Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 70 (2014) 93 – 102

Procedia Engineering

www.elsevier.com/locate/procedia

12th International Conference on Computing and Control for the Water Industry, CCWI2013

Improving the performance of water demand forecasting models by using weather input

M. Bakker[a,b,]*, H. van Duist[c], K. van Schagen[b], J. Vreeburg[d,e], L. Rietveld[a]

[a]*Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands*
[b]*Royal HaskoningDHV, P.O. Box 1132, 3800 BC, The Netherlands*
[c] *PWN Waterleidingbedrijf Noord-Holland, P.O. Box 2113, 1990 AC, Velserbroek, The Netherlands*
[d]*Wageningen University, P.O. Box 17, 6700 AA Wageningen, The Netherlands*
[e]*KWR Watercycle Research Institute, P.O. Box 1072, 4330 BB Nieuwegein, The Netherlands*

Abstract

Literature shows that water demand forecasting models which use water demand as single input, are capable of generating a fairly accurate forecast. However, at changing weather conditions the forecasting errors are quite large. In this paper three different forecasting models are studied: an Adaptive Heuristic model, a Transfer/-noise model, and a Multiple Linear Regression model. The performance of the models was studied both with and without using weather input, in order to assess the possible performance improvement due to using weather input. Simulations with the models showed that when using weather input the largest forecasting errors can be reduced by 11%, and the average errors by 7%. This reduction is important for the application of the forecasting model for the control of water supply systems and for anomaly detection.

Selection and peer-review under responsibility of the CCWI2013 Committee

Keywords: Demand forecasting; Short term; Weather input; Transfer/-noise model; MLR model

* Corresponding author. Tel.: +31 88 348 2327; fax: +31 88 348 2801.
E-mail address: martijn.bakker@rhdhv.com

1877-7058
Selection and peer-review under responsibility of the CCWI2013 Committee
doi:10.1016/j.proeng.2014.02.012

1. Introduction

There is an on-going trend towards the fully automated centralized operation of water supply systems (Worm et al., 2010; PWN, 2006). When utilities implement centralized automatic control, they aim to reduce costs and at the same time improve the quality of the operations or at least keep the same quality of operations executed by motivated operators. This goal can be achieved by implementing models for the operational control of the systems. Short term water demand forecasting models are an example of those kinds of models, where the forecast of the demand can be used for the optimal overall quantity control or for optimal pump scheduling. Forecasting models are used for this purpose by a number of utilities around the world: In the Netherlands for instance, in 2012 57% of all supplied water is controlled based on a short term water demand forecast (Bakker et al., 2013a). And the penetration of forecasting models is expected to rise over 90% in 2016, due to implementation projects which are executed currently at two large utilities. Other examples of implementations of a control model based on a short term water demand forecasts, are in the United States at four large utilities (Bunn and Reynolds, 2009).

The accuracy of the water demand forecast model is important, to avoid suboptimal control of the system, and to prevent that operators overrule the control settings (necessary or unnecessary) in order to meet all operational conditions (e.g. to avoid a reservoir to overflow or to run empty). Not the average forecast errors but the largest errors –underestimates and overestimates– of the forecasting model play an important role. Large forecast errors might induce undesired or unacceptable adjustments in control, or even in violating the operational conditions in case of limited (over)capacity of the water supply system. Despite the importance of the largest errors, many papers describing water demand forecasting models only report the average performance, expressed as the average error or as the coefficient of determination (R^2). Lertpalangsunti et al. (1999) and Jain et al. (2001) mention the largest errors explicitly, though both do not indicate whether the largest errors are underestimates or overestimates.

