

# SMaRT-OnlineWDN D5.5: Development of demand forecast software

Mathias Braun, Thomas Bernard, Idel Montalvo Arango, Jochen Deuerlein,

Olivier Piller

# ► To cite this version:

Mathias Braun, Thomas Bernard, Idel Montalvo Arango, Jochen Deuerlein, Olivier Piller. SMaRT-OnlineWDN D5.5: Development of demand forecast software. [Research Report] irstea. 2014, pp.8. hal-02601405

# HAL Id: hal-02601405 https://hal.inrae.fr/hal-02601405

Submitted on 16 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





# Fraunhofer





# Deliverable 5.5

Development of demand forecast software

Dissemination level: Public

WP5 Online Simulation Module

27th May 2014











SMaRT-Online WDN

Online Security Management and Reliability Toolkit for Water Distribution Networks

ANR reference project: BMBF reference project: ANR-11-SECU-006 13N12180

Contact persons Fereshte SEDEHIZADE Olivier PILLER

Fereshte.Sedehizade@bwb.de Olivier.Piller@irstea.fr

Development of demand forecast software 27th May 2014

### WP 5 – Online Simulation

#### D5.5 Development of demand forecast software

#### List of Deliverable 5.5 contributors:

#### From 3S [WP5 leader]

Jochen Deuerlein (<u>deuerlein@3sconsult.de</u>) Idel Montalvo (<u>montalvo@3sconsult.de</u>)

#### From Irstea

Olivier Piller (olivier.piller@irstea.fr)

#### From IOSB

Thomas Bernard (<u>Thomas.Bernard@iosb.fraunhofer.de</u>) Mathias Braun (<u>Mathias.Braun@iosb.fraunhofer.de</u>)

Work package number	5.5		Start date:	01/10/2012							
Contributors	3S Consult	Irstea	IOSB								
Person-months per partner	2	1	1								
Keywords Time series analysis – Comp	outational Intell	igence – Dema	and Predictior	1							
Objectives											
Description of the development and implementation of the demand prediction module as part of the online simulation module.											

### **Summary**

The main objective of the SMaRT-Online<sup>WDN</sup> project is the development of an online security management toolkit for water distribution networks which is based on sensor measurements of water quality as well as water quantity. The development of a demand prediction module is one objective of the "online simulation" project work package 5 of. The task is the development of a prediction module that allows for the prediction of a single future water demand and the determination of a complete demand pattern for the following 24 hour cycle.

The objective of this deliverable is to give a report of the development and implementation of the demand prediction module. Initially, section 1 gives the context of the demand prediction module in the project. Section 2 explains the objectives of the module and how they are integrated in the workflow of the SMaRT-Online<sup>WDN</sup> process. Section 3 gives the details on Prediction algorithm. Section 4 proceeds with details on the software implementation. A conclusion on the development is given in section 5.

List of Figures:	
Figure 1: Workflow for One Step Demand Prediction in Online Calibration	6
Figure 2: Workflow for Demand Pattern Prediction in SMaRT Online	6
Figure 3: Demand Prediction GUI using the DemandForecasting.dll	8

#### Contents:

1.	Background	. 5
2.	Aims and Workflow of the Demand Prediction Module	5
3.	Algorithm of the Demand Prediction Module	6
4.	Software-Implementation and Interfaces	. 7
5.	Conclusion	8
6.	Referred bibliography	8

# 1. Background

The SMaRT-Online<sup>WDN</sup> security management toolkit is based on sensor measurements (hydraulic and water quality sensors) and online hydraulic and transport models and will allow an estimation of the localization of the contamination source and the simulation of short-time future scenarios in order to estimate the impact of a contamination source. In order to make an accurate short-term prediction for contamination scenarios a good prediction for the hydraulic state of the network is needed. The hydraulic state is strongly influenced by the water demand in the observed area what poses the need for an accurate prediction and calibration of the water demand (Piller *et al.*, 2014 and Preis *et al.*, 2010). Since water demand is subject to high seasonality it is prudent to use the tools of time series analysis to model the water demand and predict the consumption in a district or consumption nodes of the network.

In time series analysis autoregressive integrate moving average (ARIMA) models have been used for decades and due to their success they have been used in a wide variety of scientific applications. In recent years a growing popularity of machine learning algorithms like the artificial neural network (ANN) and support vector machine (SVM) have led to new approaches in time series analysis.

# 2. Aims and Workflow of the Demand Prediction Module

There are two basic aims for the demand prediction module called the One-Step-Prediction and the Pattern-Prediction. The Objective of the One-Step-Prediction is to predict the water demand for an area or node of the network for a predetermined time step (i.e. fifteen minutes) into the future. The objective of the Pattern-Prediction is to predict the water demand in an area or node of the network over a predetermined timespan in a discrete number of steps (i.e. during the next 24 hours in hourly steps).

