International Conference on Hydroinformatics

8-1-2014

Real-Time Runoff Prediction Based On Data Assimilation And Model Bias Reduction

Nemanja Branisavljevic

Christopher Hutton

Zoran Kapelan

Lydia Vamvakeridou-Lyroudia

Dragan A. Savić

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

REAL-TIME SEWER FLOW PREDICTION BASED ON DATA ASSIMILATION AND BIAS REDUCTION

NEMANJA BRANISLAVJEVIĆ (1,2), CHRISTOPHER HUTTON (1,3), ZORAN KAPELAN (1), LYDIA VAMVAKERIDOU-LYROUDIA (1), DRAGAN SAVIĆ (1)

(1) Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Harrison Building, North Park Road, Exeter, EX4 4QF

(2) Now at: University of Belgrade, Faculty of Civil Engineering, Belgrade, Serbia

(3) Now at: Water and Environmental Management Research Centre, Department of Civil Engineering, Queen's School of Engineering, University of Bristol, Bristol, UK.

The increased availability of (near) real-time observations in urban sewer systems provides the potential for control of actuators, such as gates, in response to current system states. Predictive models can be applied to inform such control, however the predicted states of such models can diverge from the true states of the system, leading to local bias in the model predictions because of errors in the model, and in the input drivers (e.g. rainfall) Data Assimilation (DA) methods can incorporate (near) real-time observations to improve real-time model performance; however the success of their application is complicated by the presence of the very control structures that such models may be used to optimize. Although DA methods may be used to improve the states of the system that are further used to improve other model predictions (downstream) in general, such model predictions can still contain bias, even after correction. This paper investigates the application of a Data Assimilation technique - the Extended Kalman Filter (EKF) - to assimilate observations into a conceptual urban rainfall runoff model of a real-life system that contains control structures. The EKF state correction is extended with an additional stage whereby a bias correction extension is added to detect and reduce local model bias at the downstream flow prediction. The proposed method is tested and verified using an urban rain-runoff model for a town in Northern Europe. In comparison to model application without data assimilation, the EKF led to improved estimates of the model states using the observations from the observed location (retention basin), and therefore better estimation of the downstream water flow that is used to control gate operation. The additional local bias correction method led to further improvements in predictive performance, and in some cases the bias reduction provided better results when applied singly, without data assimilation.

INTRODUCTION

Near real-time observations in urban sewer systems, such as observations of water level and pipe flow rate, provide data that may be used to inform the control of actuators such as gates in response to current system states. Moreover, identifying a near optimal control strategy can be facilitated by the application of a numerical model of the system, when coupled with an optimization algorithm [2,15]. Such operation may help to militate against the impacts of surface flooding and Combined Sewerage Overflow (CSO) [18]. Although models have shown some potential in facilitating the identification of real-time control strategies, a number of sources of uncertainty affect the performance of urban rainfall-runoff models, including uncertainties in model structure, input rainfall, and uncertainties in model parameters [5, 13]. The combined effect of these uncertainties is to reduce the accuracy of model predictions, which undermines the ability of numerical models to support real-time system control in a robust manner. A number of techniques may be applied to reduce uncertainties in off-line calibration [9, 6], notably in model parameter values. However, such off-line techniques cannot reduce the effect of uncertainties on real-time modelling, which lead to divergence between what the model predicts, and the true system states. Data Assimilation (DA) refers to a range of techniques which combine model predictions with real-time observations to update the states of the model and derive an estimate of the current system states [10]. Such techniques include ensemble approaches, including the Ensemble Kalman Filter and Particle Filter [12, 8]. Deterministic Data Assimilation methods include MIKE UPDATE [11], which adjusts water levels in the sewer system using local observations, and a method developed by *Borup et al* [1], which applied downstream measurements to update the states of the slow changing flow components upstream in the model.

A key issue when applying DA techniques is the method by which observational information is mapped to state space to update the states of the model. In runoff systems, as long as the potential time-lag in the system is accounted for between upstream states and downstream observations [9], the covariance may provide an adequate mapping. However, sewer systems contain control structures, which when activated may act to de-couple upstream from downstream states. The result is that the 'information' that propagates from upstream to the observation location that may be used to correct the state of the system is lost, hindering the potential application of DA methods to the very systems that improved models for real-time system control can be applied. This issue was previously investigated for a synthetic case study [2], which demonstrated that storage tanks, by de-coupling upstream and downstream states in the system, can hinder the application of DA at certain points of a simulation. This paper extends this work through the application of the Extended Kalman Filter (EKF) to assimilate observations in a real-life urban rainfall-runoff system.

