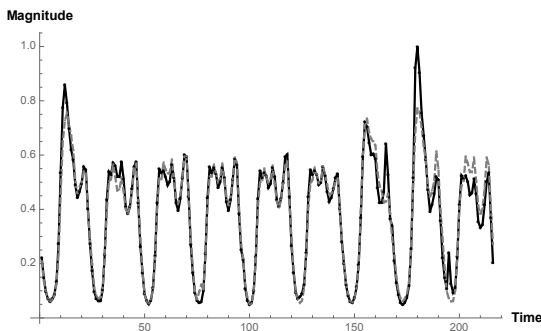


Fig. 3. DSHW Model. The solid black line corresponds to the real data, and the dotted gray line to the DSHW model.

The obtained residuals, are modeled with an ANN. The design of the architecture of the ANN, and the training process was developed by GA. The parameter of the GA used to optimize the ANN function (6) are shown in Table I.

The result of modelling the residuals using ANN is depicted in the Figure 5, where the solid black line is the residual time series and the dotted grey line correspond to the ANN prediction. It is important to notice that in spite of having noisy residuals with unclear dynamic structure, ANN was capable to model them quite accurately. The experiments

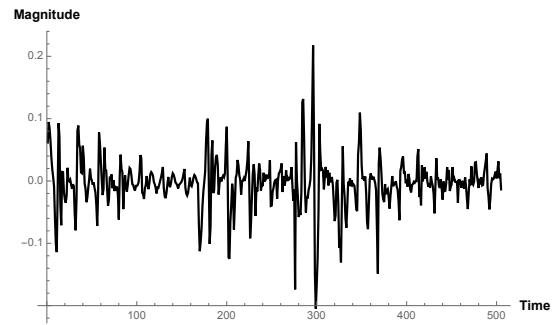


Fig. 4. Holt-Winters Residuals

TABLE I  
PARAMETERS OF THE GA

Concept	Value
Length of data	720
Length of the validation set	216
range of inputs of the ANN	10 to 70
range of neurons in the hidden layer	20 to 70
Individual mutation probability	55%
Gen mutation probability	0.05%
Crossover probability	70%

for modeling the residuals were performed under a *Python* platform [24], using the library of Evolutionary Algorithms [25].

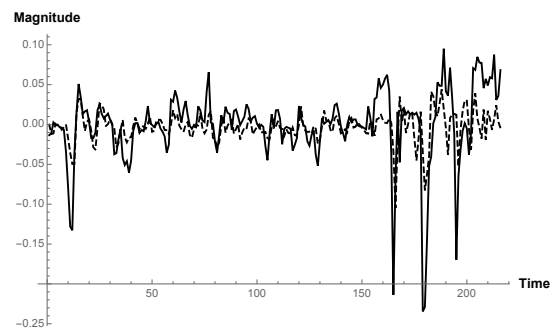


Fig. 5. ANN Model Residuals

The ANN residuals model is used together with the DSHW model to improve the performance of the single DSHW. The resultant performance of the DSHW+ANN (14) is presented in Figure 6.

Figure 6 presents four signals: the upper and lower dotted lines define the uncertainty threshold, the solid black line corresponds to the real time series, and the dotted gray line corresponds to the obtained model using DSHW+ANN. To define the uncertainty threshold, we use the statistical measure of the confidence intervals introduced in (8).

The histogram of the errors (defined as in (16)) is depicted in the Figure 7. From Figure 7, it can be observed two vertical lines that define the uncertainty threshold with a 95 % of confidence. All data outside this boundary is considered false and then replaced with the forecast estimate of the DSHW+ANN model.

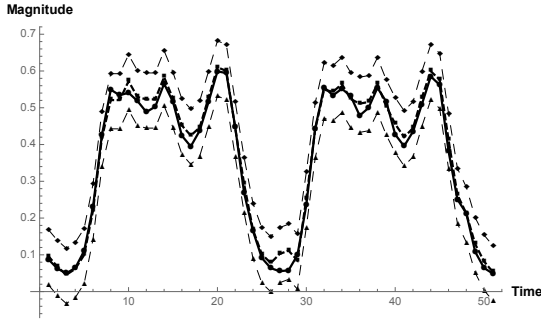


Fig. 6. HW+ANN Model

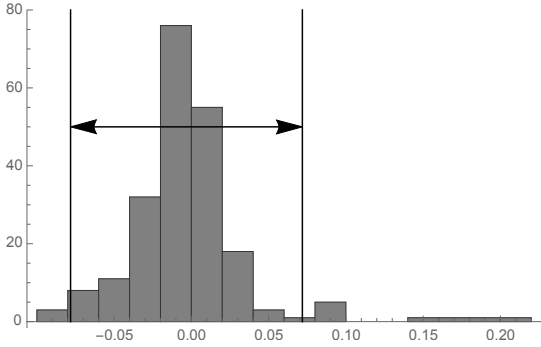


Fig. 7. Residual histogram

Once the DSHW+ANN model is available, we proceed to quantify the improvements of the model. Reducing uncertainty, we will allow improving the quality of the data validation approach proposed in Section III. Table II and III show the improvements in terms of accuracy and precision. The first column of Table II and III, shows the different forecasting models for comparison. The second column expresses the certainty in area. The third column expresses the inferior and superior limit of the confidence intervals, and finally, the fourth and fifth column report the accuracy in terms of MSE and MAPE.