These two modules are integrated into the SMaRT-Online<sup>WDN</sup> toolkit in two different workflows. The results of One-Step-Prediction are used as initial estimation by the online demand calibration. Calibration of the demand for the last observation period consists of seeking demande estimates that minimize the sum of squared residuals between the observed and hydraulic and transport model values. Then the initial demand prediction are corrected thanks to the last measurements (flow rates, pressures, concentrations....). To do this the customers in the network model are divided into a number of customer groups. Based on the adapted prediction model and historic demand data the water consumption for each customer is first predicted by the One-Step-Prediction tool then corrected (adjusted) thanks to additional information in the measurement and the simulation models. The workflow of the Online-Calibration is shown in Figure 1.



Figure 1: Workflow for One Step Demand Prediction in Online Calibration

The second one is the Look-Ahead simulation shown in Figure 2. In case of a contamination event detected by the alarm generation the SMaRT Online process starts a Look-Ahead simulation to determine the propagation of the contamination throughout the network. For this task the hydraulic state of the network has to be determined based on the prediction of the water consumption.



Figure 2: Workflow for Demand Pattern Prediction in SMaRT Online

# 3. Algorithm of the Demand Prediction Module

For the development of the prediction module different algorithms have been researched and evaluated. Namely the classes of autoregressive integrated moving average (ARIMA) models and support vector regression (SVR). Due to their superior ability in modeling non-linear effects the SVR is chosen as basis for the demand forecasting module.

Support Vector Machines are supervised learning algorithms from statistical learning theory that are used to recognize patterns in data for classification. Support Vector machines for function estimation are also referred to a Support Vector Regression (SVR). For the demand prediction algorithm the  $\varepsilon$ -SVR is used. The goal of  $\varepsilon$ -SVR is to find a function that has a maximum distance of  $\varepsilon$  from the actual measurement points  $y_i$  and at the same time is as flat as possible. For points that do not fall within the margin  $\varepsilon$  an additional cost is added proportional to its distance from the margin. These points define the support vectors. The model has two independent parameters C and  $\varepsilon$  that have to be adjusted for fitting the model to the data. The cost factor C > 0 which determines the trade-off between the flatness of the regression line and the distance to the margin  $\varepsilon$  that is tolerated. The SVR is linear process, however by applying

Development of demand forecast software 27th May 2014

kernel functions the feature space can be transformed and the same method can be used for nonlinear problems. The kernel function can be defined by polynomial functions, Gaussian radial basis, hyperbolic tangent and many more. In the following the radial basis Gaussian kernel will be used.

The inputs to the model are given by past demand values with a time lag of  $t - lag_{min}$ , t - 24, t - 25, t - 168 and t - 169. This pays respect to how far into the future the prediction goes, the daily and the weekly cycles in water demand. Additional information is given by the time of day, the day of the week and the month. Based on these factors and the model the water demand is predicted. More detailed information on the algorithm selection and application are given in (Braun *et al.*, 2014)

# 4. Software-Implementation and functionality

The software is implemented using C# and the .NET Framework. For the implementation the Accord.NET Framework is used under the GNU Lesser General Public License (LGPL) license.

The core of the demand prediction is implemented in the DemandForecasting.dll. It contains the class *CForecastingProblem* that contains the kernel support vector machine and a list of normalized historic data elements which contain the information defined in the previous chapter. The central methods are given by *DoTrain* and *DoPredictSingle*. The first method trains the kernel support vector machine with the input data. The second method predicts the demand for a single normalized input data element. Three further functions are given by *DoPredictMany* which predicts the demand for a list of input elements, *Save* to save the instance of the *CForecastingProblem* in a binary data format and *Load* to load a particular instance of *CForecastingProblem*.

For testing and demonstration purposes the DemandForecasting.dll has been integrated into a simple GUI application shown in Figure 3. It contains two tabs called "Training Data" and "Testing Data". Both windows are designed similarly, showing the input data elements in the table on the left and the visualization of the time series on the upper right. For training the SVM, a prediction span is chosen from the combo box (here reaching from 1 up to 14 hours) a set of training data is selected on the button "Load Training Data" and the training is started by the button "Train". Afterwards the SVM parameter estimation is saved. For testing, the SVM is selected by choosing the time span from the combo box. A testing set different from the training set is selected by clicking on "Load Testing Data". The prediction is performed by clicking "Test". The results are shown in the Visualization together with the correct data and the relative error.

