

# Gaussian-Process-based Demand Forecasting for Predictive Control of Drinking Water Networks

Ye Wang, Carlos Ocampo-Martínez, Vicenç Puig, and Joseba Quevedo

Automatic Control Department, Technical University of Catalonia, Rambla Sant Nebridi, 10, 08222 Terrassa, Spain

{ywang, cocampo, vpuig}@iri.upc.edu, {joseba.quevedo}@upc.edu

**Abstract.** This paper focuses on water demand forecasting for predictive control of Drinking Water Networks (DWN) in the short term by using Gaussian Process (GP). For the predictive control strategy, system states in a finite horizon are generated by a DWN model and demands are regarded as system disturbances. The goal is to provide a demand estimation within a given confidence interval. For the sake of obtaining a desired forecasting performance, the forecasting process is carried out in two parts: the expected part is forecasted by Double-Seasonal Holt-Winters (DSHW) method and the stochastic part is forecasted by GP method. The mean value of water demand is firstly estimated by DSHW while GP provides estimations within a confidence interval. GP is applied with random inputs to propagate uncertainty at each step. Results of the application of the proposed approach to a real case study based on the Barcelona DWN have shown that the general goal has been successfully reached.

**Keywords:** Gaussian Process; Water Demand Forecasting; Drinking Water Networks; Double-Seasonal Holt-Winters; Predictive Control

## 1 Introduction

Water demand forecasting has been discussed and explored in the past decades for long-term and short-term forecasts. Short-term demand forecasting plays a significant role in the optimal operational control of Drinking Water Networks (DWN). To manage DWN, which are complex and large-scale systems, obtaining an accurate model of water demand is of great significance. In the model predictive control (MPC) strategy for DWN, water demands can be regarded as disturbances, being necessary to obtain the water demand evolution over a given prediction horizon.

As electricity demand forecasting, water demand forecasting is strongly influenced by meteorological factors, such as temperature and humidity. Even though other factors related to water demand evolution can be considered, it is still difficult to forecast water demand taking into account meteorological factors depending on time as well. In other words, if the temperature is chosen as a factor for forecasting water demands, the forthcoming information is probably

from weather forecast. As a result, the other factors are not always available and water demand is usually characterized as a time series model.

Gaussian Process (GP) regression model has been treated as the state-of-the-art regression methodology and applied in many different real cases such as electricity forecasting [9], [16] and disturbance forecasting in greenhouse temperature control system [13], among other fields. There are some other methodologies for electricity forecasting that have been discussed in the past decades, such as artificial neural networks [5], [8]. These algorithms have also been employed for the water demand forecasting [1], [11]. The superiority of GP regression comes from the use of the Bayesian Inference theory, which is able to update parameters of GP model in real time. In a GP model, it is assumed that the regression variables have a multivariate Gaussian distribution.

The idea of combining MPC and GP was proposed by [10]. It is suggested that GP could be an approach to model and forecast system disturbances and then being used to implement to MPC for a real system. The main difficulty of only applying GP to forecast system disturbances is that multiple-step ahead forecasts are required. At each step, some previous values will be used as testing inputs of GP regression model in order to obtain the estimation but probably some inputs are unknown at current time. If unknown values are replaced by estimates from GP at previous steps, the next estimation would be more inaccurate. Hence, modelling demand is divided into two parts: expected and stochastic parts.

Exponential smoothing methods are originally used to manipulate financial market and economic data and then widely applied to time series data [3], [7]. Together with complementary components of level, trend and seasonality, a short-term forecasting can be performed. Double-Seasonal Holt-Winters (DSHW) is an extended exponential smoothing method with two seasonalities. It is suitable for forecasting water demand with daily and weekly period at hourly time scale. Unlike GP regression, DSHW for multiple-step ahead forecasts is only based on the last known value that is regarded as the initial value.

Leading to a combined forecasting method, a quite proper mean estimation for expected water demand is forecasted by using DSHW. The stochastic water demand is found by subtracting expected water demands. The random inputs with a Gaussian distribution are considered as the testing inputs for GP [14]. The uncertainty propagation is carried out during multiple-step forecasts as well.

The main contribution of this paper consists in proposing a new algorithm denoted *DSHW-GP* to forecast short-term water demand for the purpose of incorporating it into an MPC-based closed-loop control topology. The advantage of this approach is to make use of accurate forecasting by DSHW as the expected part to avoid the drawback of GP for multiple-step ahead. After applying this approach, the forecasting uncertainty evolution of demand over the MPC prediction horizon will be used for propagating uncertainty of system states. Going even further, a robust MPC controller can be designed to deal with uncertainty propagation of system states being alternative to the one proposed in [18].