For the one day lead demand forecast, various methods have been developed and tested. House-Peters and Chang (2011) and Donkor et al. (2013) present extensive overviews of existing methods. One of the first described methods was based on linear regression of observed values of the daily water demand combined with transfer functions for rainfall and air temperature as independent variables (Maidment et al., 1985). Other papers describe forecasting models, based on the assumption that water demand is made up of base consumption, seasonal consumption, and weather dependent consumption (Zhou et al., 2000, 2002; Gato et al., 2007a, 2007b; Alvisi et al., 2007). In the models, different methods are applied to transfer the independent observations like temperature and rainfall, to forecasted water demand, combined with a persistence component of the observed water demands. The application of Artificial Neural Networks (ANN) is most popular to forecast water demand. Initially, conventional ANN models were applied (Joo et al., 2002; Jain et al., 2001), but as the development of ANN models proceeded, more complex and dynamic ANN models were applied (Ghiassi et al., 2008). Hybrid models which combine ANN methods with other methods like Fuzzy Logic and Fourier Transformations are described in several other papers (Lertpalangsunti et al., 1999; Bárdossy et al., 2009; Odan and Reis, 2012; Adamowski et al., 2012). All water demand forecasting models found in literature are based on mathematical techniques, either conventional regression / transfer functions, or more advanced data driven techniques. The models are often complex and abstract, and hard to understand for operators. This may be a disadvantage, though many operators gain confidence in a model when they experience that it is performing good. So far, no heuristic models based on general observations of water demands have been developed.

Different inputs may be used to generate water demand forecasts. A limited number of papers describe forecasting models which use measured water demand as single input (Jowitt and Chengchao, 1992; Alvisi et al., 2007; Cutore et al., 2008; Caiado, 2010; Bakker et al., 2013b). Most models also use weather information as input, like temperature (Lertpalangsunti et al., 1999; Ghiassi et al., 2008), temperature and precipitation (Maidment et al., 1985; Jain et al., 2001; Bárdossy et al., 2009; Adamowski et al., 2012), and temperature, precipitation, and evaporation, wind speed and/or humidity (Zhou et al., 2002; Joo et al., 2002; Babel and Shinde, 2011). Although the relation between water demand and weather conditions seems obvious, some papers report that the performance of the forecasting model does not improve when using weather inputs (Ghiassi et al., 2008; Odan and Reis, 2012). In some papers, the difference in forecasting accuracy of models with and without using weather input is not reported, which limits the determination of the performance improvement due to using weather input. For an easy and reliable implementation of forecasting models, it is preferred to use measured water demand as single input (Bakker et al., 2013b). The reason for this is that connecting weather data as input to the model, results in extra costs and possible risks (depending on how the input of weather data has been implemented) of missing input for the model. In order to make good decision to use weather input or not in a forecasting model, both the performance improvement of the model and the extra costs and risks need to be considered.

Mostly, the performance of the model is assessed by doing simulations with historic water demands. Generally, the dataset is split into two parts: a part for developing the model and setting the parameters of the model ("training" or "calibration" set); and a part for assessing the model's performance on an independent dataset ("testing" or "validation" set). The data for testing the model should describe all normal variations in the water demand, in order to obtain a reliable performance evaluation of the model. Because water demand varies over the year, and also the variability (or rather the forecastability) varies, the dataset for testing a model should at least be one year and preferably three years. When analyzing three year, it can be avoided to draw conclusions from variations in demand which only occur a single year which might not be representative. However, the data for testing a model was shorter than four months in several papers (Zhou et al., 2002; Jain et al., 2001; Ghiassi et al., 2008; Bárdossy et al., 2009; Babel and Shinde, 2011; Caiado, 2010).

In this paper we compare the performance of an adaptive heuristic forecasting model with two mathematical models for the one day lead water demand forecast. All models are evaluated both with and without using weather inputs, in order to assess the performance improvement of using weather data. The performance of the models are evaluated on water demand data of six different water supply zones in the Netherlands.

2. Methods

2.1. Study areas and data

The same data is used as in Bakker et al. (2013b). The data contains the water demand in six different areas in the Netherlands in the period 2006-2011. We used the data we collected as input in simulations to assess the accuracy of the water demand forecasting models. The weather conditions in the Netherlands in the whole country are more or less similar, and can be characterized as moderate with an average daily maximum temperature in summer of around 19 ºC and in winter of around 3 ºC. For each area, all water flows supplied to the area (from treatment plants, pumping stations and reservoirs) were summed in order to derive the net water demand in the area. Each number in the datasets represents the water consumption by all customers in the area, including water losses in the area. Each dataset consists of the water demand per day in m^3 per day over a period of six years (2,192 values). The models were trained with three years of data (2006-2008), and tested with a subsequent dataset of three years (2009-2011). The characteristics of the areas are shown in Table 1 and Fig. 1.

