Short-term forecasting of urban storm water runoff in real-time using extrapolated radar rainfall data

S. Thorndahl and M. R. Rasmussen

ABSTRACT

Model-based short-term forecasting of urban storm water runoff can be applied in real-time control of drainage systems in order to optimize system capacity during rain and minimize combined sewer overflows, improve wastewater treatment or activate alarms if local flooding is impending. A novel online system, which forecasts flows and water levels in real-time with inputs from extrapolated radar rainfall data, has been developed. The fully distributed urban drainage model includes auto-calibration using online in-sewer measurements which is seen to improve forecast skills significantly. The radar rainfall extrapolation (nowcast) limits the lead time of the system to 2 hours. In this paper, the model set-up is tested on a small urban catchment for a period of 1.5 years. The 50 largest events are presented.

Key words | auto-calibration, forecasting, radar, real-time control, urban drainage modeling

S. Thorndahl (corresponding author)
M. R. Rasmussen
Aalborg University,
Department of Civil Engineering,
Aalborg,
Denmark
E-mail: *st@civil.aau.dk*

INTRODUCTION

In the future, climate change will have severe impacts on the function of urban drainage systems. Rain events will intensify and water levels of receiving waters will rise. Furthermore, catchment surfaces will grow in size and density due to increased urbanization. This will increase the risks of urban flooding and increase the discharge of untreated (or poorly treated) waste and storm water to receiving waters. The hygienic, environmental, and economic consequences are obvious. One solution is to increase drainage system treatment and storage capacity. Another solution, which, in many cases, is more economically feasible, is to apply modern technology to optimize the existing sewer systems. This can be implemented by forecast-based real-time control of drainage systems, in order to better utilize the capacity of drainage systems or to activate alarm systems if flooding or overflow is impending.

Real-time flow forecasting has been studied quite intensively with regard to rivers and streams (e.g. Collier 2007). These systems are primarily implemented by running models online to predict certain water levels and risk of flooding. Many rural systems have a runoff time larger than 12–24 hours, and it is thus sufficient to run the flow forecasting models with real-time rainfall data (either radar *Quantitative Precipitation Estimates* (QPE) or rain gauge data). Of course, if a very long lead time is required, rainfall inputs from a *Numerical Weather Prediction* (NWP) model can extend lead times up to a few days, however with considerable uncertainty. Urban drainage systems are different from the rural systems since the systems are smaller with significantly shorter runoff time. Therefore, in order to forecast flows or water levels in urban drainage systems, the rainfall input to the runoff models needs to be predicted, in order to extend lead times beyond the time of concentration of the system in question. One way of forecasting with very short lead times (commonly called nowcasting) is to extrapolate observed radar rainfall as, for example, done by Rinehart & Garvey (1978), Golding (1998), Germann & Zawadzki (2002), and Bowler *et al.* (2006).

The concept of flow forecasting in urban drainage systems is a rather unexplored area of research, but in recent years more focus has been given to the area (Achleitner *et al.* 2009; Leitão *et al.* 2010; Liguori *et al.* 2012; Liguori & Rico-Ramirez 2012; Schellart *et al.* 2012; Thorndahl

doi: 10.2166/hydro.2013.161

et al. 2012a; Wang *et al*. 2012). The increase in publications in recent years might be due to the fact that given the quality of today's models, it is possible to do reliable real-time control based on model results.

This study will investigate the potential for radar-based short-term flow forecasting in drainage systems – using a small urban catchment as its case. The concept will demonstrate how flow forecasting can be implemented in real-time operation. Real-time control of the drainage system based on the forecast is not considered in the present approach. Whether this system could actually prevent urban flooding, combined sewer overflow (CSO), or activate alarms is thus unconsidered.

Since November 2010, the system has been operated online in real-time, with the purpose of studying real-time application of drainage models. The objective is to investigate the reliability and robustness of these types of online systems, before initiating actual real-time control relying on radar-based short-term forecasts.

The system consists of three different parts: (1) a radar-based *Quantitative Precipitation Forecast* (QPF) model, (2) an urban drainage model (in this case the commercial *MOUSE* model by DHI), and (3) an *auto-calibration* (optimization) routine based on online flow measurements.

In order to perform reliable flow forecasting, a well-calibrated runoff model is required. The model needs to be tested with historical inputs, doing hindcast simulations before being implemented in real-time operation. In terms of forecasting, it is assumed that the best possible performance is obtained by implementing a real-time auto-calibration (optimization) routine running continuously, updating the system variables according to flow measurements in the system. This concept is also known as real-time data assimilation. When the systems and the model have the same state at the time of the initiation of the forecast, the best possible flow forecast is ensured. This method also enables the system to handle uncertainties in the rainfall intensity levels estimated by the radar. If the rain intensities are too small or too large, the auto-calibration routine will ensure that the runoff is adjusted accordingly. When analyzing historical data, it is preferable to apply radar data which have been calibrated against a number of rain gauges; either as daily mean field bias correction, e.g. Krajewski & Smith (2002), Seo *et al*. (1999), Seo *et al*. (2011), Krajewski *et al*. (2011) or event-based bias calibration as suggested by Thorndahl & Rasmussen (2012). This is not possible to do in real-time radar data bias calibration however, unless data from a dense network of rain gauges are available in real-time, which they are not in this case. Therefore, the bias calibration of the radar data is integrated into the overall auto-calibration of the drainage model against flow observations. Normally, two calibration procedures would be initiated separately. Since it is the prediction of the flow in the drainage system which is of major interest in this approach, it is considered a reasonable assumption to omit the middle step of certifying the exact rainfall input as long as the flow output from the runoff model corresponds to the observed flow. It is most certain that more reliable forecasts can be obtained by implementing the auto-calibration routine, compared with simulating the model in real-time with default parameter values, no updating of system states, and uncalibrated radar rainfall inputs.