Considering the two experiments, we obtain a general improvement of more than 10% in terms of the uncertainty

TABLE II  
RESULTS TIME SERIES OF FLOWMETER P10017

Model/Meas.	Precision		Accuracy	
	AREA	CI	MSE	MAPE
DSHW	37.25	[-0.0877, 0.0878]	0.0019	7.0%
DSHW+ANN	32.25	[-0.0781, 0.0718]	0.0014	6.3%
ARIMA	108.575	[-0.2520, 0.2539]	0.0165	22.9%

TABLE III  
RESULTS TIME SERIES OF FLOWMETER P10013

Model/Meas.	Precision		Accuracy	
	AREA	CI	MSE	MAPE
DSHW	85.27	[-0.222, 0.2172]	0.01251	19.18%
DSHW+ANN	82.06	[-0.1966, 0.2236]	0.0116	20.2%
ARIMA	287.4	[-0.7123, 0.7543]	0.1397	83.89%

threshold area and more than 11% in terms of accuracy.

The results demonstrates that the water demand time series have a complex dynamics not fitted completely by state-of-the-art algorithms such as DSHW. The water demand time series requires complex forecasting modeling for having a higher performance required in the industry.

## V. CONCLUSIONS AND FUTURE WORK

In this work, we proposed a GA trained ANN forecasting model that complements and improves the DSHW one-step ahead forecasting prediction accuracy and precision. We also proposed to include a validation mechanism based on confidence intervals that allows detecting missing and erroneous data. The data validation is useful to robustify the DSHW predictions since it is sensitive to inconsistent data performing poorly.

Although there exist generalised Holt-Winters models that capture a third or more seasonalities to improve the forecast accuracy such as in [26], they are focused on improving the global modeling for larger time series that put on evidence the existence of a third seasonality (e.g., Besides the intraday and intraweek cycles, it is also observed an intrayear seasonal cycle in hourly time series), rather than improving and refining the local modeling of the dynamics as we do in this paper. We believe that the proposed approach can be implemented straightforward to complement these kind of models to improve the prediction performance in the short-term.

The computation of the confidence intervals have further practical implementations. They are useful for the safe operation in control (e.g., distribution of the water considering the worst case), and also for improving the precision of the billing. Confidence intervals are defined by the upper and lower bound, which delimit the certainty of the information (certainty threshold) provided by the flow meters. Among the uses of the certainty threshold, for example, the upper bound could be used to determine the potential demand for the next period, and the lower bound can be used to define the certainty of the information provided by the flowmeter.

As future work we propose to prove the proposed approach for other methods such as the triple seasonal methods and include a comparison with the Gaussian Process DSHW proposed in [6], and also to explore more reliable methods for obtaining more accurate confidence intervals. Precise confidence intervals improve the forecast precision providing safety to the operation of the drinking water demand minimising the sacrifice of resources to keep a desired quality of service. The authors suggest the use of different probability distributions for specific water demand hour since the uncertainty might be different at different moments; e.g., the uncertainty of the water consumption is less during the night while the people sleeps than during the the middle of the day while the people are working and the activity is more unpredictable. Possibly the probability distribution functions might obey to a Gaussian or another kind of complex probability distribution shapes that are present at different moments.

## ACKNOWLEDGMENT

Hector R. Rangel thanks the CSIC-UPC (Institut de Robòtica i Informàtica Industrial) for all the resources and accommodations that made his visit at the University possible.

This work has been partially funded by the Spanish Government (MINECO) through the project CICYT ECOCIS (ref. DPI2013-48243-C2-1-R), by MINECO and FEDER through the project CICYT HARCRICS (ref. DPI2014-58104-R).