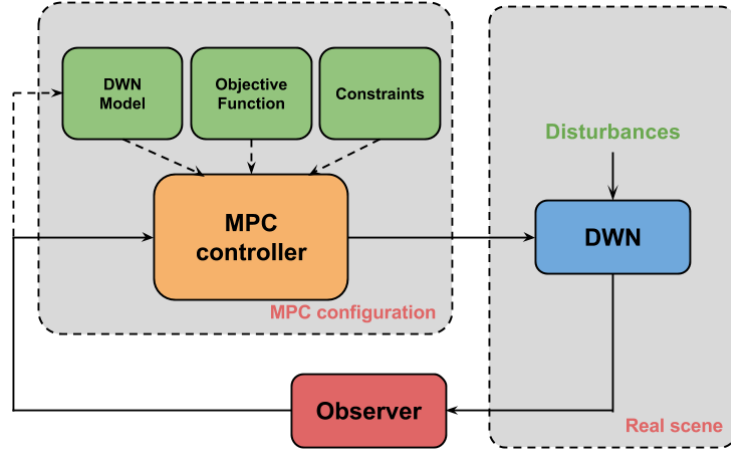
The reminder of this paper is structured as follows. In Section 2, the proposed approach including detailed equations of DSHW and GP for regression and the

DSHW-GP algorithm are presented. In Section 3, a real case study based on the Barcelona DWN is used for testing the proposed methodology in this paper and simulation results are also shown. Finally, main conclusions are drawn in Section 4.

## 2 Proposed Approach

### 2.1 MPC Framework and DWN Control-oriented Model

Fig. 1 shows the general MPC closed-loop scheme for DWN. In the labelled *Real scene* block, measurement sensors in the DWN are often influenced by disturbances. Current system states are estimated by the observer that depends on measurements obtained from the system sensors. In the *MPC configuration* block, a DWN model including system disturbances is required, which will be used to predict both the system states and outputs over a given time horizon. The general MPC controller design for DWN can be found in [12].



**Fig. 1.** Model Predictive Control (MPC) scheme for DWN

The control-oriented model for DWN considered in this paper is described by the following set of linear discrete difference-algebraic equations for all time instant  $k \in \mathbb{N}$  [6]:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{B}_d\mathbf{d}_k, \quad (1a)$$

$$\mathbf{0} = \mathbf{E}_u\mathbf{u}_k + \mathbf{E}_d\mathbf{d}_k, \quad (1b)$$

where  $\mathbf{x}_k$ ,  $\mathbf{u}_k$ ,  $\mathbf{d}_k$  denote the state vector, the manipulated flows through actuators and the demanded flow as additive measured disturbances, respectively. Moreover, (1a) describes the dynamics of storage tanks and (1b) presents the static relations within the DWN at network nodes.

**Assumption 1** *The water demands over the MPC prediction horizon  $H_p$  from the current time  $k$  are decomposed as*

$$\hat{\mathbf{d}}_{k+i} = \bar{\mathbf{d}}_{k+i} + \Sigma_{\mathbf{d}_{k+i}} \quad i = 1, 2, \dots, H_p, \quad (2)$$

where  $\bar{\mathbf{d}}_{k+i}$  is the vector of expected water demand, and  $\Sigma_{\mathbf{d}_{k+i}}$  is the vector of probabilistic independent uncertainty forecasting, i.e., stochastic demand.

As aforementioned, the expected demand  $\bar{\mathbf{d}}_{k+i}$  could be forecasted by using DSHW, and the stochastic demand  $\Sigma_{\mathbf{d}_{k+i}}$  could be forecasted by using GP. Moreover, the GP could also generate a confidence interval considering the demand forecasting errors.

## 2.2 GP Regression with Uncertainty Propagation

GP is regarded as a supervised learning algorithm widely used for different domains in the past decades. GP regression can be used for identifying the model of a dynamic system. The model identified by GP regression is so called *non-parametric model* [4], which does not mean there are no parameters inside the model but the model has flexible parameters that can be adapted from input data. Hence, GP regression is used for the state-of-the-art regression methods [4] and includes non-parametric model with Bayesian inference methods. As for the so-called parametric model, parameters impose a fixed structure or value in advance upon the model. However, the GP regression is an optimal approach to make the model more flexible. With different training data, the GP model can be adapted accordingly.