🛛 SVR_I	emand_Prediction	-																						3
File																								
Training	Data Test Data																							
	DataTimeStamp	DemandValue	DemandForecast	DemandValuePredi	DemandValueDayA	DemandValueWeel	PredictionError *																	
	04.04.2010 17:00	7101	7408,169211598	6950	8139	9964	4,3257176679	35000	T			1								Т		Demi Demi	andValue andForeca	ter
	04.04.2010 18:00	8045	8041,022537409	7101	8950	11126	-0,0494401813															- Predi	ctionError	÷.,
	04.04.2010 19:00	8044	8625,690417043	8045	8503	11162	7,2313577454	30000		-										-				
	04.04.2010 20:00	7236	7924,593583662	8044	7574	9040	9,5162186796!																	
	04.04.2010 21:00	6435	6951,015924861	7236	6503	7476	8.0188954912	05000																
	04.04.2010 22:00	6472	6411,472112977	6435	6489	6131	-0.9352269317	25000	1															
	04.04.2010 23:00	5712	5590,997392809	6472	6223	4738	-2,1183929830																	
	05.04.2010	4412	4103,920584636	5712	4701	4008	-6.9827610009	2000	+	-							- 1							
	05.04.2010 01:00	3877	3541,977344643	4412	3983	3611	-8,6412859261							A			. 1		1					
	05.04.2010 02:00	3442	3443,262142290	3877	3774	3513	0,03666886371	15000							11	Λ.								
	05.04.2010 03:00	3353	3401,235270138	3442	3315	3472	1,4385705379		Λ		1 . 1				$\Lambda \Lambda$			V	1 I V					
	05.04.2010 04:00	3394	3829,189272254	3353	3696	3598	12,822312087				$\Lambda \Lambda \Lambda$	V	11		1 1	V		1	110					
	05.04.2010 05:00	3599	4539,935051830	3394	4031	5234	26,144347091	10000			+					+ +			++-	+				
	05.04.2010 06:00	4423	5881,868216658	3599	4455	8188	32,983681136																	
	05.04.2010 07:00	6131	6337,333003536	4423	6082	10728	3,3654053749;	500	•₩_			$\downarrow$					Ų		ų_	_				
	05.04.2010 08:00	8540	8542,190793655	6131	9140	11292	0,0256533214		V	ľ	v	v .	•	ľ	*	v.	Č.,	ľ						
	05.04.2010 09:00	11117	10116,85715358	8540	11287	10887	-8,9965174634																	
	05.04.2010 10:00	11433	9924,461519128	11117	11023	9633	-13,194598800	· ·		-				-										
	05.04.2010 11:00	11229	10210,13484317	11433	10230	9134	-9.0735164023																	
	05.04.2010 12:00	10465	10455,83883035	11229	9778	8247	-0.0875410381	-5000		1		L						1						
	05.04.2010 13:00	9456	9452,953688927	10465	8675	8050	-0.0322156416		20	06 2010	1 22.06	2010	24.06	2010	26.06	2010	28.06	2010	30 (	画 06 201	0			
	05.04.2010 14:00	8551	8608,896081743	9456	7651	7327	0,6770679656		20.	00.2010	, 22.00		21.00	2010	20.00	20.0	20.00		50.1	00.201	×			
	05.04.2010 15:00	8117	8115,968781998	8551	6991	7295	-0,0127044228																	
	05.04.2010 16:00	7994	7997,010578986	8117	6950	7731	0,0376604826	1	•															
	05.04.2010 17:00	8830	8325,356269054	7994	7101	8752	-5,7151045407	Load Test D	lata															
	05.04.2010 18:00	9424	9323,467811387	8830	8045	9087	-1.0667677059	Tert																
	05.04.2010 19:00	10291	9377,197691854	9424	8044	9577	-8.8796259658 -	ICSL																
•			m				۱.																	
8739																								
									-				_	_										

Figure 3: Demand Prediction GUI using the DemandForecasting.dll

# 5. Conclusion

The following list summarizes the different steps that were performed in this study with the provided resources:

- Comparison of different time series modelling approaches with SARIMA and SVM based on the consumption data from Berlin östl. Hochstadt. It was found that the SVM framework has better goodness-of –fit.
- Definition of two workflows for the online simulation and look-ahead simulation using "One Step Prediction" and "Pattern Prediction".
- Implementation of a SVM-based demand prediction module in C# using Accord.NET resulting in DemandForecasting.dll
- Implementation of the Demand Prediction GUI for demonstration and testing purpose.

# 6. Referred bibliography

Piller, O, Montalvo, I., Deuerlein, J., Braun, M., & Ung, H. (2014). A gradient-type method for real-time state estimation of water distribution networks. International Conference on Hydroinformatics.

Preis, A., Whittle, A. J., Ostfeld, A., & Perelman, L. (2010). Efficient hydraulic state estimation technique using reduced models of urban water networks. Journal of Water Resources Planning and Management, 137(4), 343-351.

Braun, M., Bernard, T., Piller, O., & Sedehizade, F. (2014): "24-hours demand forecasting based on SARIMA and support vector machines. Water Distribution Systems Analysis (WDSA) 2014