Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems

Löwe, Roland; Vezzaro, Luca; Mikkelsen, Peter Steen; Grum, Morten; Madsen, Henrik

Published in:
Environmental Modelling & Software

Link to article, DOI:
10.1016/j.envsoft.2016.02.027

Publication date:
2016

Document Version
Peer reviewed version

Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems

Roland Löwe[a, c,*,**], Luca Vezzaro[b, c], Peter Steen Mikkelsen[b], Morten Grum[c], Henrik Madsen[a]

[*]present address: Department of Environmental Engineering, Technical University of Denmark (DTU Environment), Miljøvej B115, Kgs. Lyngby, 2800, Denmark
[**]correspondence to: rolo@env.dtu.dk, +45-27600355

[a]Department of Applied Mathematics and Computer Science, Technical University of Denmark (DTU Compute), Matematiktorvet B303, Kgs. Lyngby, 2800, Denmark, hmad@dtu.dk
[b]Department of Environmental Engineering, Technical University of Denmark (DTU Environment), Miljøvej B115, Kgs. Lyngby, 2800, Denmark, luve@env.dtu.dk, psmi@env.dtu.dk
[c]Krüger A/S, Veolia Water Solutions and Technologies, Gladsaxevej 363, Søborg, 2860, Denmark, mg@kruger.dk

ABSTRACT

This article demonstrates the incorporation of stochastic grey-box models for urban runoff forecasting into a full-scale, system-wide control setup where setpoints are dynamically optimized considering forecast uncertainty and sensitivity of overflow locations in order to reduce combined sewer overflow risk.
The stochastic control framework and the performance of the runoff forecasting models are tested in a case study in Copenhagen (76 km2 with 6 sub-catchments and 7 control points) using 2-hour radar rainfall forecasts and inlet flows to control points computed from a variety of noisy/oscillating in-sewer measurements.
Radar rainfall forecasts as model inputs yield considerably lower runoff forecast skills than "perfect" gauge-based rainfall observations (ex-post hindcasting). Nevertheless, the stochastic grey-box models clearly outperform benchmark forecast models based on exponential smoothing.
Simulations demonstrate notable improvements of the control efficiency when considering forecast information and additionally when considering forecast uncertainty, compared with optimization based on current basin fillings only.

KEYWORDS

stochastic grey-box model, probabilistic forecasting, real-time control, urban hydrology, radar rainfall, storm water management

Revised
2016/02/14

1 INTRODUCTION

This article investigates the application of probabilistic multistep runoff forecasts generated by simple, conceptual stochastic models (in the form of so-called stochastic grey-box models) in system-wide, forecast-based optimization for real-time control (RTC) of urban drainage networks. A drainage network is considered to be controlled in real time if process variables are monitored in the system and used to operate actuators affecting the flow process (Schütze et al., 2004). RTC is an efficient tool for responding to changing demands that are defined for

urban drainage systems (Rauch et al., 2005, Vanrolleghem et al., 2005) and is increasingly applied to operate these infrastructures in an efficient manner (for example, Mollerup et al., 2013, Nielsen et al., 2010; Pabst et al., 2011; Pleau et al., 2005; Puig et al., 2009 and Seggelke et al., 2013). In particular, RTC can support the operation of combined sewer systems, which are used in most of the larger European cities and are constantly challenged by increased impervious area and changing rainfall patterns (Arnbjerg-Nielsen et al., 2013; Willems et al., 2012).

Most RTC implementations aim to minimize the volume of combined sewer overflows (CSO). This is achieved by dynamically controlling flows in the system to achieve an optimal exploitation of the available storage volume, especially in cases with an uneven spatial rainfall distribution over the catchment. RTC is classically performed using static if-then-else rules (Seggelke et al., 2013, for example) that are optimized off-line based on heuristics and model simulations, but mathematical optimization routines are also applied (Pleau et al.,2005, Puig et al.,2009).

Clearly, information on the future evolution of the urban drainage system (i.e., the runoff expected in the near future) should contribute to a more efficient optimization of the controlled system. Significant developments have been made in the last decade in terms of radar-based rainfall forecasting (Krämer et al., 2005, 2007; Thorndahl et al., 2014; Vieux and Vieux, 2005) and radar-based urban runoff forecasting (Achleitner et al., 2009; Löwe et al., 2014a; Schellart et al., 2014; Thorndahl and Rasmussen, 2013), paving the way for the application of radar-based online runoff forecasts in RTC.

However, multiple sources of uncertainty affect the runoff forecasts generated by models (see the discussions in Deletic et al. (2012), Schilling and Fuchs (1986) and Sun and Bertrand-Krajewski (2013)): input uncertainty, model structure uncertainty, parameter uncertainty and measurement uncertainty (e.g., level and flow). The examples in Schilling and Fuchs (1986), Schilling (1991) and Schellart et al. (2011) demonstrate that uncertainty of the measured and forecasted rainfall input is often the major factor affecting the online performance of runoff forecast models. Previous studies have evaluated the accuracy of online runoff forecasts based on radar rainfall input in an urban setting and found the forecast performance diminished for lead-times greater than 90 minutes (Achleitner et al., 2009) and between 60 and 120 minutes (Thorndahl and Rasmussen, 2013).

Considering the large uncertainties of urban runoff forecasts, it has been hypothesized that the uncertainties may adversely impact the efficiency of forecast-based RTC schemes (Breinholt et al., 2008; Schütze et al., 2004). As a result, RTC algorithms that account for these uncertainties in mathematical optimization have recently emerged. Examples include the tree-based control algorithm, which was proposed for control of (non-urban) drainage water systems by Maestre et al. (2013), and the dynamic overflow risk assessment (DORA; Vezzaro and Grum, 2014) for urban drainage systems that performs a system-wide optimization based on the computed risk of overflow.

Accounting for the uncertainty of runoff forecasts in RTC requires that an estimate of this uncertainty is provided as an input to the control algorithm. The literature on uncertainty quantification in rainfall runoff modelling is abundant. Informal approaches (GLUE) are popular in urban hydrology (e.g., Dotto et al., 2012; Freni et al., 2009; Vezzaro and

Mikkelsen, 2012), while more formal Bayesian approaches without (Del Giudice et al., 2015a; Kavetski et al., 2006) and with data assimilation routines (Moradkhani et al., 2012; Vrugt et al., 2013) were developed mostly for natural catchment hydrology. Model estimation and updating in these approaches are commonly based on Monte Carlo simulations, and they can therefore be difficult to apply in an online context (Del Giudice et al., 2015b).

Recent research in the Storm- and Wastewater Informatics Project (SWI, 2015) has therefore focused on the application of so-called stochastic grey-box models for probabilistic online runoff forecasting over multiple prediction horizons. This type of model combines a simple and fast stochastic model structure with a data assimilation routine in the form of an extended Kalman filter, allowing the user to generate probabilistic forecasts with time-dynamic uncertainty quantification. The application of such models in urban hydrology was first tested by Carstensen et al. (1998) and Bechmann et al. (1999). Breinholt et al. (2011, 2012) developed rainfall-runoff model structures, and the performance of these for probabilistic flow predictions was assessed by Thordarson et al. (2012). Finally, Löwe et al. (2014a) analysed the influence of different rainfall inputs on runoff forecast performance, while different options for parameter estimation were compared in Löwe et al. (2014b).

The work presented here combines these recent developments: probabilistic, radar-rainfall based runoff forecasts from stochastic grey-box models have been combined with a risk-based optimization algorithm that accounts for time-dynamic forecast uncertainty (DORA, Vezzaro and Grum, 2014) and integrated into a full-scale, system-wide RTC setup, providing a proof of concept for the case of applying stochastic forecasts in RTC. The setup is tested in a case study with noisy real-world measurements and six sub-catchments with distinctly different characteristics. The purpose of this article is to

- demonstrate this new, stochastic, system-wide real-time control setup for urban drainage systems,
- evaluate how the consideration of runoff forecast uncertainty influences the efficiency of the RTC scheme, and
- evaluate what runoff forecast performance and what control efficiency can be obtained with stochastic grey-box models and radar rainfall input under realistic conditions in a variety of catchments.

The new control setup applies stochastic grey-box models for runoff forecasting. However, other probabilistic forecasting methods (such as the ones presented by Todini (2008), Van Steenbergen et al. (2012), Vrugt et al. (2005) or Weerts et al. (2011)) could easily be implemented. Thus, the proposed framework is generic in this respect.

2 METHODS

2.1 STOCHASTIC REAL-TIME CONTROL SETUP

2.1.1 General Setup

A system-wide control setup was applied. Control points need to be defined by the users and are typically located at major actuators, such as the outlet of storage basins or pumping stations. Runoff forecasts were generated by a separate stochastic model (Section 2.1.2) for the inflow to each control point. Based on the inflow forecasts and online observations of the current basin fillings, the DORA algorithm was then used to optimize the outflow from all of the control points, aiming to minimize the overall overflow risk in the catchment (Section

2.1.3). A control time step of 2 minutes was applied and a maximum forecast horizon of 2 hours was considered. Correspondingly, new runoff forecasts were generated every 2 minutes for 2 hours into the future with a resolution of 60 time steps (intervals of 2 minutes).

The online operation of the framework is illustrated in Figure 1. It can be split into 5 steps that are executed every 2 minutes:

1. Data collection – the runoff forecast models apply rainfall forecasts as an input and flow observations for updating the model states. In addition, the current basin filling is required as an input to the control algorithm. Depending on the source, these data are either downloaded as text files through FTP connections or directly imported from the SCADA system through the standard OPC UA (Unified Architecture) protocol (Mahnke et al., 2009).
2. Pre-processing – flow observations are required to update the states of the runoff forecast models (Section 2.1.2). However, for many control points, no direct inflow measurements are available. Instead, these need to be constructed by "software sensors" from a combination of indirect measurements (such as level in and outflow from a storage basin). Catchment specific pre-processing routines (see appendix A) are therefore implemented in this module. The software WaterAspects (Grum et al., 2004) was applied for this step in our work, while future implementations will apply JEP and R scripts.
3. Runoff forecasting – a separate stochastic grey-box model (Section 2.1.2) is applied for forecasting the inflow volume to each control point. The model output is a distribution of forecasted runoff volume for each considered horizon, discretized in 50 quantiles from 1 to 99 %. Each model in our work was implemented as an executable (FORTRAN-based) that communicates with the control server via text files. An R-based setup that directly communicates with the database is currently being implemented.
4. Identifying set points for the actuators using the DORA algorithm (Section 2.1.3) – this module is implemented in JAVA. The overflow risk for each control point is computed based on the current basin filling and the forecasted distribution of runoff volumes in the form of quantiles.
5. The new outflow set points for the actuators are sent to the SCADA system through the standard OPC UA protocol.

A control software is required to manage the execution of the tasks mentioned above, the communication with external data sources and actuators, data storage in a database and quality control of measurements and simulation results. In our case, the STAR® Utility Solutions™ framework (Nielsen and Önnerth, 1995) was used. The framework is implemented in JAVA but allows for the execution of external programs written in, for example, R. The framework can be installed either on a dedicated control server, on a cloud server or within the end-user's own virtual server environment.