The general GP regression model can be defined as

$$f(\mathbf{z}) \sim \mathcal{GP}(m(\mathbf{z}), k(\mathbf{z}, \mathbf{z}')), \quad (3)$$

where  $\mathbf{z}$  is the feature vector (inputs) of the GP model while  $m(\mathbf{z})$  and  $k(\mathbf{z}, \mathbf{z}')$  are mean and covariance functions for GP, whose formats should be firstly defined with some certain parameters called *hyperparameters*. These hyperparameters can be selected by using Bayesian Inference methods with training data. Usually, GP is used for modelling and forecasting a set of random variables [15]. The mean function is usually null. Then, the GP model is rewritten in the following form:

$$f(\mathbf{z}) \sim \mathcal{GP}(\mathbf{0}, k(\mathbf{z}, \mathbf{z}')). \quad (4)$$

The forecasts of GP can be performed with noises, i.e.,  $y = f(\mathbf{z}) + \epsilon$ . It is assumed that the noise  $\epsilon$  obeys a Gaussian distribution  $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ . The joint distribution of the observation outputs  $y$  and the testing outputs  $\mathbf{f}_*$  is defined as

$$\begin{bmatrix} y \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left( \mathbf{0}, \begin{bmatrix} K(\mathbf{z}, \mathbf{z}) + \sigma_n^2 I & K(\mathbf{z}, \mathbf{z}_*) \\ K(\mathbf{z}_*, \mathbf{z}) & K(\mathbf{z}_*, \mathbf{z}_*) \end{bmatrix} \right), \quad (5)$$

where  $\mathbf{z}_*$  is a set of testing inputs and  $I$  denotes the identity matrix of suitable dimensions. Moreover,  $K(\mathbf{z}, \mathbf{z})$ ,  $K(\mathbf{z}, \mathbf{z}_*)$ ,  $K(\mathbf{z}_*, \mathbf{z})$  and  $K(\mathbf{z}_*, \mathbf{z}_*)$  are covariance matrices. The detailed definitions of covariance matrices can be found in [16].

Deriving the conditional distribution, it is possible to arrive at the key forecasting expression for the GP regression as

$$\mathbf{f}_* | \mathbf{z}, \mathbf{y}, \mathbf{z}_* \sim \mathcal{N}(m(\mathbf{f}_*), k(\mathbf{f}_*)), \quad (6)$$

where  $m(\mathbf{f}_*)$  and  $k(\mathbf{f}_*)$  are posterior mean and covariance functions, respectively, which are given as

$$m(\mathbf{f}_*) \triangleq K(\mathbf{z}_*, \mathbf{z})[K(\mathbf{z}, \mathbf{z}) + \sigma_n^2 I]^{-1} \mathbf{y}, \quad (7a)$$

$$k(\mathbf{f}_*) \triangleq K(\mathbf{z}_*, \mathbf{z}_*) - K(\mathbf{z}_*, \mathbf{z})[K(\mathbf{z}, \mathbf{z}) + \sigma_n^2 I]^{-1} K(\mathbf{z}, \mathbf{z}_*). \quad (7b)$$

For selecting the feature vector, the candidate feature variables come from previous target variables in a time series model. For the time series data, previous  $N$  data ( $d(k-1), \dots, d(k-N)$ ) from current time  $k$  are chosen as the feature vector.

For multiple-step ahead forecasts, the difficulty of applying the aforementioned method is that the previous real demand is unknown at each step ahead. One solution is provided by using a random inputs as feature vector that obeys a Gaussian distribution [14]. In this way, uncertainty could be propagated during the process of multiple-step ahead forecasts. The testing inputs are  $\mathbf{z}_* \sim \mathcal{N}(\mu_{\mathbf{z}_*}, \Sigma_{\mathbf{z}_*})$ , whose definitions can be found in [14]. Performing a Taylor expansion around  $\mathbf{z}_*$  in (7a) and (7b), the final forecasts are shown as

$$m(\mu_{\mathbf{z}_*}, \Sigma_{\mathbf{z}_*}) = m(\mathbf{f}_*), \quad (8a)$$

$$\begin{aligned} k(\mu_{\mathbf{z}_*}, \Sigma_{\mathbf{z}_*}) &= k(\mathbf{f}_*) + \frac{1}{2} \mathbf{Tr} \left\{ \frac{\partial^2 k(\mathbf{z}_*)}{\partial \mathbf{z}_* \partial \mathbf{z}_*^T} \Big|_{\mathbf{z}_* = \mu_{\mathbf{z}_*}} \Sigma_{\mathbf{z}_*} \right\} \\ &\quad + \frac{\partial m(\mathbf{z}_*)}{\partial \mathbf{z}_*} \Big|_{\mathbf{z}_* = \mu_{\mathbf{z}_*}}^T \Sigma_{\mathbf{z}_*} \frac{\partial m(\mathbf{z}_*)}{\partial \mathbf{z}_*} \Big|_{\mathbf{z}_* = \mu_{\mathbf{z}_*}}, \end{aligned} \quad (8b)$$