Table 1. Characteristics of the six investigated areas.

Area	Water utility	Average demand (*m^3 per day*)	# consumers	Type
1. Amsterdam	Waternet	179,800	950,000	urban
2. Rijnregio	Dunea	55,000	305,000	urban / (rural)
3. Almere	Vitens	28,200	193,000	urban
4. Helden	WML	7,100	39,000	rural
5. Valkenburg	WML	1,760	9,200	rural
6. Hulsberg	WML	440	2,400	rural

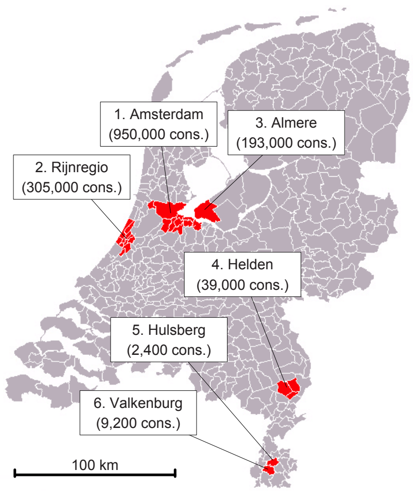

Fig. 1. Location of the six investigated areas in the Netherlands.

2.2. Adaptive heuristic model

Based on the experience of operators observing water demands, we developed a heuristic adaptive water demand forecasting model (Bakker et al., 2013b). The model for forecasting the daily water demand was setup using the following three main observations: 1. The variation in the daily water demand from one day to the next is limited; 2. Subsequent daily water demands describe a weekly pattern; 3. Changes in daily average temperature, result in changes in the water demand. The heuristic model forecasts the water demand for the next day (Q_i^*) primarily based on the measured water demand in the previous two days. In order to correct for the day of the week, the measured water demand on day i (Q_i) is divided by the typical day of the week factor of day i ($f_{dotw,typ,i}$); the water demand is forecasted by applying the day of the week factor to the corrected water demand of the previous two days:

$$Q_i^* = f_{dotw,typ,i} \cdot \left(C_1 \cdot \frac{Q_{i-1}}{f_{dotw,typ,i-1}} + C_2 \cdot \frac{Q_{i-2}}{f_{dotw,typ,i-2}} \right) \qquad [m^3/day] \qquad (1)$$

The model parameters C_1 and C_2 are by default set at 0.8 and 0.2 respectively, making the more recent measured water demands weigh heavier than the older demands. By using equation (1) the forecasted average water demand is based on a relative short period of measured demands (previous two days, with emphasis on the last day). This results in a rapid adjustment of the forecasted water demand, after a change of the measured water demand. The day of the week factor for day type ti ($f_{dotw,typ,ti}$) is adaptively learned by the model, using measurements of the daily water demand in the previous m (default 10) weeks:

$$f_{dotw,typ,ti} = \frac{\frac{1}{m}\sum_{i=1}^{i=m} Q_{\{typ=ti\}}}{\frac{1}{m \cdot 7}\sum_{i=1}^{i=m \cdot 7} Q_i} \qquad [-] \qquad (2)$$

By applying equation (1) in combination with the day of the week factor derived with equation (2), the model automatically corrects for the variation in the daily demands which occur in a weekly pattern. As a result, the lower demands in the weekend are forecasted accurately. In Bakker et al. (2013b), a sensitivity analysis of these model parameters C1, C2 and m is presented, showing optimal results with values of 0.8, 0.2 and 10 respectively.