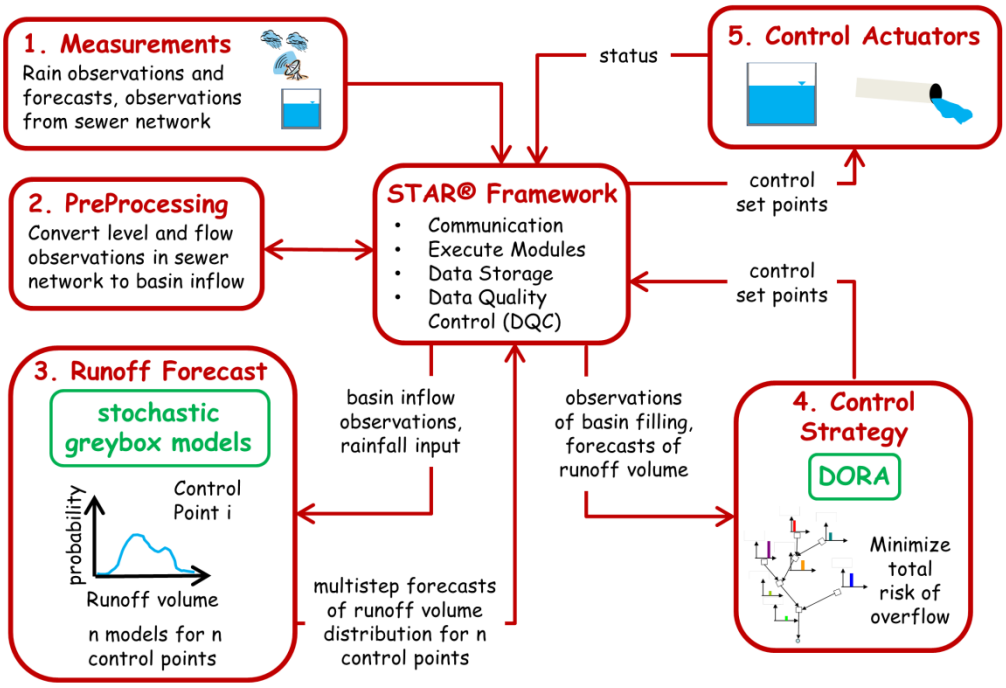

Figure 1. Technical integration of stochastic grey-box models and DORA in a STAR Utility Solutions™ control setup.

2.1.2 Runoff Forecasting Using Stochastic Grey-Box Models

Model Structure

A simple cascade of three linear reservoirs was applied for forecasting runoff volume in the inflow to a single control point. We did not consider more elaborated model structures as the purpose of this article is to provide a proof of concept. The model was set up as a stochastic grey-box model in a state-space layout as described by Breinholt et al. (2011) and shown in state equations (1), which are implemented using stochastic differential equations (SDEs) and observation equation (2). The setup includes an extended Kalman filter, which updates the model states whenever new flow observations become available (Kristensen et al., 2004). The model was implemented in the open source software CTSM-R (Juhl et al., 2013).

$$d\begin{bmatrix} S_{1,t} \\ S_{2,t} \\ S_{3,t} \\ a_0 \end{bmatrix} = \begin{bmatrix} A \cdot P + a_0 - \frac{1}{K} S_{1,t} \\ \frac{1}{K} S_{1,t} - \frac{1}{K} S_{2,t} \\ \frac{1}{K} S_{2,t} - \frac{1}{K} S_{3,t} \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma_1 S_{1,t} & & & \\ & \sigma_2 S_{2,t} & & \\ & & \sigma_3 S_{3,t} & \\ & & & \sigma_4 \cdot I \end{bmatrix} d\omega_t \quad (1)$$

$$Y_k = \frac{1}{K} S_{3,k} + D_k + e_k \quad (2)$$

S_1, S_2 and S_3 correspond to the storage states, A to the effective catchment area, P to the rain intensity, a_0 to the mean dry weather flow and K to the travel time constant. The uncertainty of model predictions is described in the so-called diffusion term, which is driven by a vector Wiener process $d\omega_t$ (Iacus, 2008). Considering a time step Δt, an increment $\Delta\omega_t$ of this process is Gaussian with mean 0 and covariance $diag(\Delta t, \Delta t, \Delta t, \Delta t)$. The parameters σ_i scale the standard deviation of the diffusion process, which here increases linearly with the state value S_i. We have included the mean dry weather flow a_0 as a state to allow the model to adapt to varying dry weather flows, which we have observed in some of the catchments considered in our case study. The index I was 1 during the updating step of the extended Kalman filter and 0 when generating runoff forecasts. The last-known estimate of a_0 was thus applied during the generation of multistep runoff forecasts.

The observation equation (2) relates time-continuous model predictions and flow observations Y_k at discrete time steps k. This equation additionally includes a trigonometric function D to describe the variation of dry-weather flows (Breinholt et al., 2011) and the observation error e_k with standard deviation σ_e.

A Lamperti transformation (Iacus, 2008) was applied to the state equations (1) to remove the dependency of the noise description on the state (Breinholt et al., 2011), as state-dependent SDEs are difficult to simulate numerically (Iacus, 2008).

The diffusion term in Equation 1 accounts for the combined effects of input and model structure uncertainty. The observation error e_k in Equation (2) can account for deficiencies in the sensor measurements as well as for oscillations resulting, for example, from varying pumping discharges. The latter were treated as noise if they occurred on short time scales of only few minutes, as such variations have only little effect on the basin volumes at the control points. The parameters A and K, the uncertainty scalings σ_i of the diffusion term and the standard deviation of the observation error σ_e were estimated as part of the automated calibration routine.

Parameter estimation

The model parameters were determined in an automated calibration routine. As an objective function, we minimized the multistep probabilistic flow forecast error as described by Löwe et al. (2014b). Using the state prediction equations of the extended Kalman filter (Eq. 17 and 18 in Kristensen et al. (2004)) and subsequently inserting the state predictions into the output prediction equations (Eq. 11 and 12 in Kristensen et al. (2004)), a multistep flow forecast was generated at each time step k for forecast horizons $i = 1 \ldots 60$ with a resolution of $\Delta t = 2\ min$. The forecasts were assumed Gaussian with mean $\hat{Y}_{k+i|k}$ and forecast covariance $\widehat{\boldsymbol{R}}_{k+i|k}$. As an estimate for the probabilistic forecast error, we computed the continuous ranked probability score $CRPS_{i,k}$ (Gneiting and Raftery, 2007) for each forecast horizon i as

$$CRPS_{i,k} = \int_{-\infty}^{\infty} \left(\hat{F}_{k+i|k}(s) - \mathcal{H}(s > Y_{k+i}) \right)^2 ds \,, \tag{3}$$

where $\hat{F}_{k+i|k}(s)$ is the cumulative normal distribution function of the flow forecast, Y_{k+i} is the observed flow for the corresponding time step and $\mathcal{H}$ is the Heaviside function that takes the value 0 if $s < Y_{k+i}$ and 1 otherwise. A closed-form solution of the $CRPS$ is available for Gaussian $\hat{F}_{k+i|k}(s)$. However, we chose to evaluate the integral in Eq. 3 numerically for

quantiles from 1 to 99 % in steps of 2 % to make the approach flexible for other distributional assumptions. A measure of average performance over all forecast horizons was defined as

$$CRPS_k = \frac{1}{\sum_{i=1}^{60}(60-i+1)} \left(\sum_{i=1}^{60}((60-i+1)\cdot CRPS_{i,k})\right). \quad (4)$$

The RTC scheme requires forecasts of runoff volume as an input (see Section 2.1.3). Therefore, more weight is put on flow forecasts for shorter forecast horizons in Eq. 4. These have a stronger influence on forecasts of runoff volume, which are generated as an integral over flow forecasts for several horizons. Finally, averaging the $CRPS_k$ over all time steps k provided the objective function for parameter estimation, which we aimed to minimize.

We applied the heuristic optimization algorithm described by Tolson and Shoemaker (2007) with 2500 objective function evaluations for automated parameter estimation. The dry weather flow variation D was fixed during the parameter estimation process. The corresponding parameters were estimated separately during a dry weather period.

On-line forecast generation

To generate probabilistic runoff forecasts online, we performed scenario simulations of the model equations (1), starting with the updated states provided by the extended Kalman filter at time step t and ending at the maximum considered forecast horizon $t + j$. We considered $N = 1000$ scenarios. The forecasted flow for each scenario was integrated into a runoff volume. The resulting empirical distribution of forecasted runoff volumes served as input to the control algorithm in the form of quantiles with a resolution of 2 %. The approach was described in more detail by Löwe (2014) and Löwe et al. (2014a).

The generation of on-line runoff forecasts was based on scenario simulations of the stochastic process without distributional assumption, while assumed-Gaussian forecasts were generated using the extended Kalman filter during parameter estimation. This inconsistency is a shortcoming of the current setup, which was caused by the need to generate forecasts with limited computational effort during parameter estimation.

2.1.3 Real-time Control under Uncertainty

We applied the dynamic overflow risk assessment (DORA, see Vezzaro and Grum (2014) and Vezzaro et al. (2014)) in this study. This approach, in the terminology of Mollerup et al. (2015), acts on the optimization layer of the real-time control setup, aiming for a system-wide (across the entire catchment) reduction of the risk of CSO using a forecast-based mathematical optimization routine that accounts for both forecast uncertainty and impact cost.

The overflow risk for each controlled point is calculated by

1. subtracting the basin outflow volume over the forecast horizon and the currently free basin volume from the forecasted probability distribution of runoff volume, and
2. multiplying the resulting probability distribution of overflow by a constant CSO unit cost that is user-defined for each overflow location (which reflects the sensitivity of the different receiving waters). More sensitive control points (e.g., discharging to bathing areas) are given higher CSO unit costs than less sensitive control points (e.g., discharging close to the wastewater treatment plant (WWTP) inlet).

The motivation for DORA is that stochastic forecasts are needed because a deterministic forecast only leads to optimal control decisions if the loss function applied in optimization does not depend on the uncertainty range associated with the forecasted variable. Even for the simple CSO unit cost applied here, this is clearly not the case because the overflow risk is a discontinuous function that is zero for small forecasted runoff volumes and increases linearly for larger forecasted runoff volumes that would lead to an overflow of the basin.
At each control time step (in this study set to 2 minutes, i.e., each time a new set of measurements from the catchment becomes available), DORA executes the following loop (Figure 2):

- Step 1: The available storage volume for each basin is calculated using online measurements.
- Step 2: Runoff forecasts (and the associated uncertainty) are used to estimate the overflow risk for each controlled point. The probability density for the forecasted runoff volume is here described empirically by a set of quantiles provided by the stochastic grey-box model. This is different from the approach in Vezzaro and Grum (2014), who described forecast uncertainty analytically by a Gamma distribution with roughly fixed parameters.
- Step 3: A genetic algorithm (Meffert et al.) is used to identify the optimal set of flows between all of the basins in the catchment, minimizing the total CSO risk. The settings of the algorithm were defined for the study area after off-line tests, which focused on convergence (especially in dry weather conditions, when CSO risk is low and several solutions form a Pareto front). By initializing the algorithm from the solution obtained at the previous time step, a population size of 100 and a maximum of 50 evolutions were sufficient to obtain the desired convergence and reliability.
  When the CSO risk is low (e.g., after the end of a rain event with no new rainfall within the forecast horizon), DORA empties the controlled system as quickly as possible, with the highest priority on the control points with the largest CSO cost.
- Step 4: Optimal set points for each basin outflow are sent to the actuators in the system.