Beside random errors that the traditional data assimilation and real time calibration methods effectively handles, a non-zero mean prediction error, also known as bias is often present in prediction results. Although it is well known that bias in model predictions significantly contributes to model inaccuracy, the problem of properly handling it has received little attention so far. Nevertheless, several error reduction methods that are related to the problem of existing bias in the model predictions, have been developed and applied to some water related problems. In [11] the authors applied a simplified DA methodology for river flow prediction and coupled this to a simple error correction procedure, which took the form of a linear or trigonometric error model. The error model significantly improved the results of the predictions by reducing the local bias introduced by the tidal wave. In [17] a DA procedure was applied not on the state-space model of hydrologic process, but on its linear error prediction model. That way the errors were reduced instead the discrepancy of the predicted and real states of the system. In *Vojinovic et al.* [19] the authors applied radial based ANN to model the bias of the sewage flow model and obtained promising results, while in [4] authors compared several bias models (constant, input dependent and output dependent) combined with some data transformation methods, and gave some recommendations how to handle bias in hydrologic models.

Guided by the success of the previous research in this field we introduced the extension to the existing data assimilation procedure in the form of local bias correction utilising a exponential error model.

METHODOLOGY

The prediction methodology is divided in two steps: first, data assimilation using EKF, and second error correction of the most recent data. The basic Kalman Filter provides an optimal system estimate in linear systems with Gaussian errors. The EKF is an extension of the KF that may be better applied to nonlinear systems, and may be considered as a predictor-corrector scheme. A single model, or estimate of the states of the system x is propagated forwards in time alongside the state error covariance matrix P. Once the observations are obtained, both the state vector and error covariance matrix are updated using the Kalman gain, which updates the model states considering the relative error between the model states and the observations (for further details see [16]).

In the second step, bias correction is performed by correcting the predicted value using an exponential term with two parameters - Δy - linear transformation parameter, α - stretch parameter:

$$y_{M,corr}^{i} = y_{M}^{i} + \Delta y \times e^{\alpha \times y_{M}^{i}} \tag{1}$$

where $y_{M,corr}^{i}$ is corrected predicted value and y_{M}^{i} is predicted value before correction. The parameters Δy and α are determined by minimising the mean square error between the observed and the predicted value for the last n predicted and observed values:

$$\underset{\Delta y,\alpha \in R}{\arg\min} \left\{ \sum_{i=i_0}^{i_0+n} \left[y_{O}^{i} - \left(y_{M}^{i} + \Delta y \times e^{\alpha \times y_{M}^{i}} \right) \right]^{2} \right\} \tag{2}$$

The number n is determined to be the number of the last predicted values with bias of the same sign, and i determines the time index in the sets of last n predicted and observed values.

CASE STUDY

In order to apply and test the developed methodology based on DA and bias/error correction, a coupled conceptual-physical model of an urban rainfall-runoff system for a city in Northern European was developed. The model consists of a conceptual hydrological representation of the processes in the upper catchments, coupled with a physical hydraulic representation of the main trunk sewer and control structures.

A conceptual diagram of the system analysed is depicted in Figure 1. The catchment S_{TB}, which is the most upstream from the rest of the system, consists of four sub-catchments and a storage tank/basin (TB on Figure 1), to which flow is diverted via a number of weirs during high flows. The control structures, gates and pumps are used for water control in the system, and while the gates are used to divert water to the retention basins during the extreme rain events, the pumps are used to empty them during low flows.

The system is controlled to minimize flow to the catchment outlet, and subsequently to a waste-water treatment plant, and also to minimize the risk of CSO discharge, via the tanks. The hydraulic system of the sewage is loaded from four sub-catchments with different areas and different percentage of pervious and impervious covers.

The sub-catchments in the system (S_1, S_2, S_4 and S_5) are modelled using a conceptual linear reservoir approach, while the flow through the sewer system including the overflow over the weirs and flow under the gates is modelled using physical hydraulic equations, and a Muskingum approach [3] applied for flow routing within the main sewer between S_{TB} and the next sub-catchment downstream.