When a change in the temperature occurs, the model calculates the corrected forecast (Q_i^{**}) by multiplying the original forecast with the temperature correction factor (f_T) and the change in temperature between the forecasted and previous day ($T_i - T_{i-1}$, note that for T_i a forecasted temperature from a weather bureau needs to be used):

$$Q_i^{**} = Q_i^{*} \cdot f_T \cdot \left(T_i - T_{i-1}\right) \qquad [m^3/day] \qquad (3)$$

The temperature correction factor is calculated by the model, by deriving a relation using a least squares fit between the forecast errors (of the uncorrected forecast) and the change in Temperature (see Fig. 2).

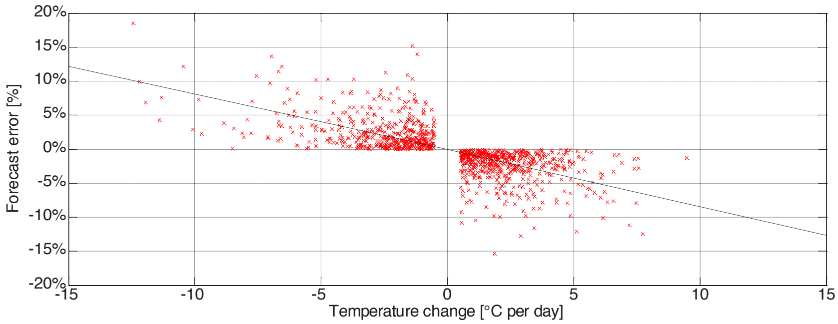

Fig. 2. Forecast error as function of Temperature change. The derivative of the fitted line is the f_T factor, which is derived for both increases and decreases of the Temperature.

2.3. Transfer- /noise model

In modelling a time series, Box and Jenkins (1976) combined the benefits of both an Auto Regressive Integrated Moving Average (ARIMA) model and a transfer model, in a so called transfer-/noise model. The basic assumption in this model is that the output signal of a dynamical system is driven by several signals, including white noise (Fig. 3). These kinds of models have successfully been applied in many scientific areas, including hydrology (Castellano-Méndez et al., 2004) and economics (Grillenzoni, 2000).

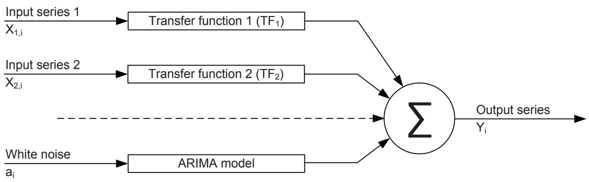

Fig. 3. Setup of transfer-/noise model.

In case that the output series y_i has the same time interval as the input series x_i, the dynamical relation between two series can be modelled with a transfer function of order (r,s):

$$\left(1-\delta_1 B-\delta_2 B^2-\ldots-\delta_r B^r\right)\cdot y_i=\left(\omega_0-\omega_1 B-\omega_2 B^2-\ldots-\omega_s B^s\right)\cdot x_i \qquad (4)$$

where δ_1 to δ_r are the autoregressive parameters, ω_0 to ω_s the moving-average parameters and B the backshift or lag operator ($B \cdot x_i = x_{i-1}$), or simplified $y_i = \omega(B)/\delta(B) \cdot x_i = \mathrm{TF}(B,x_i)$.

The ARIMA model is a special form of the transfer model. The output signal y_i is modelled as a linear function of white noise only. The general form is:

$$\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p\right)\cdot\left(\nabla^d y_i-c\right)=\left(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q\right)\cdot a_i \quad (5)$$

where ϕ_1 to ϕ_r are the autoregressive parameters, θ_1 to θ_q the moving-average parameters, a the white noise, ∇ the difference operator ($\nabla y_i = y_i - y_{i-1} = (1-B)\cdot y_i$), d the number of differences ($\nabla^d y_i = y_i - y_{i-d} = (1-B^d)\cdot y_i$) and c a constant. The transfer-/noise model is generated by linear superposition of the transfer function(s) and the noise function. The role of the latter is to account for any influences on the system output, which are not covered by any of the input series $x_{n,i}$. In case of two inputs, the relation between output y_i and the inputs $x_{n,i}$ becomes:

$$y_i = TF_1(x_{1,i}) + TF_2(x_{2,i}) + N_i \quad (6)$$

where TF_n is the Transfer Function and N_i is the white noise component. When using the Box-Jenkins methodology to forecast water demand, a number of modifications were applied. Firstly, only the Moving-Average parameters turned out to be valuable (no significant δ parameters were found during the model building phase, and therefore not used in the model). Secondly, water demand is not a stationary series but usually shows a weekly modulation. Therefore, we differentiated the original demand series (Q_i) twice, and used the resulting output series y_i as the output of the transfer-/noise model:

$$y_i = \nabla\nabla^7 Q_i = \left(Q_i - Q_{i-1}\right) - (Q_{i-7} - Q_{i-8}) \quad (7)$$