where  $\mathbf{Tr}$  denotes the trace operator. Moreover, the first and second order derivatives are computed as

$$\frac{\partial m(\mathbf{z}_*)}{\partial \mathbf{z}_{*d}} \Big|_{\mathbf{z}_* = \mu_{\mathbf{z}_*}} = \left[ -\frac{1}{2l^2} (\mathbf{z}_d - \mu_{\mathbf{z}_{*d}}) K(\mu_{\mathbf{z}_*}, \mathbf{z}) \right]^T K^{-1}(\mathbf{z}, \mathbf{z}) \mathbf{y}, \quad (9a)$$

$$\begin{aligned} \frac{\partial^2 k(\mathbf{z}_*)}{\partial \mathbf{z}_{*d} \partial \mathbf{z}_{*e}^T} \Big|_{\mathbf{z}_* = \mu_{\mathbf{z}_*}} &= -2 \left( -\frac{1}{2l^2} \right)^2 \left\{ M(\mathbf{z}_{*d})^T K^{-1}(\mathbf{z}, \mathbf{z}) M(\mathbf{z}_{*e}) \right. \\ &\quad \left. + [(\mathbf{z}_d - \mu_{\mathbf{z}_{*d}}) (\mathbf{z}_e - \mu_{\mathbf{z}_{*e}}) K(\mu_{\mathbf{z}_*}, \mathbf{z})]^T K^{-1}(\mathbf{z}, \mathbf{z}) K(\mu_{\mathbf{z}_*}, \mathbf{z}) \right\} \\ &\quad + 2 \left( -\frac{1}{2l^2} \right) K(\mu_{\mathbf{z}_*}, \mathbf{z})^T K^{-1}(\mathbf{z}, \mathbf{z}) K(\mu_{\mathbf{z}_*}, \mathbf{z}) \delta_{de}, \end{aligned} \quad (9b)$$

$$M(\mathbf{z}_{*i}) = (\mathbf{z}_i - \mu_{\mathbf{z}_{*i}}) K(\mu_{\mathbf{z}_*}, \mathbf{z}), \quad (9c)$$

where  $l$  is a parameter of covariance function and  $\mathbf{z}_d \in \mathbb{R}^d$  and  $\mathbf{z}_e \in \mathbb{R}^e$  are different column vectors of input data. Further detailed calculations can be found in [14].

### 2.3 Double-Seasonal Holt-Winters

The exponential smoothing method was firstly introduced by R. G. Brown in 1956 and later improved by C. C. Holt and P. R. Winters with trend and seasonal components, which is called *Holt-Winters* (HW) method. This method is usually applied to time series data in order to generate short-term forecasts [3].

Simple exponential smoothing did not consider the time series data with tendency and periodicity. HW method contains these two features but only with one period in an additive or multiplicative way. Afterwards, it is extended the single period to double multiplicative seasonality [17]. This method is the so-called *Double-Seasonal Holt-Winters* (DSHW).

Some comparisons with several exponential smoothing methods have been discussed in [2] and it is concluded that DSHW method can provide forecasting results with more robustness and accuracy.

The DSHW model for water demand is built as follows:

$$\hat{d}(k+j|k) = (L(k) + jT(k))S_1 \left( k + j - \left\lfloor \frac{j}{s_1} \right\rfloor s_1 \right) S_2 \left( k + j - \left\lfloor \frac{j}{s_2} \right\rfloor s_2 \right), \quad (10)$$

where  $L(k)$ ,  $T(k)$ ,  $S_1(k)$ ,  $S_2(k)$  denote *level*, *trend* and *two seasonalities*, respectively:  $S_1(k)$  is the first season  $s_1$  while  $S_2(k)$  is the second season  $s_2$  and  $j$  is the forecasting index within a given horizon. To compute these components, the following expressions are used:

$$L(k) = \alpha \frac{d(k)}{S_1(k-s_1)S_2(k-s_2)} + (1-\alpha)(L(k-1) + T(k-1)), \quad (11a)$$

$$T(k) = \gamma(L(k) - L(k-1)) + (1-\gamma)T(k-1), \quad (11b)$$

$$S_1(k) = \delta_1 \frac{d(k)}{L(k)S_2(k-s_2)} + (1-\delta_1)S_1(k-s_1), \quad (11c)$$

$$S_2(k) = \delta_2 \frac{d(k)}{L(k)S_1(k-s_1)} + (1-\delta_2)S_2(k-s_2), \quad (11d)$$

where  $\alpha, \gamma, \delta_1, \delta_2$  are smoothing parameters that can be obtained by using least-squared methods with given training data. In principle, a collection of training dataset in two suitable periods should be achieved at initial forecasting time  $k_{ini}$ .