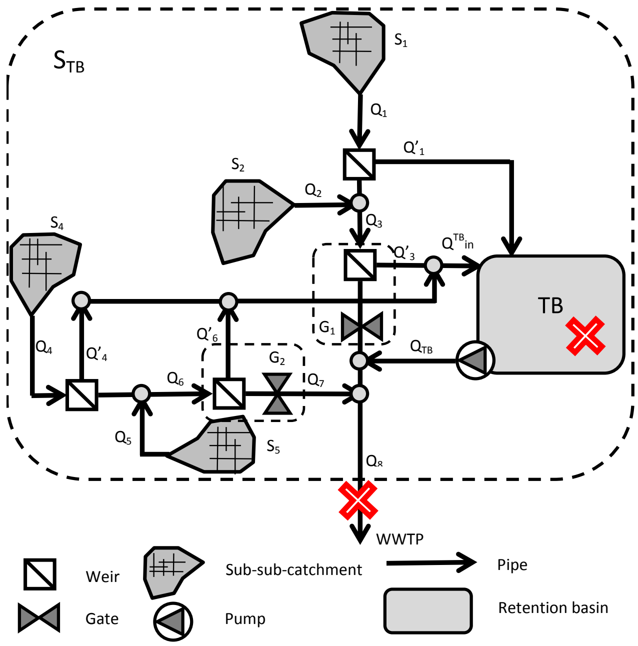

Figure 1. Case study sewage system with observation locations indicated as crosses

The sub-catchments are modelled with two linear reservoirs in parallel; the first is used for modelling overland flow and the latter for modelling flow in the sewage system in the sub-catchments. The runoff is considered only from impervious areas with the specified loss function that was used to model the evaporation and other losses of the runoff water. Since the main purpose of the model is to provide predictions during rain events, especially when the controls and retention structures are operational (e.g. larger flows), base flow is modelled as constant flow through the whole simulation, equal to the mean diurnal flow calculated from historical data.

The weirs in the system are divided into two groups, static weirs and weirs with gates. Static weirs are modelled to overflow when the water level in the pipe reaches their crest. The weirs with gates are modelled with the orifice equation combined with the weir equation(s) accompanied with the water balance equation, solved using Newton-Raphson method.

Water levels in TB retention basin and the flows in the most downstream pipe (P_8) are observed (crosses on Figure 1). The process of DA was performed in two ways: 1) the levels in the retention basin are used and 2) the flows in pipe P8 are used as observed values to update the states in linear reservoirs of sub-catchments. The historical observed values from those locations were also used for model calibration. Other data, than those used for calibration were used in the DA process.

RESULTS AND DISCUSSION

The EKF is applied to assimilate observations of retention basin water level to update the states of the four linear reservoirs representing the upstream sub-catchments in the system for a real rainfall runoff event with sufficient size to trigger activation of TB retention tank. The tank level start rising (Figure 2) when the water level exceeds the static weir levels and continues to rise as gates G_1 and G_2 (Figure 1) close. G_1 and G_2 re-open and close again during the simulation, which explains the plateaus observed at 17:00. Gates G_1 and G_2 finally re-open around 21:30 when the pumps are also started to empty the tanks. It may be noticed that for the most of the simulation time, without DA, the model consistently either over-predicts or under-predicts the water level in the tank thus demonstrating the local (conditional) prediction bias. Applying the model with EKF, which is applied to update the upstream linear reservoirs in the system, reduces the model water level prediction error in the tank. Additional local bias correction, which incorporates information on recent prediction bias, leads to further improvements in predictive performance.