DORA does not currently account for transport times in the optimization step 3 (see Vezzaro and Grum (2014)). Instead, an immediate transfer of outflow volumes is assumed between the control points.

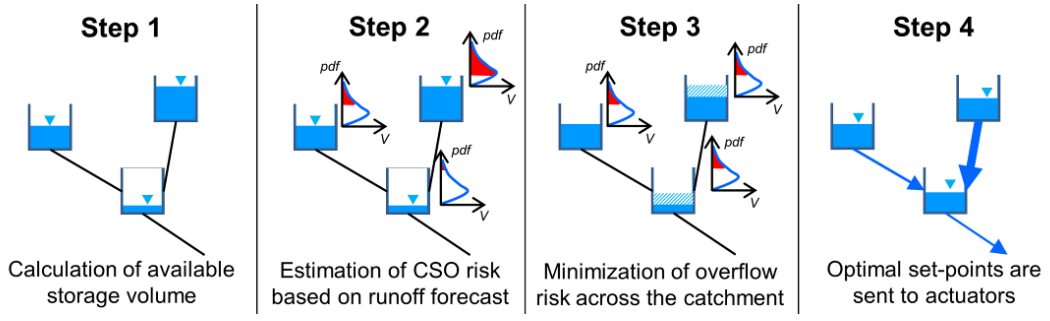

Figure 2. Schematic representation of the principal steps in DORA. The runoff volume V is forecasted with a probability density function pdf. The part of the pdf used for computing the probability of overflow is marked in red.

2.2 PERFORMANCE EVALUATION

We validated the stochastic forecasting and control setup in a two-step procedure. First, we evaluated runoff forecasting performance by comparing forecasts and observations. Second, we determined the efficiency of the control setup with and without forecast uncertainty and considering different rainfall inputs.

2.2.1 Evaluation of Forecast Quality

In the evaluation of forecast performance, we focused solely on lead times of 120 minutes (60 time steps) into the future because this is the longest horizon considered in the system-wide control scheme and may be considered as the worst case.

Point Forecast Skill

To assess point forecast quality, we applied a skill score defined as:

$$SPI = 1 - \frac{\sum_{k=1}^{N}\left(\widehat{V}_{k+60|k,50\%} - \sum_{k=1}^{60} Y_{k+i}\cdot\Delta t\right)^2}{\sum_{k=1}^{N}\left(\sum_{i=1}^{60}\left((1-\lambda)\cdot Y_{SM,k-1} + \lambda\cdot Y_k\right)\cdot\Delta t - \sum_{i=1}^{60} Y_{k+i}\cdot\Delta t\right)^2}. \quad (5)$$

In equation 5, the numerator of the fraction is the mean squared error of the runoff volume forecasts generated by the stochastic grey-box models. $\widehat{V}_{k+60|k,50\%}$ is the median of the probabilistic forecast of runoff volume generated by the stochastic grey-box models at time step k for a forecast horizon of 60 time steps. Y_k are the flow observations for the same period. These are available in intervals of $\Delta t = 2\ min$ and for a total of N time steps during an event for which the forecast skill is computed.

The denominator of the fraction in equation (5) is the mean squared error of a reference (or benchmark) forecast. As a reference, we considered locally constant runoff volume forecasts derived using exponential smoothing (Brown and Meyer, 1961). $Y_{SM,t-1}$ is the smoothed flow observation obtained for the previous time step and λ is the smoothing parameter, which was tuned to minimize the 60 step forecast error shown in the denominator in equation (5) during the calibration events described in section 3 and which can vary between 0 and 1.

We denote the resulting skill score as the smoothed persistence index (SPI) because it resembles the persistence index described in Bennett et al. (2013). However, a smoothed value is applied as the reference forecast instead of the last observation to make the score more robust towards the noisy flow measurements we encountered in our study. Ideally, the SPI would take a value of 1 for a perfect runoff forecast. Values smaller than 0 indicate that the forecasts generated by the stochastic grey-box models have a bigger mean squared error than the locally constant forecast based on exponential smoothing.

Forecast Reliability

In a probabilistic sense, it is desirable for the runoff forecasts to be reliable. An α % prediction interval should empirically include α % of the observations, i.e., have an observed coverage rate of α %. This property of the probabilistic forecasts can be assessed by plotting predicted (or nominal) and observed coverage rates against each other in reliability diagrams (Murphy and Winkler, 1977). Such diagrams are easier to understand and simplify

communication with practitioners and were therefore preferred over the probability integral transform used by, for example, Hemri et al. (2013) and Renard et al. (2010). Ideally, predicted and observed coverage rates should be equal. Predicted coverage rates smaller than the observed coverage rates indicate an overestimation of forecast uncertainty by the model, while the reverse indicates an underestimation of forecast uncertainty.

Sharpness of Forecasts

Finally, given a reliable probabilistic forecast, it is desirable for it to be as sharp (or "narrow") as possible. A common measure is the sharpness or average width of an α % prediction interval. Jin et al. (2010) normalized this measure with the observation to obtain the average interval width *ARIL*. The observation, however, is not related to the forecast and *ARIL* will be difficult to evaluate if the observations approach zero, for example. We therefore applied a modified version of *ARIL* in which we normalized by the absolute value of the forecast median. We applied this version for the 90 % prediction interval as a measure of forecast uncertainty:

$$ARIL^{*} = \frac{1}{N}\sum_{t=1}^{N}\frac{\hat{V}_{95\%,k+60|k} - \hat{V}_{5\%,k+60|k}}{\left|\hat{V}_{50\%,k+60|k}\right|} \qquad (6)$$

In (6), $\hat{V}_{95\%,k+60|k}$, $\hat{V}_{50\%,k+60|k}$ and $\hat{V}_{5\%,k+60|k}$ correspond to the 95 %, 50 % and 5 % quantiles of the probabilistic runoff volume forecasts generated at time step t for a lead time of 120 minutes (60 time steps). Smaller values of $ARIL^{*}$ indicate narrower prediction intervals.

2.2.2 Evaluation of Control Efficiency

To evaluate the effect of different forecast inputs on the efficiency of the system-wide control algorithm, simulations need to be performed in a model that describes flows in all relevant parts of the catchment, includes all actuators and allows for the evaluation of CSO in different scenarios (as demonstrated by Seggelke et al., 2013, for example). In the evaluation, this model (Section 3.2) replaces the actuators in Figure 1 and provides current basin fillings as input to the DORA algorithm.

To compare the performance of the setup in different scenarios, we focused on the evaluation of overflow volumes and cost accumulated over a number of rain events. Reduced overflow volumes in a scenario indicate an improved performance of the control system. The best performing setup minimizes the total overflow cost, which corresponds to the overflow volume weighted according to the expected environmental impact at the location of the overflow structures. The weighting factors correspond to the CSO unit cost defined in DORA for the different overflow structures (see Section 2.1.3 and Table 1 in Section 3)

2.2.3 Considered Scenarios

Five scenarios were simulated to (i) evaluate the influence of runoff forecast uncertainty on the efficiency of system-wide control and (ii) estimate what forecast performance and what control efficiency can be achieved under realistic conditions:

1. **AU – Rain gauge based runoff forecast with uncertainty**: The inputs for the stochastic grey-box models were the rain gauge measurements averaged for each sub-catchment (see Section 3.3.1).
   Rainfall forecasts are required as model input for the generation of runoff forecasts. In

this scenario, perfect rainfall forecasts derived from the rain gauge measurements for the forecast period where applied, both when calibrating the parameters of the runoff forecast models and when evaluating runoff forecasting performance and control efficiency.

2. **ANU – Rain gauge based runoff forecast without uncertainty**: Runoff forecasts were generated in the exact same way as in scenario AU. However, runoff forecast uncertainty was neglected when evaluating control performance by defining a forecast distribution with negligible standard deviation (the forecast median divided by 2500) around the forecast median.
3. **BU – Radar based runoff forecast with uncertainty**: Radar rainfall measurements and forecasts (see Section 3.3.1) were used as model input for calibrating the runoff forecast models, for evaluating runoff forecasting performance and for evaluating control efficiency.
4. **BNU – Radar based runoff forecast without uncertainty**: Runoff forecasts were generated in the exact same way as in scenario BU. However, runoff forecast uncertainty was neglected when evaluating control performance by defining a forecast distribution with negligible standard deviation (the forecast median divided by 2500) around the forecast median.
5. **REF – No forecast**: This is a reference scenario for the evaluation of control efficiency only. In this scenario, DORA was used with a zero forecast as described by Vezzaro and Grum (2014). The control algorithm in this case simply attempts to equalize the basin fillings in the different sub-catchments, weighted according to the CSO unit cost at the overflow points (Table 1).

Scenario AU provides a base case with near-perfect rainfall forecast. Scenario BU, on the other hand, illustrates the runoff forecast quality and control efficiency that can be achieved with more realistic rainfall forecasts. As the skill of radar rainfall forecasts strongly decreases with the forecast horizon (Achleitner et al.,2009; Thorndahl and Rasmussen, 2013), scenario BU would be expected to yield lower runoff forecasting skill and reduced control efficiency as a result of the larger uncertainty of the rainfall input applied in this case.

If the consideration of forecast uncertainty has a (positive) impact on the performance of system-wide control (as hypothesized by Vezzaro and Grum (2014) and Löwe et al. (2014b)), then scenarios AU and BU should yield better control results than their counter parts ANU and BNU.

Finally, the reference scenario REF provides a reasonable benchmark for the control performance obtained when applying DORA with and without runoff forecasts as input.

3 CASE STUDY

3.1 CATCHMENT

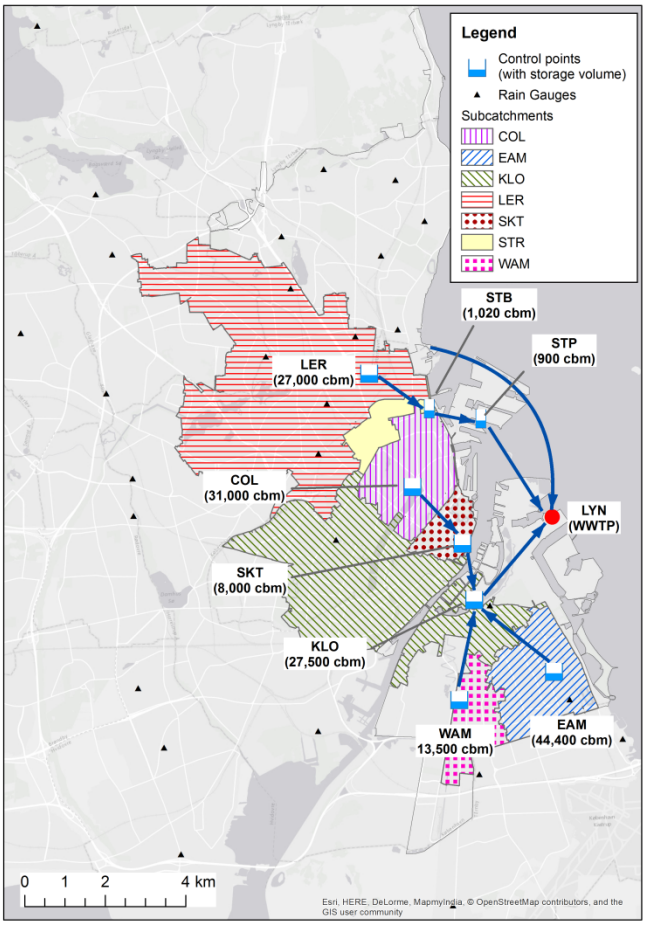

Figure 3. Catchment of the Lynetten wastewater treatment plant (WWTP) with control points in the combined sewer system and their respective sub-catchments.