### 2.4 DSHW-GP Approach

In this paper, the proposed DSHW-GP approach is shown in Algorithm 1. Since both of DSHW and GP forecasting models need to be trained before forecasting, it is assumed that a collection of past data is available. Meanwhile, the DSHW

loop can be run with past data in a certain time in order to obtain the training data for GP loop. In this approach, the effectiveness and efficiency are both considered. Assuming the periodicity of the water demand with period  $\Delta_p$ , due to the accuracy of the DSHW method, the calculation process can be reduced to be executed each  $2\Delta_p$ , which will be partially chosen as estimations of expected water demand with the horizon of  $\Delta_p$ . Hence,  $H_p$  is considered equal to  $\Delta_p$  in this case.

---

**Algorithm 1** DSHW-GP Algorithm
 

---

```

1:  $n \leftarrow$  Simulation Days
2:  $k \leftarrow$  Current hour
3: for  $i \leftarrow 1 : n$  do
4:    $tdd \leftarrow$  Get past a set of real demands ▷ DSHW Loop
5:   Training DSHW by  $tdd$  with two periods  $(s_1, s_2)$ 
6:    $dm \leftarrow dshw(k, 2\Delta_p)$ :  $2\Delta_p$ -step ahead at time  $k$ 
7:   for  $j \leftarrow 1 : \Delta_p$  do ▷ GP Loop
8:      $dmp \leftarrow$  Past a set of estimations of DSHW
9:      $dsto \leftarrow d_{total} - dmp$ 
10:    Training GP model with  $dsto$ 
11:     $mean, cov, lb, ub \leftarrow GP(k + j, \Delta_p)$  ▷ Prediction by GP with random inputs
12:     $\bar{\mathbf{d}}(i, j) \leftarrow dm(j : j + \Delta_p - 1) + mean$ 
13:     $\Sigma_{\mathbf{d}(i, j)} \leftarrow cov$ 
14:     $\mathbf{dlb}(i, j) \leftarrow lb$ 
15:     $\mathbf{dub}(i, j) \leftarrow ub$ 
16:   end for
17:    $k \leftarrow k + \Delta_p$ 
18: end for

```

---

**Remark.** For daily forecasts, DSHW loop is only executed at the first hour of a day with training data. The forecasting results include  $2\Delta_p$  demand estimations that will be regarded as the expected demand of hourly forecasts. The procedure is as follows: at time  $k$ , expected estimations are selected from  $k + 1$  to  $k + \Delta_p$ . At time  $k + 1$ , expected estimations are selected from  $k + 2$  to  $k + \Delta_p + 1$ . Until time  $k + \Delta_p$ , expected estimations are selected from  $k + \Delta_p + 1$  to  $k + 2\Delta_p$ . The DSHW loop is executed daily while the GP loop is executed hourly. The total estimation contains two parts coming from selected DSHW and GP loops, respectively. The total mean estimation is the sum of results from DSHW and GP. Upper and lower forecasting bounds are produced by GP.

### 3 Case Study: Barcelona Drinking Water Network

#### 3.1 Case-Study Description

The proposed approach is applied to the case study of Barcelona DWN. The Barcelona DWN supplies  $237.7 \text{ hm}^3$  water to approximately 3 million consumers every year, covering a  $424 \text{ km}^2$  area. The entire network is composed of 63 storage tanks, 3 surface sources, 7 underground sources, 79 pumps, 50 valves, 18 nodes and 88 water demands. The topology of Barcelona DWN is described in Fig.

2. Currently, AGBAR<sup>1</sup> is in charge of managing the entire network through a supervisory control system with sampling period of one hour. It is necessary to forecast the water demands of the whole network within an MPC strategy with a prediction horizon of 24 hours. The improved forecasts of water demands could lead to obtain huge economic benefits. The quality of gathered real data has much influence on demand forecasting results due to unexpected noise from sensors. After comparing and selecting different sets of real data, the real water demand data of *C10COR* during the year 2013 will be used to illustrate the proposed approach. Similar results can be obtained in case of other water demands of the case study.