## REFERENCES

- [1] W. H. Organization, Guidelines for drinking-water quality: recommendations, Vol. 1, World Health Organization, 1993.
- [2] E. a. Donkor, T. a. Mazzuchi, R. Soyer, J. A. Roberson, Urban Water Demand Forecasting : Review of Methods and Models, *Journal of Water Resources Planning and Management* 140 (February) (2014) 146–159. doi:10.1061/(ASCE)WR.1943-5452.
- [3] P. Cutore, a. Campisano, Z. Kapelan, C. Modica, D. Savic, Probabilistic prediction of urban water consumption using the SCEM-UA algorithm, *Urban Water Journal* 5 (2) (2008) 125–132. doi:10.1080/15730620701754434.
- [4] S. Zhou, T. McMahon, A. Walton, J. Lewis, Forecasting operational demand for an urban water supply zone, *Journal of Hydrology* 259 (14) (2002) 189 – 202. doi:http://dx.doi.org/10.1016/S0022-1694(01)00582-0.  
URL <http://goo.gl/Fcrc34>
- [5] N. Kourentzes, Intermittent demand forecasts with neural networks, *International Journal of Production Economics* 143 (1) (2013) 198–206. doi:10.1016/j.ijpe.2013.01.009.
- [6] Y. Wang, C. Ocampo-Martinez, V. Puig, J. Quevedo, Gaussian-process-based demand forecasting for predictive control of drinking water networks, in: *Proceedings of the 9th International Conference on Critical Information Infrastructures Security (CRITIS)*, 2014.
- [7] J. Quevedo, V. Puig, C. G., J. Blanch, J. and Aguilar, D. Saporta, G. Benito, M. Hedro, A. Molina, Validation and reconstruction of flow meter data in the barcelona water distribution network, *Control Engineering Practice* 18 (6) (2010) 640 – 651. doi:http://dx.doi.org/10.1016/j.conengprac.2010.03.003.  
URL <http://goo.gl/2bMbGE>
- [8] J. Quevedo, J. Saludes, V. Puig, J. Blanch, Short-term demand forecasting for real-time operational control of the barcelona water transport network, in: *Control and Automation (MED), 2014 22nd Mediterranean Conference of*, 2014, pp. 990–995. doi:10.1109/MED.2014.6961503.
- [9] E. Cadena, W. Rivera, Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model, *Renewable Energy* 35 (12) (2010) 2732–2738. doi:10.1016/j.renene.2010.04.022.
- [10] G. Zhang, B. E. Patuwo, M. Y. Hu, Forecasting with artificial neural networks:: The state of the art, *International Journal of Forecasting* 14 (1) (1998) 35 – 62. doi:http://dx.doi.org/10.1016/S0169-2070(97)00044-7.  
URL <http://goo.gl/fy100z>
- [11] S. Weisberg, *Applied Linear Regression*, Vol. 528, 2005. doi:10.2307/1269895.  
URL <http://goo.gl/puZxFe>
- [12] P. Goodwin, The Holt-Winters Approach to Exponential Smoothing : 50 Years Old and Going Strong, *Foresight* (2010) 30–33.
- [13] J. W. Taylor, L. M. de Menezes, P. E. McSharry, A comparison of univariate methods for forecasting electricity demand up to a day ahead, *International Journal of Forecasting* 22 (1) (2006) 1–16. doi:10.1016/j.ijforecast.2005.06.006.
- [14] J. Hoskins, D. Himmelblau, Process control via artificial neural networks and reinforcement learning, *Computers and Chemical Engineering* 16 (4) (1992) 241 – 251, neural network applications in chemical engineering. doi:http://dx.doi.org/10.1016/0098-1354(92)80045-B.  
URL <http://goo.gl/Qf5nlc>
- [15] M. R. Smith, J. T. A. (Eds.), *Multiple Model Approaches to Modelling and Control*, Taylor and Francis, London, 1997.
- [16] M. Orr, *Introduction to radial basis function networks*, University of Edinburg (1996) 1–7.  
URL <http://goo.gl/b4RjIA>
- [17] D. E. Goldberg, J. H. Holland, Genetic algorithms and machine learning, *Machine learning* 3 (2) (1988) 95–99.
- [18] H. Rodriguez, V. Puig, J. J. Flores, R. Lopez, Flow meter data validation and reconstruction using neural networks: Application to the barcelona water network, in: *European Control Conference 2016*, 2016.
- [19] J. J. Flores, H. Rodriguez, M. Graff, Reducing the search space in evolutive design of arima and ann models for time series prediction, in: *Advances in Soft Computing*, Springer, 2010, pp. 325–336.
- [20] J. J. Flores, R. Loaeza, H. Rodríguez, E. Cadenas, Wind speed forecasting using a hybrid neural-evolutive approach, in: *MICAI 2009: Advances in Artificial Intelligence*, Springer, 2009, pp. 600–609.
- [21] J. J. Flores, M. Graff, H. Rodriguez, Evolutive design of arma and ann models for time series forecasting, *Renewable Energy* 44 (2012) 225–230.
- [22] S. Sociedad General de Aguas de Barcelona, Annual corporate governance report (2008).
- [23] R. Hyndman, Package forecast, <https://goo.gl/gtyA8V>, [Online; accessed 19-March-2016] (2016).
- [24] P. S. Foundation., *Python language reference*, version 2.7., Available at <http://www.python.org> [Online].
- [25] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, C. Gagné, DEAP: Evolutionary algorithms made easy, *Journal of Machine Learning Research* 13 (2012) 2171–2175.
- [26] J. W. Taylor, Triple seasonal methods for short-term electricity demand forecasting, *European Journal of Operational Research* 204 (1) (2010) 139–152. doi:10.1016/j.ejor.2009.10.003.