The case study was designed to test the setup in a situation where the runoff forecast models need to cope with a variety of sub-catchments with different characteristics (Table 1), where realistic rainfall forecasts are applied (Section 3.3.1) and where outflow measurements are far from perfect (Section 3.3.1 and Appendix C). We considered the catchment of the Lynetten wastewater treatment plant (WWTP), which covers the central area of Copenhagen (Denmark) and has a total area of approximately 76 km^2. The system-wide control strategy for the Lynetten catchment considers seven sub-catchments and nine overflow structures (see Figure 3), discharging to recipients with different sensitivities to CSO. Large storage basins were implemented in the catchment over the past three decades as a result of efforts to minimize CSO and secure bathing water quality in the harbour. The total storage capacity is approximately 153,000 m^3.

Separate stochastic grey-box models were implemented to forecast runoff volumes for the inflow to each control point. No runoff forecasts were generated for the sub-catchments discharging to the St. Annæ basin (SKT) and to the WWTP inlet (LYN) due to the very poor quality of the available flow and water level observations. Only the current filling rate at these control points was included in the optimization strategy to calculate the system-wide CSO risk and no control decisions were determined for the corresponding actuators. The

Strandvænget sub-catchment comprises two control points at the basin outlet (STB) and the pumping station (STP) to the WWTP. Runoff forecasts were only generated for the basin inflow because the pumping station only receives inflows from STB. The characteristics of the sub-catchments are summarized in Table 1.

Table 1. Main characteristics of the control points considered. Points not controlled by DORA are used to calculate the CSO risk, but they are not considered as actuators in the optimization algorithm

Sub-catchment	Imper-vious area [ha]	Storage available for RTC [m^3]	Max outflow [m^3/s]	CSO unit cost [€/m^3]	Controlled by DORA	Typology
Colosseum (COL)	211	30,914	0.9	5	X	basin, pumped outflow
East Amager (EAM)	228	44,425	2.1	25	X	storage pipes, pumped outflow
Kloevermarken (KLO)	777	27,500	7.5	5	X	pumping station with storage in upstream pipe network
Lersoeledning (LER)	733	27,000	1.1	25	X	storage pipe with gate
Lynetten WWTP (LYN)	564	76	5 (6.4[a])	1		CSO at WWTP inlet
St. Annæ (SKT)	77	7,987	1.3	5		basin, pumped outflow
Strandvaenget						
Basin (STB)	92	1,020	3.9	25	X	CSO structure, pumped outflow
Pumping station (STP)	-	900	2.4	1	X	pumping station
West Amager (WAM)	97	13,490	1.0	5	X	basin, pumped outflow
Total	2,279	153,312				

3.2 Catchment Simulation Model for the Evaluation of Control Efficiency

We used a conceptual model of the Lynetten catchment (implemented in WaterAspects - Grum et al., 2004) to evaluate the control efficiency. Following the procedure presented by Borsanyi et al. (2008), this model was calibrated against a detailed MIKE URBAN model of the catchment. A sketch of the model together with a comparison of simulated and observed inflows to the control points EAM, COL, KLO, LER, SKT and WAM is provided in Appendix C for all rain events.

The generation of runoff was described using the time area method, and a simple time delay was applied for routing in pipes. Local controls existing in the catchment (e.g., pumping based on filling degree in basins) were implemented in the model. They were overridden by the DORA set points when system-wide control strategies were simulated.

Rain gauge measurements averaged over each sub-catchment (see Section 3.3.1) were used as input for the catchment simulation model.

3.3 Data and Simulation Periods

3.3.1 Rain Data and In-Sewer Observations

A time step of 2 minutes was adopted for all of the datasets in this work, corresponding to the control time step of the existing control setup. Data available at higher temporal resolution were averaged, while data with lower temporal resolution were assumed constant in between observations ("zero order hold"). Online measurements were available for the period from November 2011 until September 2014.

Rain measurements from 29 gauges in the area (Figure 3) with a temporal resolution of 1 minute were available from the network of the Danish Water Pollution Committee (SVK), which is operated by the Danish Meteorological Institute (Jørgensen et al., 1998). A time series of mean areal rainfall was determined for each of the sub-catchments shown in Figure 3 using Thiessen polygons.

Radar rainfall measurements and forecasts were available from the C-band radar of the Danish Meteorological Institute in Stevns. The data had a resolution of 10 minutes in time and 2x2 km in space. The radar data were time-dynamically adjusted to rain-gauge data at every time step as described in Löwe et al. (2014a), Thorndahl et al. (2013) and Thorndahl and Rasmussen (2013). A mean areal rainfall series was calculated for each sub-catchment from the radar data by computing a weighted average of the rainfall measured in different pixels. The weighting factors for this process were determined from the intersecting area between a pixel and the corresponding sub-catchment.

Historical radar rainfall forecasts were made available for forecast horizons of 10, 20, 30, 60 and 90 minutes. We interpolated the forecasts for horizons of 40, 50, 70 and 80 minutes and assumed that the rainfall forecasts for the 100 to 120 minute horizons were equal to the forecast for the 90 minute horizon. This is a limitation in our work caused by the data that were made available to us. In reality, a radar-based flow forecasting setup would be expected to perform slightly better than presented here.

Various level and flow measurements from the sewer network were available for the considered period (see Appendix A). In most sub-catchments, no direct measurements of the inflow to the control point were provided. However, inflow measurements are required to update the stochastic runoff forecasting models (see Equation 2) and to evaluate forecast performance. They were computed from the available data using the water balance for each control point and (in some cases) rating curves (see Appendix A). This approach led to noisy flow measurements (see Appendix C) and proved problematic in the LER and STB catchments, where negative measurements were obtained after rain events because the water balance was not closed in some situations. Such systematically negative data were excluded from the updating of the forecast models and from the evaluation of forecast performance.

3.3.2 Selection of Rain Events

Rain events were identified from the mean areal radar rainfall measurements for the six sub-catchments where stochastic runoff forecasting models were implemented. An event was

considered to start when any of the mean areal rainfall series exceeded a threshold intensity of 0.2 mm/10 min. The event was considered to end when the mean areal rainfall series for all sub-catchments were below this threshold for a period of at least 10 hours.

Based on these criteria, a total of 422 rain events were identified between Nov 2011 and Sep 2014. Many of these events were unlikely to cause CSO due to the small rainfall volumes involved. In addition, significant data gaps were observed for many events. The number of events under consideration was reduced in the three-stage procedure shown in Table 2.

Appendix B lists all 130 rain events identified after the first two stages of data inspection, while Appendix C depicts the observed inflow to the control points for these events. Rain events that were identified as problematic during visual inspection were excluded from the evaluation of forecast performance in the corresponding catchment as well as from the evaluation of control efficiency. These events are marked in the table in Appendix B and with a grey background in Appendix C.

In total, between 114 and 127 rain events were considered for the evaluation of forecast performance in the different sub-catchments, and 98 events were considered for the evaluation of control efficiency. Four rain events were selected for estimating parameters of the forecast models. These were chosen to cover different rainfall characteristics (short, intense and localized storms as well as widespread, long lasting rainfall) in different seasons and are marked in Appendix B.

Table 2. Procedure for selecting rain events for simulation. The table shows the criteria applied in different stages together with the number of rain events removed from the dataset according to the different criteria.

Events removed according to criterion	**Criterion**
Stage 1 (automated) – Remove insignificant events	
251	observed maximum inflow at any of the considered control points (after smoothing) did not exceed the peak dry weather flow by at least a factor of 1.5, **or** a simulation with a conceptual model of the whole catchment without system-wide RTC (Section 3.2) did not yield CSO and the maximum rain intensity averaged over the whole catchment was below 1 mm/30 min
Stage 2 (automated) – Remove events with bad data quality	
32	at least 10 % of the in-sewer measurements were missing in at least one of the considered sub-catchments,
	or
7	the maximum radar rain intensity, averaged over a 30 min interval and the whole Lynetten catchment, was higher than 30 mm/30 min while no corresponding increased runoff was observed,
	or
2	both of the above issues
Stage 3 (manual) – Visual inspection of the remaining events	
3 to 16 (depending on sub-catchment under	inflow measurements had no relation to the rainfall measured by radar and gauges

consideration)

4 RESULTS

4.1 FORECAST PERFORMANCE

This section focuses on the evaluation of runoff forecast performance obtained for the stochastic grey-box models. As explained in Section 2.2.1, all of the results shown in the following were derived for forecasts of runoff volume for a forecast horizon of 120 minutes, corresponding to 60 control time steps.

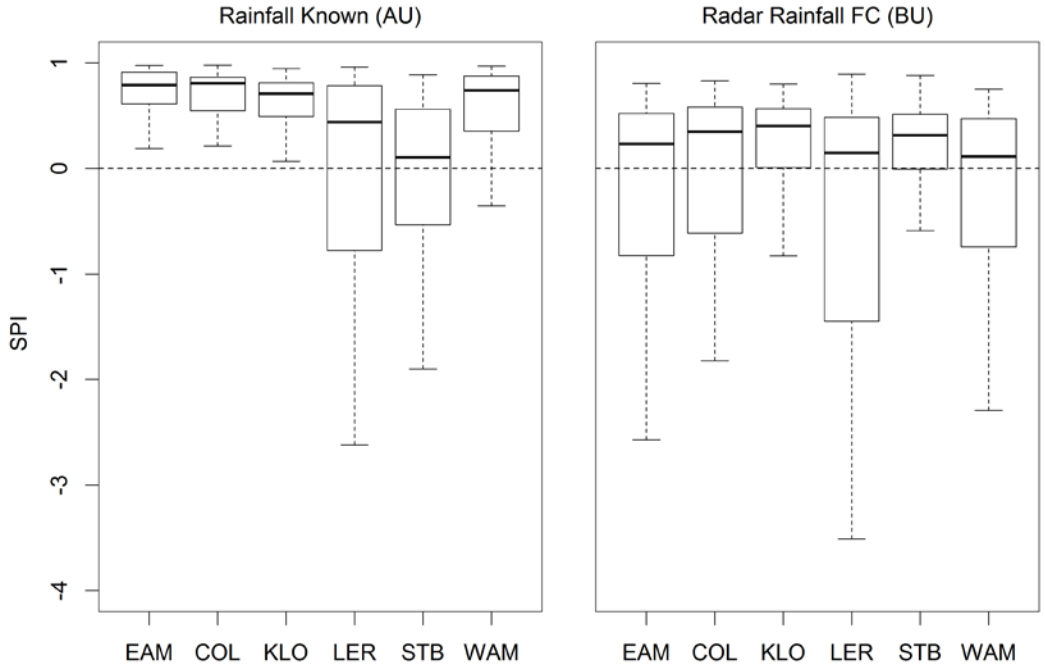

Figure 4. Boxplot of point forecast skill (SPI) for all considered events in the different catchments using rain gauge observations (scenario AU, left) and radar rainfall observations and forecasts (scenario BU, right) as input for runoff forecasting.