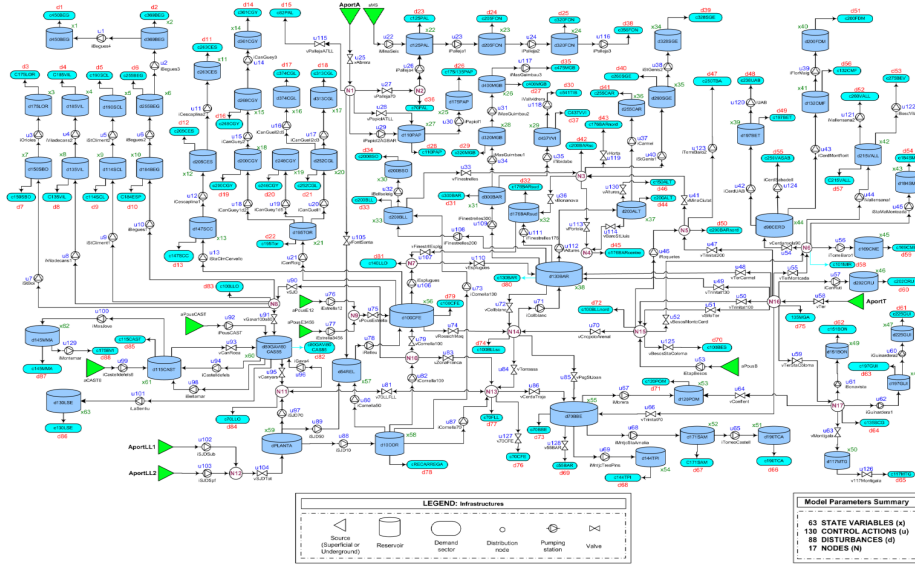


Fig. 2. Barcelona DWN topology

### 3.2 Results

Looking into the dataset of real water demands available, there are approximately one year's data available. From this set of data, the daily and weekly periods can be clearly observed. For the simulation in this paper, one month and half data set is divided into the testing data set and the remaining data set is used for the validation. The simulation is running for a scenario of two days (48 hours). Comparing the forecasts and real values of water demands, the error measurements are calculated by using the key performance indicators (KPIs) defined as

*Mean Squared Error (MSE):*

$$MSE = \frac{1}{n} \sum_{t=1}^n (R_t - P_t)^2, \quad (12)$$

<sup>1</sup> AGBAR: Aguas de Barcelona, S. A. Company which manages the drinking water transport and distribution in Barcelona (Spain).



Mean Absolute Error (MAE):

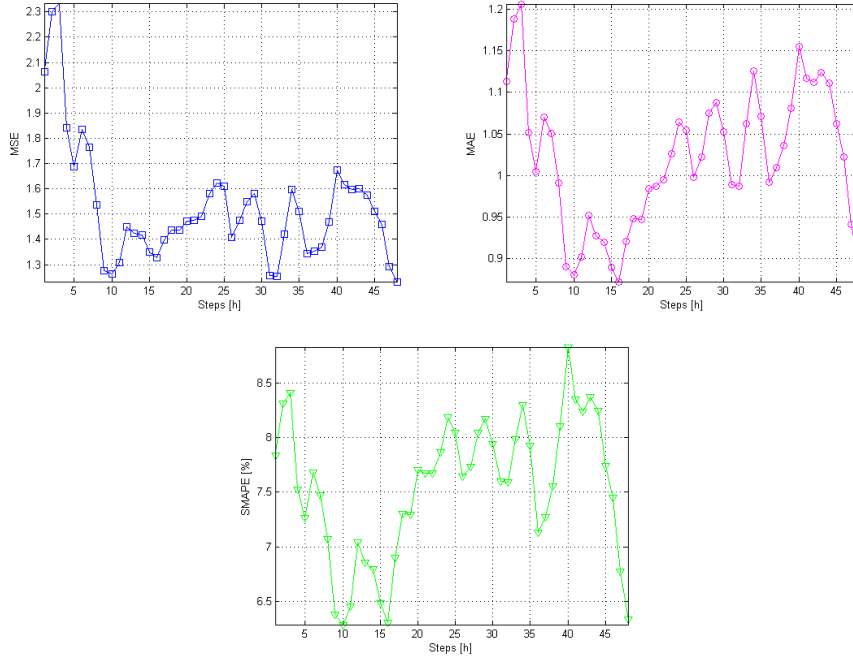
$$MAE = \frac{1}{n} \sum_{t=1}^n |R_t - P_t|, \quad (13)$$

Symmetric Mean Absolute Percentage Error (SMAPE):

$$SMAPE = \frac{100}{n} \sum_{t=1}^n \frac{|R_t - P_t|}{R_t + P_t}, \quad (14)$$

where  $R_t$  denotes real value of the drinking water demand from validation data and  $P_t$  denotes the forecasting mean value of the water demand obtained by the DSHW-GP algorithm. In terms of MSE and MAE, they are representing the difference between the actual observation and the observation values forecasted by the model. Moreover, SMAPE is an accurate measurement based on percentage errors, which are adapted to compute time-series data. In this case study,  $\Delta_p=24$  hours.