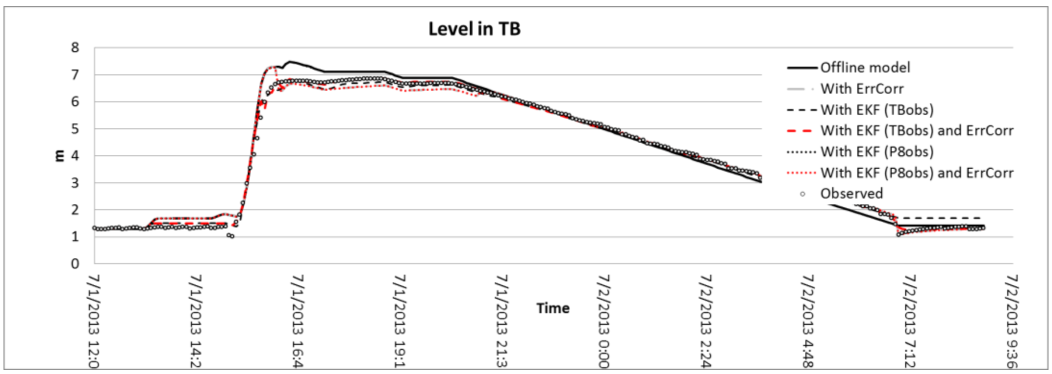

Figure 2. Water level in retention basin TB

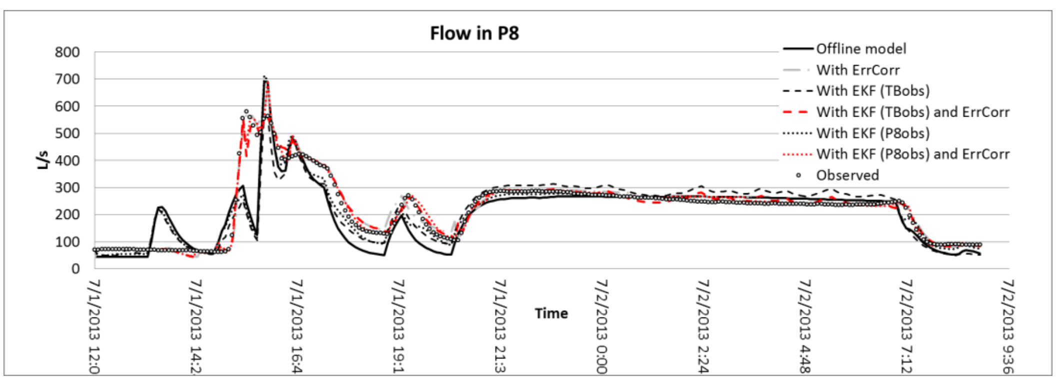

Figure 3. Flows in pipe P8

Figure 3 shows the results of the model simulated flows with and without DA at the most downstream pipe (P8 on Figure 1). The model sometimes underestimates and sometimes overestimates the flows at the downstream end of the model reach. The application of the EKF to assimilate observations into the model reduces the effect of this over/under-prediction to a certain point. The principal effect of the assimilation, to reduce the error in peak runoff magnitude, is partly met due to the limited amount of water removed from the TB retention basin. The over-prediction at certain time steps seems to result from errors in the input rainfall, which are derived from rainfall radar. Much of the downstream runoff signal will be reflected in the input rainfall. However, there appear to be peaks in the rainfall that result in simulated discharge peaks that do not occur in the observations at the downstream end of the reach. This could also result from the smoothing and attenuating effect of the storage tanks.