Figure 4 shows the point forecast skill SPI obtained in all of the catchments. Skill values larger than zero indicate that the stochastic grey-box models outperformed the benchmark forecast derived from exponential smoothing. This was mostly the case; however, there is a large spread of the results obtained for different rain events.
Very high forecast skill was obtained if rain gauge observations were used as input for runoff forecasting and future rainfall was assumed known (scenario AU). In the more realistic scenario based on radar rainfall forecasts (BU), the runoff forecasting skill was clearly reduced and the spread of the SPI values obtained for different rain events increased. The impact of this reduction on the efficiency of the control scheme is shown in the next section.

Lower forecast skills were generally obtained in the LER and STB sub-catchments for the AU scenario due to the complexity of this part of the catchment with multiple gates and overflow points. Such features are hard to capture with the very simple, data-driven forecast models applied here (Equation 1). In addition, the derivation of flow measurements based on multiple rating curves and with part of the basin outflows not captured by the sensors lead to significant uncertainty of the observed basin inflow.

Unexpectedly, in the STB catchment, the SPI tended to be higher in scenario BU than in scenario AU. This difference was caused by an improved forecast skill of the stochastic grey-box model during dry weather. The estimated uncertainty scaling of the model states (see Equation 1) was larger in scenario BU due to the larger forecast errors caused by the radar rainfall forecasts. As a result, the extended Kalman filter could more easily adapt the dry weather state a_0 of the model (see Equation 1) to the rather strong variations of observed dry weather flows in the STB catchment, leading to improved forecast skill.

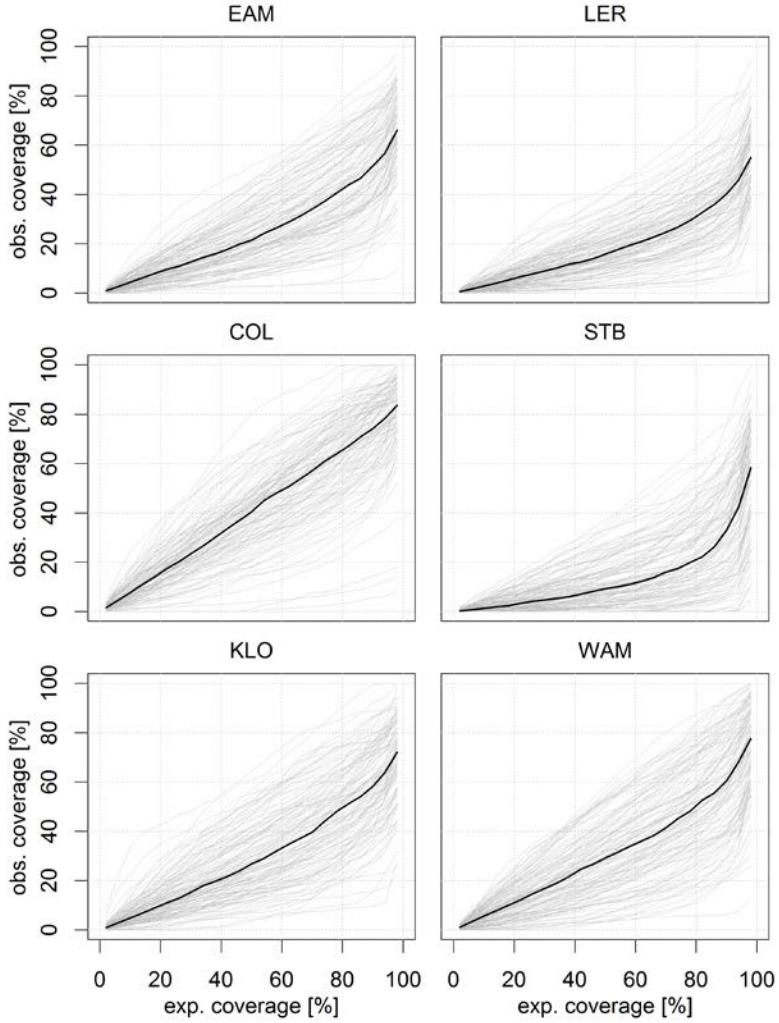

Figure 5. Reliability diagrams (expected vs. observed coverage of the observations) for scenario AU (true (observed) rainfall input from gauges in runoff forecasting) for the different catchments. The results for the single events are marked in grey, while the median coverage rates over all events are marked as black, solid lines.

Figure 5 shows reliability diagrams (expected against observed coverage rates) for scenario AU for the different sub-catchments. The grey lines (showing results for the single events) illustrate that, similar to the point forecasting skill SPI, the reliability of forecasts strongly varied from event to event. Generally, the actual uncertainty of the forecasts was underestimated. The worst results in terms of forecast reliability were obtained in the LER and the STB catchment, where the point forecast skill was also lowest.

Similar results were obtained for scenario BU (Figure 6). However, the reliability of forecasts generally improved as a result of the larger forecast uncertainty.

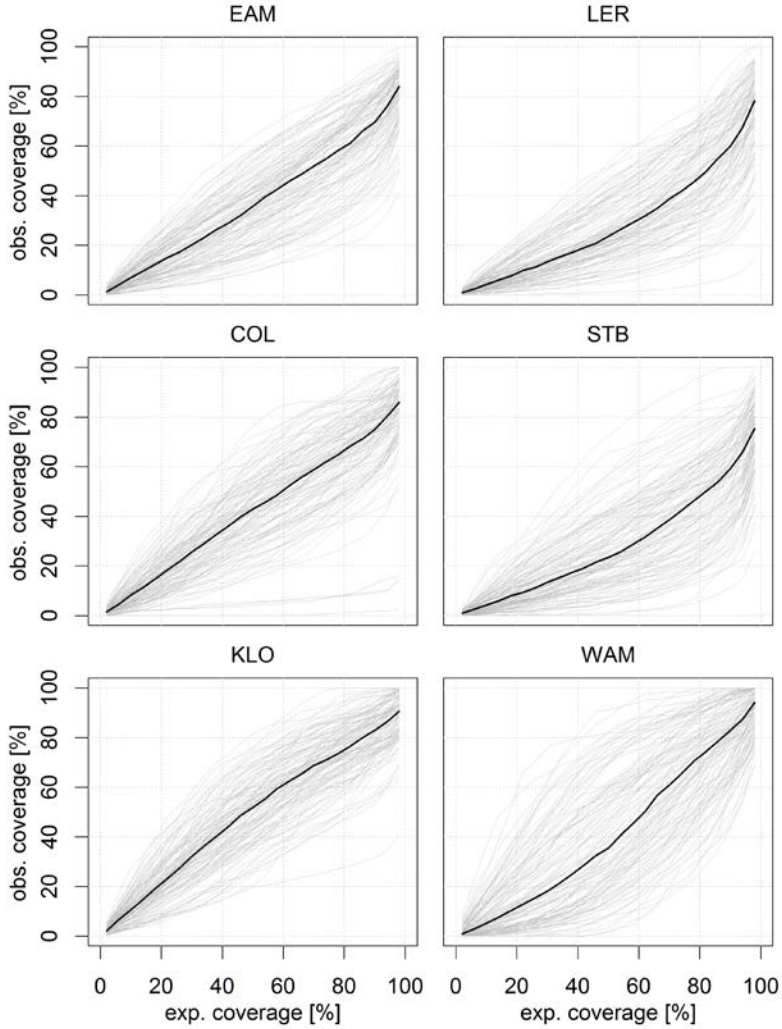

Figure 6. Reliability diagrams (expected vs. observed coverage of the observations) for scenario BU (radar rainfall forecasts as input in runoff forecasting) for the different catchments. The results for the single events are marked in grey, while the median coverage rates over all events are marked as black, solid lines.

Figure 7 shows the ARIL* values obtained for scenarios AU and BU in the different sub-catchments. ARIL* is an expression of the uncertainty of runoff forecasts (see Section 2.2.1). As expected, the ARIL* values strongly increased when radar rainfall forecasts were used as model input in scenario BU instead of rain gauge observations (with assumed perfect rainfall forecast) in scenario AU (Figure 7).

An exception was again the STB catchment, where only a very minor increase in forecast uncertainty was observed for scenario BU. This result fits well with the improved point forecasting skill obtained in this catchment.

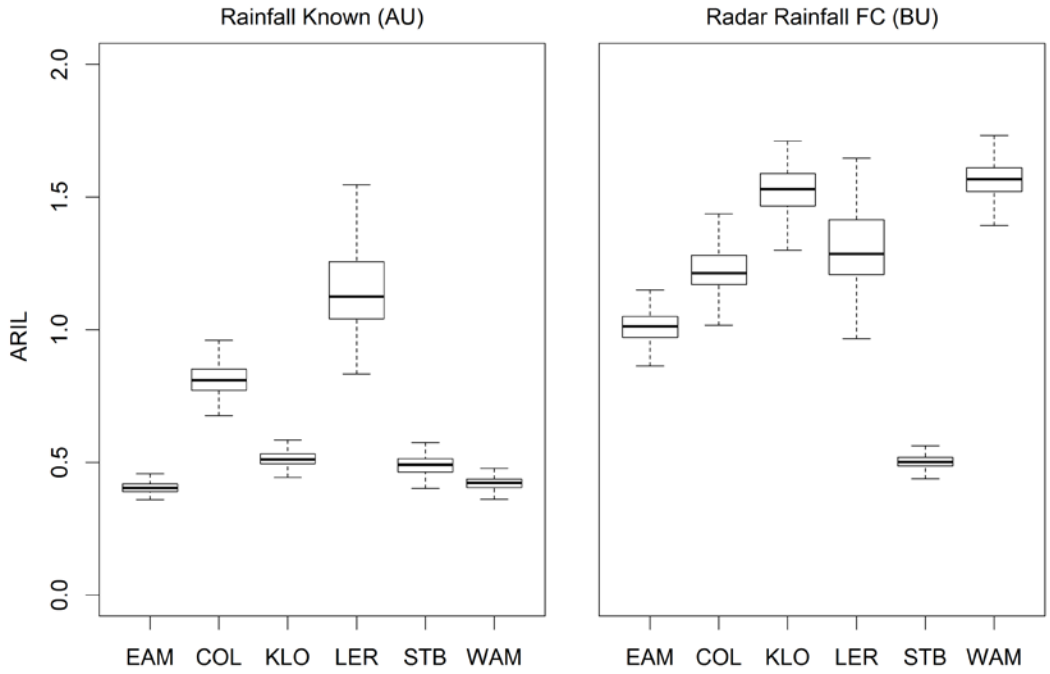

Figure 7. Boxplot of prediction interval width (ARIL*) for all considered events in the different catchments using rain gauge observations (scenario AU, left) and radar rainfall forecasts (scenario BU, right) as input for runoff forecasting.

4.2 Efficiency of System-wide Real-Time Control

The total overflow volumes and cost obtained for the considered scenarios are shown in Figure 8. In the reference scenario REF, overflow occurred for 87 of the considered rain events, leading to a total overflow volume of $0.95 \cdot 10^6 m^3$ (Figure 8, left) and $12.0 \cdot 10^6$ units of overflow cost (Figure 8, right). .

Including forecast information in the control scheme in all cases lead to a strong reduction of overflow volumes and cost. As expected, overflow volumes and cost were smallest for scenarios AU and ANU because the future rainfall was considered known during the generation of runoff forecasts. Control efficiency was reduced if radar rainfall measurements and forecasts were used as input to the stochastic runoff forecasting models (for example, scenario ANU yielded 15 % lower overflow volume and 20 % lower overflow cost than scenario BNU). Nevertheless, in scenarios BU and BNU, the amount of overflow was also greatly reduced compared to the reference scenario REF.