Like in the heuristic model, the daily average temperature is used as an independent variable to forecast the water demand. Temperature only affects water demand during the spring and summer (average temperature higher than 10 ºC), and the effect increases as temperature increases. By trial and error, we found optimal forecasting results when using a transformed value (z_i) of the original temperature values (T_i) with:

$$z_i = 0.001\cdot\left(T_i - 10\right)^3 \;\; if\; T_i > 10 \qquad Else\; z_i = 0 \quad (8)$$

To compensate for special days with abhorrent water demand (like national holidays), a dummy input series F_i was created with values of -1 on special days with expected lower demand and +1 for expected higher demands. Both abovementioned input series (transformed temperature z_i and the intervention series for special days F_i) were differentiated in the same way as the demand series (equation (7)). The noise model (N_i) was kept simple, by constructing it as an additive model with 3 moving average parameters:

$$N_i = \left(a_i - \theta_1\cdot a_{i-1}\right) - \left(\theta_7\cdot a_{i-7} - \theta_8\cdot a_{i-8}\right) \quad (9)$$

Assuming the white noise a_i is 0, the resulting forecast model is formulated as:

$$\begin{aligned} Q^*_i = \quad & Q_{i-1} + \left(Q_{i-7} - Q_{i-8}\right) + \\ & \omega_{1,0}\left(z_i^* - z_{i-1} - z_{i-7} + z_{i-8}\right) + \omega_{1,1}\left(z_{i-1} - z_{i-2} - z_{i-8} + z_{i-9}\right) + \\ & \omega_{2,0}\left(F_i - F_{i-1} - F_{i-7} + F_{i-8}\right) + \omega_{2,1}\left(F_{i-1} - F_{i-2} - F_{i-8} + F_{i-9}\right) + \\ & -\theta_1 a_{i-1} - \theta_7 a_{i-7} + \theta_8 a_{i-8} \end{aligned} \quad (10)$$

Where Q^*_i is the forecasted demand and z^*_i is the transferred forecasted average Temperature, obtained from a weather institute. The unknown parameters (θ_1, θ_7, θ_8, $\omega_{1,0}$, $\omega_{1,1}$, $\omega_{2,0}$, $\omega_{2,1}$) were estimated using the training dataset of Q, T and F values.

2.4. Multiple Linear Regression (MLR) model

When omitting the noise component, the transfer-/noise model is reduced to a simplified Multiple Linear Regression model. Experiments showed that in general the regression parameters z_{i-7} and z_{i-8} were not significant and we settled for the following forecast model:

$$Q^*_i = Q_{i-1} + Q_{i-7} - Q_{i-8} + \beta_1 z_i + \beta_2 z_{i-1} + \beta_3 F_i + \beta_4 F_{i-1} + \beta_5 F_{i-7} + \beta_6 F_{i-8} \tag{11}$$

where the Q_i, z_i and F_i are the same as in the transfer-/noise model. The β_n were estimated using the training dataset with the Least Squares Algorithm.