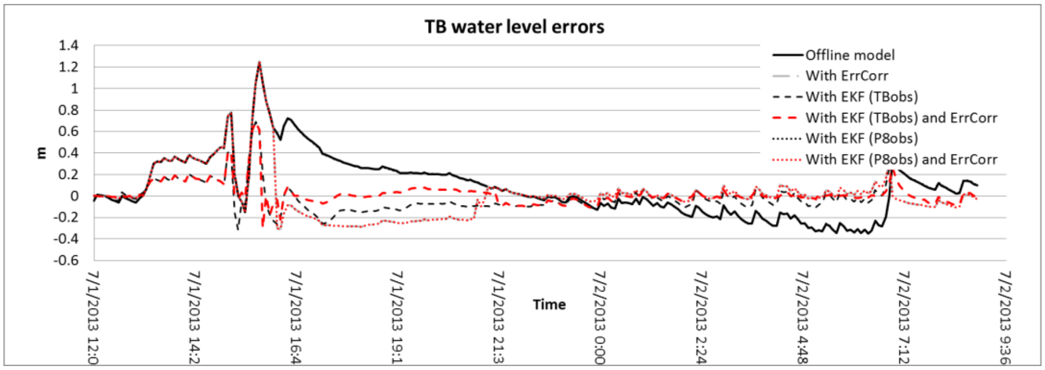

Figure 4. Errors of predicted levels in the TB retention basin

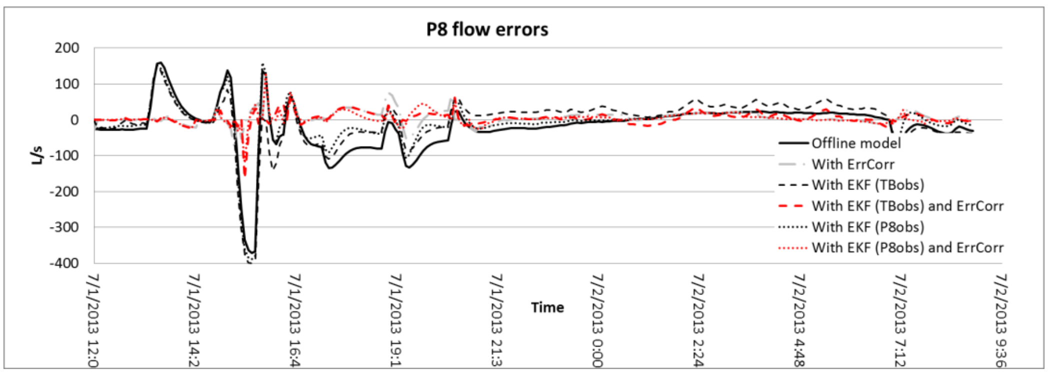

Figure 5. Errors of predicted flows in pipe P8

When the bias correction method is applied, the predictions improved significantly. Comparing the predictions with bias correction only and both data assimilation and bias correction without data assimilation it can be seen that in both cases the predictions were improved. Additionally for some data values, the prediction got worst that in the case when the bias correction was applied singly. These points were mostly when the data assimilation increased the discrepancy between the predicted and the observed data.

Figure 4 shows the errors of the level prediction in the TB retention basin. It may be noticed that the high error rates during filling of the basin were reduced only using the EKF data assimilation procedure that may lead to poor estimation of potential CSOs. Nevertheless, overestimating the CSOs may lead to more precautious operation by the system, but more accurate estimations are desired since most of the system performance may be used when the downstream volumes are exploited to the full capacity.

Figure 5 presents errors in the prediction of pipe P8 flows. As it can be seen from Figure 5, the bias/error correction procedure reduces these errors to a large extent.

Table 1. Nash–Sutcliffe and RMS statistics

	P8				TB			
	Offline model	**With EKF**	**With ErrCorr**	**With EKF and ErrCorr**	**Offline model**	**With EKF**	**With ErrCorr**	**With EKF and ErrCorr**
NS	-51.96	-43.44	-3.63	-1.83	-2.09	0.20	0.58	0.58
RMS	69.42	63.59	20.53	16.05	0.29	0.18	0.11	0.11

Table 1 presents the model prediction statistics in the form of Nash Sutcliffe [14] coefficient and Root Mean Square error coefficient. In case of the flow in pipe P8, application of the EKF and bias correction method are compared when the flow in P8 is used in the DA procedure, and in the case of water level in TB, the corresponding levels in TB are used in the EKF procedure.

It may be seen that the NS value approaches one when both DA and error correction are used and that the RMS values are also reduced significantly. It may be noticed that in the case of the level in TB basin, the EKF applied before error correction didn't significantly changed the improvement of the prediction but although it introduced significant improvement applied by singly.

CONCLUSION

This paper has demonstrated that the application of DA techniques, specifically the EKF, can lead to improved runoff predictions when applied to assimilate observations into a conceptual rainfall-runoff model of a real urban rainfall-runoff system. More specifically, it was shown that, assimilating one observation variable for sub-catchment state updates the affects of other model states, leading to more robust model predictions. Even further improvement of model predictions are accomplished with bias reduction method applied with or without DA. When the input data are available in advance, only one simulation model run is required to get reasonably acceptable results as shown by the examples in this paper. The main drawback of presented bias/error reduction method is the fact the model modification is purely geometrical and that it doesn't transfer any information about the modelling process to the result. This method showed promising results even with the abruptly changeable flows what is the case in the controlled sewer system. Also, one of the main advantages of the algorithm is its ability to be applied in real-time due its fast performance.

REFERENCES

[1] Borup, M., Grum, M., and Mikkelsen, P. S. "Real time adjustment of slow changing flow components in distributed urban runoff models". *12th International Conference on Urban Drainage*, Porto Alegre, Brasil. (2011).