The results obtained by the system-wide control scheme improved further if the uncertainty of the runoff forecasts was accounted for. The total overflow cost (i.e., the objective function of the control scheme) and volume in scenario AU were reduced by 33 % compared to scenario ANU (Figure 8). In scenario BU, the total overflow volume was reduced only minimally compared to scenario BNU (Figure 8). This result was caused by a strong increase in forecast uncertainty at control point KLO. As a result, the optimization routine frequently assigned high outflows to this control point (reducing overflow volumes almost to zero), while outflows from STP were frequently minimized (leading to a strong increase of overflow volumes at this point). However, the total overflow cost in this scenario was reduced by 20 %

compared to scenario BNU, meaning that CSO were diverted from more to less sensitive recipients.

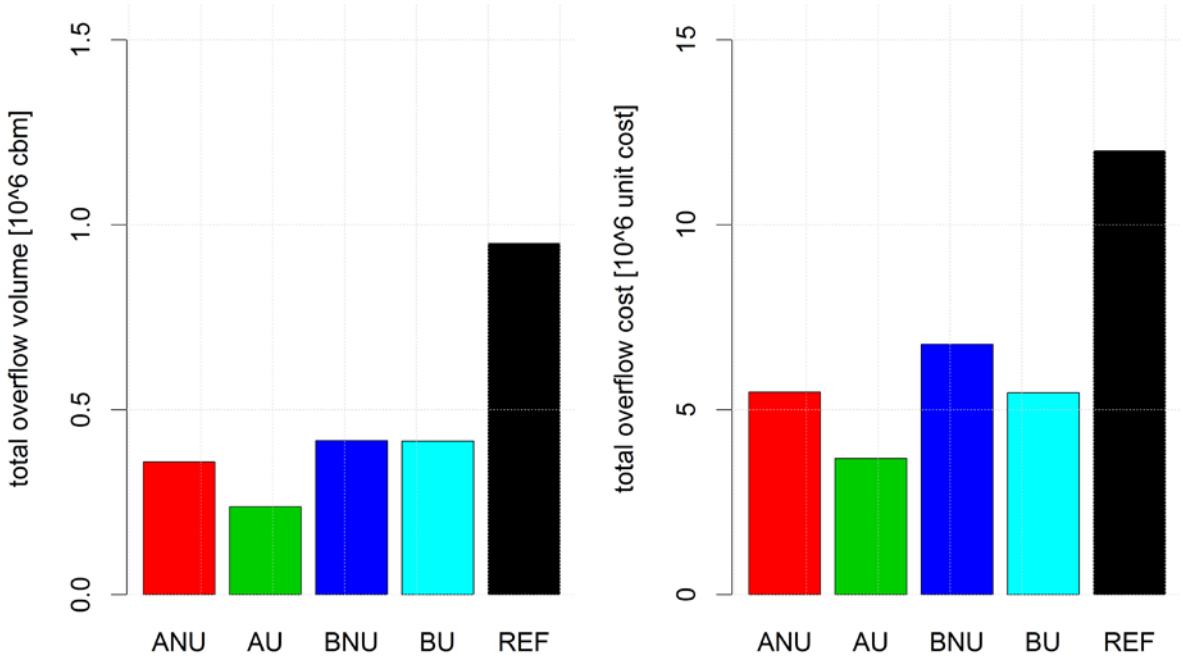

Figure 8. Total overflow volume (left) and cost (right) over all events and catchments in the different scenarios.

5 Discussion

5.1 Dependency of Runoff Forecast Skill on Catchment and Rainfall Input

On average, the stochastic grey-box models outperformed the exponential smoothing benchmark in all of the considered sub-catchments. However, the forecast skill varied strongly between catchments and rain events.

If future rainfall was assumed to be known (scenario AU), the highest forecast skill was obtained for the smaller catchments (EAM, COL, WAM – see Figure 4), where a reservoir cascade could suitably describe the runoff processes. For the more complex catchments, forecasts could be improved if somewhat more complex model structures were considered (Del Giudice et al. 2015a, Löwe et al. 2014a). However, simple models are desirable for online purposes (see the discussion in Harremoës and Madsen (1999)) and the work of Del Giudice et al. (2015a) demonstrated only limited improvement of the predictions beyond a certain level of model complexity.

The skill of the runoff forecasts (SPI, Figure 4) was strongly reduced and varied more between events if radar rainfall forecasts were used as model input (scenario BU) instead of a perfect rainfall forecast derived from gauge measurements (scenario AU). The decrease in forecast skill was most pronounced for the smallest considered sub-catchment (WAM) and less pronounced for the larger sub-catchments such as KLO. This behaviour was caused by the shorter concentration time in smaller catchments, where a runoff forecast for two hours into the future is strongly affected by the uncertainty of the rainfall forecast.

5.2 Reliability of Runoff Forecasts

Figure 5 and Figure 6 compare expected and observed coverage rates for forecasts of runoff volume on a 120 minute horizon. We identified a general tendency for the runoff forecasts to be unreliable. For example, a 90 % prediction interval covered less than 70 % of the observations in all of the sub-catchments in scenario AU.

The main reason for this result was that the stochastic grey-box approach aims to model runoff forecast uncertainty for a multitude of forecast horizons in a single model structure. This approach has the advantage of providing us with an intrinsic quantification of the correlation between forecasts for different horizons, but the model structure is currently not adapted to account for the different effects occurring at different forecast horizons.

Forecast variance increases nonlinearly from short forecast horizons (where the updating of the model to current observations has a strong influence on forecast quality) to longer forecast horizons (where uncertainty from rainfall input and model structure affects the runoff forecast most). The stochastic differential equations in Equation 1, however, assume that forecast variance increases linearly with lead time because the variance of an increment $\Delta\omega_t$ of the Wiener process driving the noise term directly corresponds to the considered time increment Δt. As a result, the stochastic forecast models tended to be reliable on short forecast horizons and unreliable on longer forecast horizons (not shown, but demonstrated in Löwe et al., 2014b).

We identified the following options for addressing this problem in the grey-box modelling framework in the future:

- Different forecast models could be applied for different forecast horizons. While this option would yield reliable forecasts, it would also lead to a strong increase in the number of parameters that need to be identified, and it would not provide the description of correlation between forecast horizons. The identification of forecast distributions of runoff volumes would then require the application of copulas (Madadgar et al., 2014; Papaefthymiou and Kurowicka, 2009) or recursive estimates of the correlation of forecast errors for different horizons (Löwe et al., 2014b, Pinson et al., 2009) to link the stochastic flow forecasts for different horizons.
- A scaling factor depending on forecast lead time could be introduced in the diffusion term of the state equations (Equation 1) and identified as a parameter in the automatic calibration routine. This option seems preferable, as it could be easily integrated in the grey-box modelling approach.

Another interesting result was that higher coverage rates were observed for scenario BU, where radar rainfall forecasts were used as input to the forecast models, than for scenario AU. The parameter estimation procedure identifies the uncertainty scaling for the model states (σ_i) based on how many observations are located how far from the centre of the forecasted distribution (see Löwe et al. (2014b)). During rain periods, runoff forecast errors are much larger if radar rainfall is used as an input to the models, leading to a strong increase in the uncertainty parameters in the model and to increased forecast uncertainties. These, in turn, lead to an increased reliability of the model during dry weather periods, explaining the more reliable pattern observed in Figure 6.

This issue can also be related to a deficiency in the structure of the stochastic grey-box model because only a single parameter σ_i is used in Equation 1 to scale the forecast uncertainty. Alternative formulations of the diffusion term should distinguish between dry weather and rain periods.

5.3 Forecast Uncertainty and System-wide Real-Time Control

The results shown in Figure 8 indicate that there is a clear benefit in using forecast information in the system-wide control algorithm. All scenarios that apply forecast information (AU, ANU, BU and BNU) yield much lower overflow volumes and cost than the reference scenario REF.

In addition, accounting for the uncertainty of runoff forecasts in the system-wide control algorithm has proven beneficial. The reduction in total overflow cost (comparing scenarios AU and ANU as well as BU and BNU) was comparable in magnitude to the increase in total overflow cost caused by the uncertainty of radar rainfall forecasts (comparing scenarios AU and BU as well as ANU and BNU).

The results also showed some limitations of the setup. Replacing perfect rainfall forecasts (scenarios AU and ANU) by radar rainfall forecasts (BU and BNU) decreased runoff forecast skill and strongly increased runoff forecast uncertainty at KLO. This resulted in high forecasted overflow cost at this point and a prioritization of outflows from KLO over those from STP (see Figure 3), strongly increasing overflow volumes at STP. Although the total overflow cost in the system could be reduced, such effects may be undesirable and can be mitigated by an adjustment of the CSO unit cost.

Generally, DORA prioritizes outflow from overflow points where runoff forecast uncertainty is high over overflow points where runoff forecast uncertainty is low. This is desirable because free storage volume is kept available at points where little is known about the future runoff, while storage volume at other control points is used to the fullest. It is, however, important that realistic estimates of forecast uncertainty are identified. In particular, combinations of over- and underestimation of forecast uncertainty at different control points are expected to negatively impact the performance of the control scheme.

5.4 General Applicability of the Setup

The aim of the article was to provide a proof of concept for a forecast- and optimization-based RTC setup that takes forecast uncertainty into account. The setup was demonstrated in a case study involving six different sub-catchments in which the performance of the runoff forecasting models was tested by comparing with observations. The process of generating stochastic runoff forecasts over a horizon of 2 hours and identifying set points using the DORA algorithm required approximately 1 minute on a standard PC (Intel i7-4930k) and is thus well feasible within a control time step of 2 minutes.

The sub-catchments had different sizes and structures (Table 1), and they therefore behaved differently hydraulically. In addition, flow observations were far from perfect and, in most of the catchments, were affected by changes in pumping discharges (Section 3.3.1 and Appendix C). These conditions correspond well to what we would expect in other urban catchments. The skilful forecasts that were obtained for most of the sub-catchments suggest that the forecast setup can be transferred to other catchments.

Current limitations of the setup are that rather unreliable forecasts are obtained for long forecast horizons (Section 5.2) and that only a very simple model structure is considered, while including effects from, e.g., overflow structures located upstream from the control point may well improve the forecast skill in some sub-catchments (Sections 2.1 and 4.1). Conversely, the radar rainfall forecasts provided as model input in our case study were incomplete. In particular, no forecast information was available for horizons beyond 90 minutes. We would therefore expect somewhat better rainfall forecasts and thus better performance of the runoff forecasts in other catchments with more complete rainfall forecasts.

The derivation of inflow measurements using the water balance of the control points proved problematic in terms of operational reliability because each inflow measurement depended on the correct operation of multiple sensors. In fact, we were able to use only 98 out of 171 relevant rain events in our data period as a result of sensors failing at one or multiple control points. This problem can be avoided by installing redundant level sensors or dedicated flow measurements. Männig and Lindenberg (2013) demonstrated that a reliable operation of a control system can also be achieved with a large number of 13 control points and more than 100 in-sewer measurements.

The effect of forecast uncertainty on the optimization-based control scheme was tested for the first time in an urban setting in this study. Raso et al. (2014) demonstrated the value of considering forecast uncertainty in reservoir operation. As we applied a full-scale catchment in our case study, our results provide a strong indication that optimization-based control schemes should consider forecast uncertainty. Nevertheless, this result needs to be verified in further studies and catchments.