2.5. Model performance evaluation

The performance of the different models and the performance improvement by using weather inputs was assessed by comparing the forecasted to the measured values. As stated before, not only the average errors are important, but also the largest errors (underestimates and overestimates). For evaluating the models we therefore chose the following evaluation parameters: The 0.5% and 99.5% confidence intervals of the Relative Error RE_i ($RE_{0.5\%}$ and $RE_{95.5\%}$); The Mean Absolute Percentage Error ($MAPE$), and The Nash-Sutcliffe Model Efficiency (R^2):

$$RE_i = \frac{(\hat{y}_i - y_i)}{\overline{y}} \cdot 100\% \tag{12}$$

$$MAPE = \frac{\frac{1}{n}\sum_{i=1}^{n} |\hat{y} - y_i|}{\overline{y}} \cdot 100\% \tag{13}$$

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{14}$$

Where $\hat{y}_i$ is the forecasted value, y_i is the measured value, and $\overline{y}$ is mean of the measured values. The 0.5% and 99.5% confidence intervals represent the underestimates and overestimates which are statistically exceeded approximately 2 days in every year. We chose this measure rather than the absolute largest underestimates and overestimates, because those might be related to isolated anomalies (e.g. caused by pipe bursts of fire fighting) which are not representative for the performance of the model. By calculating the abovementioned confidence intervals, the evaluation is not influenced by single observations, though still indicating properly the largest errors that may be expected.

3. Results and discussion

In Table 2 to Table 4 the results of the three investigated forecasting methods, both without and with using weather input, are shown.

Table 2. Results heuristic water demand forecasting model (testing period: 2009-2011).

Area	Without weather input				With weather input			
	MAPE	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$	*MAPE*	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$
1. Amsterdam	1.39%	0.756	-6.54%	6.97%	1.32%	0.790	-5.68%	6.33%
2. Rijnregio	1.88%	0.694	-8.90%	9.81%	1.73%	0.751	-8.90%	7.90%
3. Almere	2.08%	0.718	-10.96%	10.81%	1.97%	0.764	-9.41%	8.43%
4. Helden	3.72%	0.794	-24.46%	17.30%	3.33%	0.848	-19.45%	14.08%
5. Valkenburg	3.55%	0.788	-16.54%	15.12%	3.38%	0.808	-16.12%	14.48%
6. Hulsberg	5.08%	0.688	-28.73%	21.02%	4.48%	0.755	-22.49%	20.68%

Table 3. Results Transfer-/noise water demand forecasting model (testing period: 2009-2011).

Area	Without weather input				With weather input			
	MAPE	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$	*MAPE*	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$
1. Amsterdam	1.36%	0.766	-6.25%	5.88%	1.25%	0.812	-5.96%	5.23%
2. Rijnregio	1.77%	0.740	-8.55%	8.68%	1.65%	0.777	-7.38%	8.24%
3. Almere	2.03%	0.733	-9.56%	10.91%	1.87%	0.793	-8.23%	9.61%
4. Helden	3.74%	0.791	-21.57%	20.37%	3.47%	0.832	-19.40%	18.63%
5. Valkenburg	3.44%	0.807	-15.76%	14.70%	3.32%	0.818	-16.26%	14.16%
6. Hulsberg	5.04%	0.696	-26.07%	25.92%	4.77%	0.727	-22.73%	23.15%

Table 4. Results MLR water demand forecasting model (testing period: 2009-2011).

Area	Without weather input				With weather input			
	MAPE	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$	*MAPE*	R^2	$RE_{0.5\%}$	$RE_{95.5\%}$
1. Amsterdam	1.54%	0.709	-7.38%	6.28%	1.47%	0.738	-6.41%	5.98%
2. Rijnregio	1.99%	0.681	-8.32%	10.47%	1.87%	0.731	-7.43%	9.48%
3. Almere	2.37%	0.681	-10.95%	10.66%	2.26%	0.724	-10.31%	8.73%
4. Helden	4.25%	0.747	-23.21%	20.39%	4.01%	0.780	-22.36%	18.73%
5. Valkenburg	3.97%	0.756	-20.29%	15.01%	3.81%	0.772	-18.33%	14.64%
6. Hulsberg	5.97%	0.619	-29.22%	26.19%	5.73%	0.643	-28.82%	24.68%

The forecast errors of both the heuristic and the transfer/-noise model were 10-15% smaller than the MLR model. This indicates that water demand cannot be described properly with regression formulas only. The tables show comparable performance of the heuristic model and the transfer/-noise model (see Table 2 and Table 3), with on average a slightly better performance of the transfer/-noise model. When using weather input, the forecast errors were smaller for all models in all areas. The performance improvement by using weather input amounted 7% on average for the *MAPE*, and 11% on average of the $RE_{0.5\%}$ and the $RE_{99.5\%}$. This indicates that using weather input is helpful in reducing the maximum errors of forecasting models.