[2] Butler, D., and Schutze, M., "Integrating simulation models with a view to optimal control of urban wastewater systems", *Environmental Modelling & Software*, Vol. 204., No.1, (2005), pp 415-426.

[3] Cunge, J.A.,: "On the Subject of a Flood Propagation Computation Method (Muskingum Method)", *Journal of Hydraulic Research*, Vol. 7, (1969)

[4] Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., Reichert, P. and Rieckermann, J.: "Improving uncertainty estimation in urban hydrological modeling by statistically describing bias", *Hydrol. Earth Syst. Sci*., Vol. 17, (2013), pp 4209-4225

[5] Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T. D., Rauch, W., Bertrand-Krajewski, J. L., and Tait, S., "Assessing uncertainties in urban drainage models", *Physical and Chemistry of the Earth* Vol. 42–44, (2012), pp 3-10.

[6] Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and Deletic, A. "Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling" *Water Research,* Vol. 46, (2012), pp.2545-2558

[7] Hansen, L. S., Borup, M., Moller, A., and Mikkelsen, P. S. "Flow Forecasting in Urban Drainage Systems using Deterministic Updating of Water Levels in Distributed Hydraulic Models". *12th International Conference on Urban Drainage*, Porto Alegre, Brasil.

[8] Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L., & Savic, D. A. "Real-time demand estimation in water distrubtion systems under uncertainty". In*WDSA 2012: 14th Water Distribution Systems Analysis Conference, 24-27 September 2012 in Adelaide, South Australia* (p. 1374). Engineers Australia.

[9] Hutton, C., Kapelan, Z., Vamvakeridou-Lyroudia, L., and Savić, D. "The Application of Formal and Informal Bayesian Methods for Water Distribution Hydraulic Model Calibration." *J. Water Resour. Plann. Manage.,* 10.1061/(ASCE)WR.1943-5452.0000412 (Oct. 3, 2013).

[10] Hutton, C., Kapelan, Z., Vamvakeridou-Lyroudia, L., and Savić, D. "Dealing with Uncertainty in Water Distribution System Models: A Framework for Real-Time Modeling and Data Assimilation." *J. Water Resour. Plann. Manage*., Vol. 140, (2014), pp169–183.

[11] Madsen, H. and Skotner C. "Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure", *Journal of Hydrology*, Vol. 308, (2005), pp 302-312.

[12] McMillan, H. K., Hreinsson, E. Ö., Clark, M. P., Singh, S. K., Zammit, C., and Uddstrom, M. J. "Operational hydrological data assimilation with the recursive ensemble Kalman filter", *Hydrol. Earth Syst. Sci.*, Vol. 17, (2013) pp21-38.

[13] Molini A, Lanza LG, Barbera PL. "The impact of tipping-bucket rain gauge measurement errors on design rainfall for urban-scale applications", *Hydrological processes*, Vol.19, (2005), pp1073–1088.

[14] Nash, J. and Sutcliffe, J.: "River flow forecasting through conceptual models part I : A discussion of principles*", J. Hydrol*., Vol. 10, (1970), pp 282–290.

[15] Puig, V., Cembrano, G., Romera, J., Quevedo, J., Aznar, B., Ramon, G., and Cabot, J., "Predictive optimal control of sewer networks using CORAL tool: application to Riera Blanca catchment in Barcelona", *Water Science and Technology*, Vol.60, No.4, (2009), pp 869-878.

[16] Reichle, R. H., Walker, J. P., Koster, R. D., & Houser, P. R. "Extended versus ensemble Kalman filtering for land data assimilation" *Journal of hydrometeorology*, Vol. *3,* (2002), pp 728–740.

[17] Smith, P. J., Beven, K., Weerts, A., and Leedal, D.: "Adaptive correction of deterministic models to produce accurate probabilistic forecasts", *Hydrol. Earth Syst. Sci. Discuss*., Vol. 9, pp 595-627.

[18] Vezzaro, L., Lowe, R., Madsen, H., Grum, M., Mikkelsen, P.S., "Investigating the use of stochastic forecast for RTC of urban drainage systems", *Novotech 2013,* Lyon (2013)

[19] Vojnovic, Z., Kecman, V., and Babovic, V. "Hybrid Approach for Modeling Wet Weather Response in Wastewater Systems", *J. Water Resour. Plann. Manage.*, Vol. 129(6), (2003), pp 511–521.