6 CONCLUSIONS

A forecast-based, stochastic optimization setup was presented for system-wide real-time control of combined sewer systems aimed at reducing combined sewer overflows. The setup combined stochastic grey-box models for probabilistic forecasting of urban runoff online and the risk-based optimization algorithm DORA that accounts for forecast uncertainty and impact cost.

In a case study in Copenhagen, Denmark, involving 6 sub-catchments of varying sizes and 7 control points we assessed forecast performance by comparing runoff forecasts to measurements and by testing the efficiency of the control scheme in simulations. We conclude that:

1. Accounting for forecast uncertainty in the system-wide control positively affected the results of the control scheme. In the simulation study performed in this work, the reduction of total overflow cost resulting from the consideration of forecast uncertainty was comparable to the increase of total overflow cost resulting from the uncertainty of radar rainfall forecasts (comparing simulation results for the case of a perfect, rain gauge based rainfall forecast to a real-world radar rainfall forecast).
2. Higher uncertainty of the runoff forecast at a control point leads to a higher priority of this control point in DORA. It is therefore important to identify realistic estimates of forecast uncertainty. In particular, for a robust performance of DORA, forecast uncertainty must not be underestimated at some control points and overestimated at others.

3. Using radar rainfall forecasts as input to the stochastic runoff forecasting models instead of perfect rainfall forecasts based on rain gauge measurements lead to a significant decrease in runoff forecast skill. Nevertheless, an exponential smoothing model used as the benchmark forecast was outperformed in all of the considered sub-catchments. In addition, the control scheme yielded much better results than in the reference case where optimization was performed without forecast information, i.e., based on the current basin fillings only.
4. Models that forecast the inflow to the control points could be set up, although direct inflow measurements were not available for most control points. Inflow measurements were derived using the water balance of the storage basins and were in several cases strongly influenced by pumping discharges. The stochastic grey-box models were capable of handling the resulting noisy flow measurements. However, the considered measurements must be ensured to fully capture the water balance at a control point.
5. Stochastic runoff forecasting models need to consider a nonlinear increase of forecast uncertainty with forecast lead time when generating multistep forecasts.
6. Deriving flow measurements from a multitude of sensors implies that each measurement depends on the correct operation of multiple sensors. This can severely impact the reliability of the control setup, a problem that can easily be mitigated by installing redundant sensors in the most suitable locations during the implementation of the RTC system.

The present study has provided a proof of concept for considering forecast uncertainty in a risk-based optimization scheme for RTC of urban drainage systems. Future work should focus on improving rainfall forecasts as well as the development of libraries of runoff forecasting models, where the model structure performing best for a given control point can be selected automatically.

7 ACKNOWLEDGMENTS

This research has been financially supported by the Danish Council for Strategic Research, Programme Commission on Sustainable Energy and Environment through the Storm- and Wastewater Informatics (SWI) project. The catchment and flow data were kindly provided by Copenhagen Utility Company (HOFOR). We thank the Danish Meteorological Institute (DMI) and Michael Rasmussen and Søren Thorndahl from Aalborg University (Dept. of Civil Engineering) for providing data from the C-Band radar at Stevns. Luca Vezzaro was an industrial postdoc financed by the Innovation Fund Denmark under the project "MOPSUS – Model predictive control of urban drainage systems under uncertainty".

8 REFERENCES

Achleitner, S, Stefan Fach, T Einfalt, and W Rauch. 2009. "Nowcasting of Rainfall and of Combined Sewage Flow in Urban Drainage Systems." Water Science and Technology 59 (6): 1145–51. doi:10.2166/wst.2009.098.

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen, H., Nguyen, V.T. V, 2013. Impacts of climate change on rainfall extremes and urban drainage systems: A review. Water Sci. Technol. 68, 16–28. doi:10.2166/wst.2013.251

Bechmann, H, M K Nielsen, H Madsen, and N K Poulsen. 1999. "Grey-Box Modelling of Pollutant Loads from a Sewer System." Urban Water 1 (1): 71–78.

Bennett, N D., B F W W Croke, G G, J H A Guillaume, S H Hamilton, A J Jakeman, S Marsili-Libelli, L T H Newman, J P Norton, C Perrin, S A Pierce, B Robson, R Seppelt, A A Voinov, B D Fath, and V Andreassian. 2013. "Characterising Performance of Environmental Models." Environmental Modelling & Software 40: 1–20. doi:10.1016/j.envsoft.2012.09.011.

Borsányi, P., Benedetti, L., Dirckx, G., De Keyser, W., Muschalla, D., Solvi, A.-M., Vandenberghe, V., Weyand, M., Vanrolleghem, P.A., 2008. Modelling real-time control options on virtual sewer systems. J. Environ. Eng. Sci. 7, 395–410. doi:10.1139/S08-004

Breinholt, A., Santacoloma, P. A., Mikkelsen, P. S., Madsen, H., Grum, M., & Nielsen, M. K. (2008). Evaluation framework for control of integrated urban drainage systems. In Proceedings of the 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK.

Breinholt, A, F O Thordarson, J K Møller, M Grum, P S Mikkelsen, and H Madsen. 2011. "Grey-Box Modelling of Flow in Sewer Systems with State-Dependent Diffusion." Environmetrics 22 (8): 946–61. doi:10.1002/env.1135.

Breinholt, A., J.K. Møller, H. Madsen, P.S. Mikkelsen. 2012. "A formal statistical approach to representing uncertainty in rainfall-runoff modelling with focus on residual analysis and probabilistic output evaluation - distinguishing simulation and prediction". J. Hydrol., 472-473, 36-52.

Brown, R G, and R F Meyer. 1961. "The Fundamental Theorem of Exponential Smoothing." Operations Research 9 (5): 673–85. http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=7689664&site=ehost-live.

Carstensen, J, M K Nielsen, and H Strandbæk. 1998. "Prediction of Hydraulic Load for Urban Storm Control of a Municipal WWT Plant." Water Science and Technology 37 (12): 363–70.

Del Giudice, D., Reichert, P., Bares, V., Albert, C., Rieckermann, J., 2015a. Model bias and complexity - Understanding the effects of structural deficits and input errors on runoff predictions. Environ. Model. Softw. 64, 205–214. doi:10.1016/j.envsoft.2014.11.006

Del Giudice, D., Löwe, R., Madsen, H., Mikkelsen, P.S., Rieckermann, J., 2015b. Comparison of two stochastic techniques for reliable urban runoff prediction by modeling systematic errors. Water Resour. Res. 51, 5004–5022. doi:10.1002/2014WR016678

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., Tait, S., 2012. Assessing uncertainties in urban drainage models. Phys. Chem. Earth, Parts A/B/C 42-44, 3–10. doi:10.1016/j.pce.2011.04.007

Dotto, C.B.S., G. Mannina, M Kleidorfer, L Vezzaro, M. Henrichs, D.T. McCarthy, G Freni, W Rauch, and A Deletic. 2012. "Comparison of Different Uncertainty Techniques in Urban Stormwater Quantity and Quality Modelling." Water Research 46 (8): 2545–58. doi:10.1016/j.watres.2012.02.009.

938 Freni, G, G. Mannina, and Gaspare Viviani. 2009. “Assessment of Data Availability Influence
939 on Integrated Urban Drainage Modelling Uncertainty.” Environmental Modelling and
940 Software 24 (10): 1171–81. doi:10.1016/j.envsoft.2009.03.007.
941
942 Gneiting, T, and Adrian E Raftery. 2007. “Strictly Proper Scoring Rules, Prediction, and
943 Estimation.” Journal of the American Statistical Association 102 (477): 359–78.
944 doi:10.1198/016214506000001437.
945
946 Grum, M, E Longin, and J J Linde. 2004. “A Flexible and Extensible Open Source Tool for
947 Urban Drainage Modelling: www.WaterAspects.org.” In Proceedings of the 6th International
948 Conference on Urban Drainage Modelling. Dresden, Germany.
949 http://prswww.essex.ac.uk/mantissa/reports/essex/MANTISSAFinalReportJ.pdf.
950
951 Harremoës, P., Madsen, H., 1999. Fiction and reality in the modelling world–Balance
952 between simplicity and complexity, calibration and identifiability, verification and
953 falsification. Water Sci. Technol. 39, 1–8.
954
955 Hemri, S., F. Fundel, and M. Zappa. 2013. “Simultaneous Calibration of Ensemble River
956 Flow Predictions over an Entire Range of Lead Times.” Water Resources Research 49: 6744–
957 55. doi:10.1002/wrcr.20542.
958
959 Iacus, S M. 2008. Simulation and Inference for Stochastic Differential Equations : With R
960 Examples. Springer Series in Statistics. Milan, Italy.
961
962 Jin, X, C Y Xu, Q Zhang, and V P Singh. 2010. “Parameter and Modeling Uncertainty
963 Simulated by GLUE and a Formal Bayesian Method for a Conceptual Hydrological Model.”
964 Journal of Hydrology 383 (3-4): 147–55. doi:10.1016/j.jhydrol.2009.12.028.
965
966 Jørgensen, H K, S Rosenørn, H Madsen, and P S Mikkelsen. 1998. “Quality Control of Rain
967 Data Used for Urban Runoff Systems.” Water Science and Technology 37 (11): 113–20.
968
969 Juhl, R, N R Kristensen, P Bacher, J K Møller, and H Madsen. 2013. “CTSM-R User Guide.”
970 Kgs. Lyngby, Denmark. Technical University of Denmark. http://CTSM.info.
971
972 Kavetski, Dmitri, George Kuczera, and Stewart W. Franks. 2006. “Bayesian Analysis of Input
973 Uncertainty in Hydrological Modeling: 1. Theory.” Water Resources Research 42 (March):
974 W03407. doi:10.1029/2005WR004368.
975
976 Krämer, S, M Grum, H R Verworn, and A Redder. 2005. “Runoff Modelling Using Radar
977 Data and Flow Measurements in a Stochastic State Space Approach.” Water Science and
978 Technology 52 (5): 1–8.
979
980 Krämer, Stefan, Lothar Fuchs, and H R Verworn. 2007. “Aspects of Radar Rainfall Forecasts
981 and Their Effectiveness for Real Time Control -The Example of the Sewer System of the City
982 of Vienna.” Water Practice & Technology 2 (2). doi:10.2166/WPT.2007042.
983
984 Kristensen, N. R., Madsen, H., and Jørgensen, S. B. (2004) Parameter estimation in stochastic
985 grey-box models. Automatica, 40(2), 225–237.

Löwe, R. 2014. “Probabilistic Forecasting for On-Line Operation of Urban Drainage Systems.” Ph.D. thesis, Department of Applied Mathematics and Computer Science, Technical University of Denmark (DTU). Kgs. Lyngby, Denmark.

Löwe, R, S Thorndahl, P S Mikkelsen, M R Rasmussen, and H Madsen. 2014a. “Probabilistic Online Runoff Forecasting for Urban Catchments Using Inputs from Rain Gauges as Well as Statically and Dynamically Adjusted Weather Radar.” Journal of Hydrology 512: 397–407. doi:10.1016/j.hydrol.2014.03.027.

Löwe, R, P S Mikkelsen, and H. Madsen. 2014b. “Stochastic Rainfall-Runoff Forecasting: Parameter Estimation, Multi-Step Prediction, and Evaluation of Overflow Risk.” Stochastic Environmental Research and Risk Assessment 28: 505–16. doi:10.1007/s00477-013-0768-0.