According to hourly-scale forecasts repeated 48 times, KPIs are shown in the Fig. 3. Plots of MSE and MAE show that they are varying in a small interval (no more than 1). SMAPE belongs to the range between 0% and 100%. If the practical value of SMAPE is near 0%, the forecasting results are quite accurate. In this case, the general SMAPE is between 6% and 9% (never greater than 10 %).



**Fig. 3.** Error measurements: MSE, MAE and SMAPE

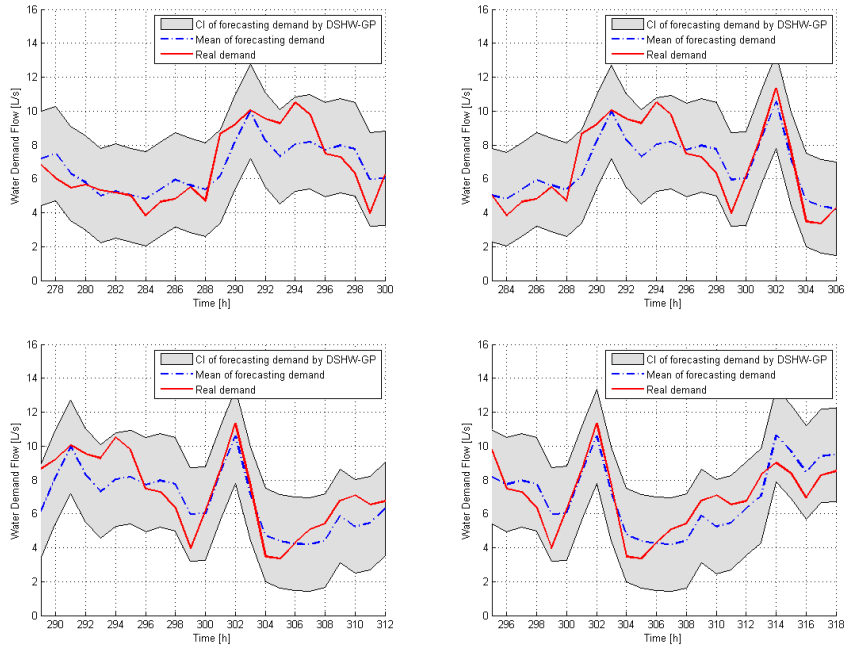
The forecasting result for each step ahead is a Gaussian distribution. The confidence interval (CI) can be obtained as follows.

$$\mathbf{d}_k \in \left[ \bar{\mathbf{d}}_k - \frac{c}{\sqrt{P}} \Sigma_{d_k}^{1/2}, \bar{\mathbf{d}}_k + \frac{c}{\sqrt{P}} \Sigma_{d_k}^{1/2} \right], \quad (15)$$

where  $P$  is the number of samples, which is equal to 1 in terms of one-step ahead forecast. Moreover,  $c$  denotes the critical value with respect to a confidence level, such as 95% or 98%. The calculation of this level is done by means of the inverse standard probability density function, which is shown as

$$c = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right), \quad (16)$$

where  $c$  is the critical value with respect to the confidence level  $(1 - \frac{\alpha}{2})$ .



**Fig. 4.** A sequence of simulation results

In many applications of GP, confidence level is chosen between 90% and 100% since a large number could imply that some unexpected noises gathered by using different sensors stay inside the confidence interval. Hence, the critical values are around 2 when the confidence levels are chosen inside the aforementioned interval. Fig. 4 shows a sequence of selected simulation results in 48-step forecasts. The gray area denotes the confidence interval with critical value equal to 2.

The real demand is approximately around the mean estimation in Fig. 4. Sometimes the mean estimation does not perfectly match the real demand since the latter probably contains some unexpected noisy measurements from sensors. In terms of GP, the challenge is how to select a proper feature vector for a real case and get the accurate testing inputs. In this case study, the goal of this work has been properly reached and the real water demands are inside the confidence interval.