The tables show increasing forecast errors with decreasing water demand in the areas. The errors in the smallest area (6. Hulsberg) are on average 400% larger than the errors in the largest area (1. Amsterdam). This indicates that the randomness in the demand is larger in smaller areas, and that therefore the demand cannot be forecasted as accurate as in larger areas.

The R^2 value indicates how well the model is able to describe the variation in the data. The R^2 values for the different areas are in the same range, varying between 0.7 and 0.85 for the heuristic and transfer/-noise model. This shows that the forecasting models perform comparable on the different datasets. The higher *MAPE* values in combination with the comparable R^2 values in the smaller areas, indicate that the variability of the water demands in those areas is larger, but that the models are able to describe those variations to a comparable degree. Based on this observation, we conclude that the models can be applied to forecast the water demand in smaller areas, but that larger forecast errors need to be taken into account. In Bakker et al. (2013b) a relation between the expected forecast accuracy and the average water demand in an area is presented.

4. Conclusions

Simulations with six different sets of water demands, showed that a heuristic model and a transfer/-noise model outperformed a Multiple Linear Regression model when forecasting the one day lead water demand. The transfer/-noise model performed somewhat better than the heuristic model. The heuristic model is easier to understand for operators than the mathematical transfer/-noise model. In the heuristic model, the adaptively learned factors which influence the forecast can be presented to the operators, which allows operators to verify the validity of the factors and to understand how future demands will be forecasted.

When using weather input, the performance of the forecasting models can be improved by 7% with respect the average errors, and 11% with respect to the largest errors. This improvement can be relevant when higher forecasting accuracies are necessary when using the forecasts for optimal control or for anomaly detection. In each implementation of a forecasting model, the evaluation can be made whether the higher costs and complexity of using weather input weighs up to the higher forecasting accuracy.

Acknowledgements

This study was carried out in the DisConTO project (Distribution Control Training & Operation). The project is a cooperation between 4 water supply companies (Vitens, Dunea, PWN and Brabant Water), Delft University of Technology, The National Institute for Public Health and the Environment (RIVM), Royal HaskoningDHV Consultancy and Engineering and UReason. The project is financially supported by the Dutch government through the “Innowater” program.

References

Adamowski, J., Fung Chan, H., Prasher, S.O., Ozga-Zielinski, B., and Sliusarieva, A. (2012)."Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada". *Water Resources Research.* 48 (1).

Alvisi, S., Franchini, M., and Marinelli, A. (2007)."A short-term, pattern-based model for water-demand forecasting". *Journal of Hydroinformatics.* 9 (1): 39-50.

Babel, M.S., and Shinde, V.R. (2011)."Identifying Prominent Explanatory Variables for Water Demand Prediction Using Artificial Neural Networks: A Case Study of Bangkok". *Water Resources Management.* 25 (6): 1653-1676.

Bakker, M., Vreeburg, J.H.G., Palmen, L.J., Sperber, V., Bakker, G., and Rietveld, L.C. (2013a)."Better water quality and higher energy efficiency by using model predictive flow control at water supply systems". *Journal of Water Supply: Research and Technology - AQUA.* 62 (1): 1-13.

Bakker, M., Vreeburg, J.H.G., Van Schagen, K.M., and Rietveld, L.C. (2013b). (submitted). "A fully adaptive forecasting model for short-term drinking water demand ". *Journal of Environmental Modelling and Software.*

Bárdossy, G., Halász, G., and Winter, J. (2009)."Prognosis of urban water consumption using hybrid fuzzy algorithms". *Journal of Water Supply: Research and Technology - AQUA.* 58 (3): 203-211.

Box, G.E.P., and Jenkins, G.M. 1976. *Time Series Analysis: Forecasting and Control*. 2 nd ed. San Francisco: Holden-Day.