Madadgar, S., Moradkhani, H., Garen, D., 2014. Towards improved post-processing of hydrologic forecast ensembles. Hydrol. Process. 28, 104–122. doi:10.1002/hyp.9562

Maestre, J. M., L Raso, P J van Overloop, and B. De Schutter. 2013. “Distributed Tree-Based Model Predictive Control on a Drainage Water System.” Journal of Hydroinformatics 15 (2): 335–47. doi:10.2166/hydro.2012.125.

Mahnke, W., S. H. Leitner, M. Damm. 2009. „OPC Unified Architecture”. Springer. ISBN:978-3-540-68898-3

Männig, F., and M. Lindenberg. 2013. “Betriebserfahrungen mit der Abflusssteuerung des Dresdner Mischwassernetzes.” KA Korrespondenz Abwasser, Abfall 2013 (12): 1036–43. doi:10.3242/kae2013.12.001.

Meffert, K. et al.: „JGAP - Java Genetic Algorithms and Genetic Programming Package”. http://jgap.sf.net

Mollerup, A.L., D. Thornberg, P. S. Mikkelsen, N.B. Johansen, G. Sin. 2013. ”16 Years of Experience with Rule Based Control of Copenhagen’s Sewer System”. 11th Int. IWA Conf. on Instrumentation, Control and Automation (ICA), Narbonne/France, 18-20 September 2013. Extended abstract, 4 pp

Mollerup, A.L., P.S. Mikkelsen, D. Thornberg, G. Sin. 2015. ”Regulatory control analysis and design for sewer systems”. Environmental Modelling & Software. Environ. Model. Softw. 66, 153–166. doi:10.1016/j.envsoft.2014.12.001

Moradkhani, H, C M DeChant, and S Sorooshian. 2012. “Evolution of Ensemble Data Assimilation for Uncertainty Quantifcation Using the Particle Filter Markov Chain Monte Carlo Method.” Water Resources Research 48 (12). doi:10.1029/2012WR012144.

Murphy, A H, and R L Winkler. 1977. “Reliability of Subjective Probability Forecasts of Precipitation and Temperature.” Journal of the Royal Statistical Society.Series C (Applied Statistics) 26 (1): 41–47. http://www.jstor.org.globalproxy.cvt.dk/stable/2346866.

1034 Nielsen, M K, and T Önnerth. 1995. “Improvement of a Recirculating Plant by Introducing
1035 STAR Control.” Water Science & Technology 31 (2): 171–80.
1036
1037 Nielsen, N H, C Ravn, and N Mølbye. 2010. “Implementation and Design of a RTC Strategy
1038 in the Sewage System in Kolding , Denmark.” In Proceedings of NOVATECH 2010. Lyon,
1039 France.
1040
1041 Pabst, M, J Alex, M Beier, C Niclas, M Ogurek, D Peikert, and M Schütze. 2011. “ADESBA
1042 - A New General Global Control System Applied to the Hildesheim Sewage System.” In
1043 Proceedings of the 12th International Conference on Urban Drainage. Vol. 11–16 Sept. Porto
1044 Alegre, Brazil.
1045
1046 Papaefthymiou, G., Kurowicka, D., 2009. Using copulas for modeling stochastic dependence
1047 in power system uncertainty analysis. IEEE Trans. Power Syst. 24, 40–49.
1048 doi:10.1109/TPWRS.2008.2004728
1049
1050 Pinson, P., Madsen, H., Papaefthymiou, G., 2009. “From Probabilistic Forecasts to Wind
1051 Power Production. Production” 51–62. doi:10.1002/we
1052
1053 Pleau, M, H Colas, P Lavalle, G Pelletier, and R Bonin. 2005. “Global Optimal Real-Time
1054 Control of the Quebec Urban Drainage System.” Environmental Modelling and Software 20
1055 (4). Elsevier: 401–13. doi:10.1016/j.envsoft.2004.02.009.
1056
1057 Puig, V, G Cembrano, J Romera, J Quevedo, B Aznar, G Ramón, and J Cabot. 2009.
1058 “Predictive Optimal Control of Sewer Networks Using CORAL Tool: Application to Riera
1059 Blanca Catchment in Barcelona.” Water Science and Technology 60 (4). IWA Publishing:
1060 869–78. doi:10.2166/wst.2009.424.
1061
1062 Raso, L., Schwanenberg, D., van de Giesen, N.C., van Overloop, P.J., 2014. Short-term
1063 optimal operation of water systems using ensemble forecasts. Adv. Water Resour. 71, 200–
1064 208. doi:10.1016/j.advwatres.2014.06.009
1065
1066 Rauch, W, K Seggelke, Rebekah Brown, and Peter Krebs. 2005. “Integrated Approaches in
1067 Urban Storm Drainage: Where Do We Stand?” Environmental Management 35 (4): 396–409.
1068 doi:10.1007/s00267-003-0114-2.
1069
1070 Renard, B, Dmitri Kavetski, George Kuczera, Mark Thyer, and S Franks. 2010.
1071 “Understanding Predictive Uncertainty in Hydrologic Modeling: The Challenge of Identifying
1072 Input and Structural Errors.” Water Resources Research 46 (5): W05521.
1073 doi:10.1029/2009WR008328.
1074
1075 Schellart, A.N.A., Shepherd, W.J., Saul, A.J., 2011. Influence of rainfall estimation error and
1076 spatial variability on sewer flow prediction at a small urban scale. Adv. Water Resour. 45, 65–
1077 75. doi:10.1016/j.advwatres.2011.10.012
1078
1079 Schellart, A N A, Sara Liguori, Stefan Krämer, A J Saul, and M A Rico-Ramirez. 2014.
1080 “Comparing Quantitative Precipitation Forecast Methods for Prediction of Sewer Flows in a

Small Urban Area." Hydrological Sciences Journal 59 (7): 1418–36. doi:10.1080/02626667.2014.920505.

Schilling, W, and L Fuchs. 1986. "Errors in Stormwater Modeling-a Quantitative Assessment." Journal of Hydraulic Engineering 112 (2): 111–23.

Schilling, W. 1991. "Rainfall Data for Urban Hydrology: What Do We Need?" Atmospheric Research 27 (1): 5–21.

Schütze, M., Campisano, A., Colas, H., Schilling, W., Vanrolleghem, P.A., 2004. Real time control of urban wastewater systems—where do we stand today? J. Hydrol. 299, 335–348. doi:10.1016/j.jhydrol.2004.08.010

Seggelke, K, R Löwe, T Beeneken, and L Fuchs. 2013. "Implementation of an Integrated Real-Time Control System of Sewer System and Waste Water Treatment Plant in the City of Wilhelmshaven." Urban Water Journal 10 (5): 330–41. doi:10.1080/1573062X.2013.820331.

Storm- and Wastewater Informatics, SWI. 2015: Project website, http://www.swi.env.dtu.dk/, visited 12th March 2015.

Sun, S., Bertrand-Krajewski, J.L., 2013. Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method. Water Resour. Res. 49, 5381–5394. doi:10.1002/wrcr.20444

Thordarson, F O, A Breinholt, J K Møller, P S Mikkelsen, M Grum, and H Madsen. 2012. "Uncertainty Assessment of Flow Predictions in Sewer Systems Using Grey Box Models and Skill Score Criterion." Stochastic Environmental Research and Risk Assessment 26 (8): 1151–62. doi:10.1007/s00477-012-0563-3.

Thorndahl, S, and M R Rasmussen. 2013. "Short-Term Forecasting of Urban Storm Water Runoff in Real-Time Using Extrapolated Radar Rainfall Data." Journal of Hydroinformatics 15 (3): 897–912. doi:10.2166/hydro.2013.161.

Thorndahl, S, T S Poulsen, T Bøvith, M Borup, M Ahm, J E Nielsen, M Grum, M R Rasmussen, R Gill, and P S Mikkelsen. 2013. "Comparison of Short Term Rainfall Forecasts for Model Based Flow Prediction in Urban Drainage Systems." Water Science and Technology 68 (2): 472–78. doi:10.2166/wst.2013.274.

Thorndahl, S, J E Nielsen, and M R Rasmussen. 2014. "Bias Adjustment and Advection Interpolation of Long-Term High Resolution Radar Rainfall Series." Journal of Hydrology 508 (January): 214–26. doi:10.1016/j.jhydrol.2013.10.056.

Todini, E., 2008. A model conditional processor to assess predictive uncertainty in flood forecasting. Int. J. River Basin Manag. 6, 123–137. doi:10.1080/15715124.2008.9635342

Vanrolleghem, P A, L Benedetti, and J Meirlaen. 2005. "Modelling and Real-Time Control of the Integrated Urban Wastewater System." Environmental Modelling and Software 20 (4): 427–42. doi:10.1016/j.envsoft.2004.02.004.

Van Steenbergen, N., Ronsyn, J., Willems, P., 2012. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. Environ. Model. Softw. 33, 92–105. doi:10.1016/j.envsoft.2012.01.013

Vezzaro, L, and P S Mikkelsen. 2012. "Application of Global Sensitivity Analysis and Uncertainty Quantification in Dynamic Modelling of Micropollutants in Stormwater Runoff." Environmental Modelling and Software 27-28: 40–51. doi:10.1016/j.envsoft.2011.09.012.

Vezzaro, L, and M Grum. 2014. "A Generalised Dynamic Overflow Risk Assessment (DORA) for Real Time Control of Urban Drainage Systems." Journal of Hydrology 515: 292–303. doi:10.1016/j.jhydrol.2014.05.019.

Vezzaro, L, M.L. Christensen, C Thirsing, M. Grum, and P S Mikkelsen. 2014. "Water Quality-Based Real Time Control of Integrated Urban Drainage Systems: A Preliminary Study from Copenhagen, Denmark." Procedia Engineering 70: 1707–16. doi:10.1016/j.proeng.2014.02.188.

Vieux, B E, and J E Vieux. 2005. "Statistical Evaluation of a Radar Rainfall System for Sewer System Management." Atmospheric Research 77 (1). Elsevier: 322–36. doi:10.1016/j.atmosres.2004.10.032.

Vrugt, J A, C G H Diks, Hoshin V Gupta, W Bouten, and J M Verstraten. 2005. "Improved Treatment of Uncertainty in Hydrologic Modeling: Combining the Strengths of Global Optimization and Data Assimilation." Water Resources Research 41 (1): W01017. doi:10.1029/2004WR003059.

Vrugt, J A, C J F Ter Braak, C G H Diks, and G Schoups. 2013. "Hydrologic Data Assimilation Using Particle Markov Chain Monte Carlo Simulation: Theory, Concepts and Applications." Advances in Water Resources 51: 457–78. doi:10.1016/j.advwatres.2012.04.002.

Weerts, A., Winsemius, H.C., Verkade, J.S., 2011. Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). Hydrol. Earth Syst. Sci. 15, 255–265. doi:10.5194/hess-15-255-2011

Willems, P., Ollson, J., Arnbjerg-Nielsen, K., Beecham, S., Pathirana, A., Gregersen, I.B., Madsen, H., Nguyen, V.T. V, 2012. Impacts of Climate Change on Rainfall Extremes and Urban Drainage. IWA Publishing, London, United Kingdom.

9 APPENDIX A

see Supporting Material

10 APPENDIX B

see Supporting Material

11 APPENDIX C

see Supporting Material