## 4 Conclusions

In this paper, the DSHW-GP algorithm has been proposed and applied to the water demand forecasting for DWN management. DSHW and GP have their own strengths and drawbacks. The approach of DSHW-GP takes advantages of two methods while avoids drawbacks of both of them. The DSHW is used for modelling expected part of water demand while GP is used for modelling stochastic part of water demand. This approach is tested in the Barcelona DWN. Results show that it is useful for forecasting water demand in a short term achieving a confidence interval at the same time. The forecasting results can be applied to robust MPC to consider for the possible worst-case demand scenario.

Further work is focused on applying this approach to an MPC-based closed-loop scheme. The mean and bounds of demand forecasting obtained by using the DSHW-GP algorithm will be used to compute estimates of system states in order to design a robust MPC controller. Besides, the demand forecasting method can be used for guaranteeing a reliable supply in the water networks by means of avoiding unexpected uncertainties in a short-term future.

**Acknowledgments.** This work is partially supported by the research projects CICYT SHERECS DPI-2011-26243 and ECOCIS DPI-2013-48243-C2-1-R, both of the Spanish Ministry of Education, by EFFINET grant FP7-ICT-2012-318556 of the European Commission and by AGAUR Doctorat Industrial 2013-DI-041. Ye Wang also thanks China Scholarship Council for providing postgraduate scholarship.

## References

1. M. Bakker, J. H. G. Vreeburg, K. M. van Schagen, and L. C. Rietveld. A fully adaptive forecasting model for short-term drinking water demand. *Environmental Modelling and Software*, 48(0):141–151, 2013.
2. J. Blanch, J. Quevedo, J. Saludes, and V. Puig. Short-term demand forecasting for operational control of the barcelona water transport network. In *Conferencia Nacional de Jóvenes Profesionales del Agua de España*, pages 1–10, Barcelona, 2010.
3. W. R. Christiaanse. Short-term load forecasting using general exponential smoothing. *IEEE Transactions on Power Apparatus and Systems*, 90(2):900–911, 1971.

4. M. P. Deisenroth. *Efficient reinforcement learning using Gaussian processes*. PhD thesis, Karlsruhe Institute of Technology, 2010.
5. Y. Fung and V. Rao Tummala. Forecasting of electricity consumption: a comparative analysis of regression and artificial neural network models. In *2nd International Conference on Advances in Power System Control, Operation and Management*, volume 2, pages 782–787, 1993.
6. J. M. Grosso, C. Ocampo-Martínez, V. Puig, and B. Joseph. Chance-constrained model predictive control for drinking water networks. *Journal of Process Control*, 24(5):504–516, 2014.
7. P. J. Harrison. Exponential smoothing and short-term sales forecasting. *Management Science*, 13(11):821–842, 1967.
8. M. Hayati and Y. Shirvany. Artificial neural network approach for short term load forecasting for illam region. *World Academy of Science: Engineering and Technology*, 22:280–284, 2007.
9. J. M. Lourenco and P. J. Santos. Short-term load forecasting using a gaussian process model: The influence of a derivative term in the input regressor. *Intelligent Decision Technologies*, 6(4):273–281, 2012.
10. J. M. Maciejowski and X. Yang. Fault tolerant control using gaussian processes and model predictive control. In *2013 Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 1–12, Nice, 2013.
11. I. S. Msiza, F. V. Nelwamondo, and T. Marwala. Water demand prediction using artificial neural networks and support vector regression. *Digital intelligence*, 3(11):1–8, 2008.
12. C. Ocampo-Martínez, V. Puig, G. Cembrano, and J. Quevedo. Application of MPC strategies to the management of complex networks of the urban water cycle. *IEEE Control Systems Magazine*, 33(1):15–41, 2013.
13. A. Pawlowski, J. L. Guzman, F. Rodriguez, M. Berenguel, and J. E. Normey-Rico. Predictive control with disturbance forecasting for greenhouse diurnal temperature control. In *The 18th IFAC World Congress*, pages 1779–1784, Milano, 2011.
14. J. Quinonero-Candela, A. Girard, and C. E. Rasmussen. Prediction at an uncertain input for gaussian processes and relevance vector machines application to multiple-step ahead time-series forecasting. Technical report, University of Glasgow, Department of Computing Science, 2002.
15. C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. ISBN 026218253X. the MIT Press, Massachusetts Institute of Technology, 2006.
16. M. Samarasinghe and W. Al-Hawani. Short-term forecasting of electricity consumption using gaussian processes. Master’s thesis, University of Agder, 2012.
17. J. W. Taylor. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*, 54(8):799–805.
18. Y. Wang. Model predictive control for drinking water networks based on gaussian processes. Master’s thesis, Technical University of Catalonia, 2014.