Bunn, S.M., and Reynolds, L. (2009)."The energy-efficiency benefits of pumpscheduling optimization for potable water supplies". *IBM Journal of Research and Development.* 53 (3).

Caiado, J. (2010)."Performance of combined double seasonal univariate time series models for forecasting water demand". *Journal of Hydrologic Engineering.* 15 (3): 215-222.

Castellano-Méndez, M., González-Manteiga, W., Febrero-Bande, M., Manuel Prada-Sánchez, J., and Lozano-Calderón, R. (2004)."Modelling of the monthly and daily behaviour of the runoff of the Xallas river using Box-Jenkins and neural networks methods". *Journal of Hydrology.* 296 (1-4): 38-58.

Cutore, P., Campisano, A., Kapelan, Z., Modica, C., and Savic, D. (2008)."Probabilistic prediction of urban water consumption using the SCEM-UA algorithm". *Urban Water Journal.* 5 (2): 125-132.

Donkor, E., Mazzuchi, T., Soyer, R., and Roberson, J. (2013). (in press). "Urban Water Demand Forecasting: A Review of Methods and Models". *Journal of Water Resources Planning and Management.*

Gato, S., Jayasuriya, N., and Roberts, P. (2007a)."Temperature and rainfall thresholds for base use urban water demand modelling". *Journal of Hydrology.* 337 (3-4): 364-376.

Gato, S., Jayasuriya, N., and Roberts, P. (2007b)."Forecasting residential water demand: Case study". *Journal of Water Resources Planning and Management.* 133 (4): 309-319.

Ghiassi, G., Zimbra, D.K., and Saidane, H. (2008)."Urban water demand forecasting with a dynamic artificial neural network model". *Journal of Water Resources Planning and Management.* 134 (2): 138-146.

Grillenzoni, C. (2000)."Time-varying parameters prediction". *Annals of the Institute of Statistical Mathematics.* 52 (1): 108-122.

House-Peters, L.A., and Chang, H. (2011)."Urban water demand modeling: Review of concepts, methods, and organizing principles". *Water Resources Research.* 47 (5).

Jain, A., Varshney, A.K., and Joshi, U.C. (2001)."Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks". *Water Resources Management.* 15 (5): 299-321.

Joo, C.N., Koo, J.Y., and Yu, M.J. (2002)."Application of short-term water demand prediction model to Seoul". *Water Science and Technology* 46 (6-7): 255-261.

Jowitt, P.W., and Chengchao, X. (1992)."Demand forecasting for water distribution systems". *Civil Engineering Systems.* 9 (2): 105-121.

Lertpalangsunti, N., Chan, C.W., Mason, R., and Tontiwachwuthikul, P. (1999)."Toolset for construction of hybrid intelligent forecasting systems: Application for water demand prediction". *Artificial Intelligence in Engineering.* 13 (1): 21-42.

Maidment, D.R., Miaou, S.-P., and Crawford, M.M. (1985)."Transfer function models of daily urban water use". *Water Resources Research.* 21 (4): 425-432.

Odan, F.K., and Reis, L.F.R. (2012)."Hybrid Water Demand Forecasting Model Associating Artificial Neural Network with Fourier Series". *Journal of Water Resources Planning and Management.* 138 (3): 245-256.

PWN. (2006)."Automated Water System Control". *Control Engineering.* 53 (10).

Worm, G.I.M., van der Helm, A.W.C., Lapikas, T., van Schagen, K.M., and Rietveld, L.C. (2010)."Integration of models, data management, interfaces and training support in a drinking water treatment plant simulator". *Environmental Modelling and Software.* 25 (5): 677-683.

Zhou, S.L., McMahon, T.A., Walton, A., and Lewis, J. (2000)."Forecasting daily urban water demand: A case study of Melbourne". *Journal of Hydrology.* 236 (3-4): 153-164.

Zhou, S.L., McMahon, T.A., Walton, A., and Lewis, J. (2002)."Forecasting operational demand for an urban water supply zone". *Journal of Hydrology.* 259 (1-4): 189-202.