In the present approach, the drainage system is simulated with a fully distributed urban drainage model (MOUSE), which moreover provides the possibility to predict CSOs to receiving waters or local surcharge or flooding of manholes. Other approaches such as Thorndahl *et al*. (2012a) have focused on forecasting flow with a much simpler (linear reservoir) runoff model, limiting the model to forecasting the flow in one single point. Obviously, a simple (non-distributed) model has the benefit of short computation time compared with a distributed model, and this is an advantage in auto-calibration in which multiple simulations for each time step would be required.

The paper is structured as follows. In the Methods and Data section, the radar data, flow observations, and catchment are introduced, along with a description of methods applied for radar data extrapolation, runoff modeling, and auto-calibration. The Results of Auto-Calibration section presents the performance of the auto-calibration routine, and in the Results of Short-Term Flow Forecasting section the runoff forecasts are evaluated and compared with observed data. The final two sections present discussions and conclusions, respectively.

METHODS AND DATA

Radar rainfall data

The online flow forecasting system is currently utilizing data from a C-band Doppler weather radar owned and operated by the Danish Meteorological Institute (DMI). The radar has a range of 240 km, and according to Gill *et al.* (2006) the quantitative range is limited to a maximum of 75–100 km. The radar data have been adjusted for attenuation according to Gill *et al.* (2006), and ground and sea clutter have been removed using methods proposed by Gill (2007), Bøvith *et al.* (2006), and Bøvith (2008). The data are sampled with a temporal resolution of 10 minutes, and a pseudo CAPPI (Constant Altitude Plan Position Indicator) product is generated based on nine elevation scans. The spatial resolution of the Cartesian pseudo CAPPI product is $2 \times 2\ \mathrm{km}^2$. The radar is located approx. 50 km north of the study catchment.

The data calibration of the radar is based on a traditional Z–R-relationship (Reflectivity-Rain rate) using standard parameters ($A = 200$, $B = 1.6$, Marshall & Palmer 1948). In other studies, e.g. Thorndahl *et al.* (2012a, 2012b), the radar is adjusted using a daily mean field bias correction (Krajewski & Smith 2002). A daily bias correction is not possible in this case however, due to real-time application as presented in the introduction. Furthermore, since the runoff model is auto-calibrated against flow-measurements in real-time, the radar data calibration becomes an integrated part of the overall calibration of the drainage model, eliminating the need for an individual *a priori* calibration of the rainfall input.

Rainfall nowcast model

The radar extrapolation algorithm has been developed at Aalborg University, Denmark and is currently running in real-time operation on several radars and radar composites. It is based on the TREC algorithm (TRacking of Echoes with Correlation, Rinehart & Garvey 1978) and the CO-TREC (Continuous-TREC, Li *et al.* 1995; Mecklenburg *et al.* 2000).

The model is structured as follows:

In the first step of the model (*Preprocessing*), the radar data are spatially averaged over a number of time steps (in this case $N = 4$). A reflectivity cutoff value is implemented ($Z_{\min} = 10$ dBz) in order to reduce noise produced by low reflectivity. In the second step (*Correlation Analysis*), a 2D cross-correlation analysis is applied on subset boxes of the radar image. These boxes are displaced in pairs and correlation coefficients are calculated for all possible pairs of boxes. In order for the correlation coefficient to be calculated, a *minimum percentage* of the pixels in the individual boxes must contain values (rainfall). The maximum correlation coefficients are found in order to derive translation vectors (cf. the TREC algorithm) corresponding to the movement of the rain.

In order to avoid defective or unrealistic advection (especially around the edges of the radar image), the field of vectors is smoothed by implementation of different filters and constraints on the variations of speeds and directions (CO-TREC). An example of the vector field in TREC and CO-TREC is shown in Figure 1. In the final step of the nowcast model (*Extrapolation*), the latest calibrated radar image is extrapolated a number of time steps (in this case 12 corresponding to 2 hours) into the future based on the field of vectors using a linear backward advection scheme. Lead times of 30, 60 and 120 minutes are applied in this study.

The model is also described and applied in Thorndahl *et al.* (2009, 2010, 2011, 2012a, 2012b). It does not include growth/decay terms and is solely an extrapolation of the observed rain. Reasonable forecast performance, within a lead time of 0–120 minutes, using C-band radar data has been shown in the above-mentioned publications. The critical success index (CSI) which indicates whether the model is able to predict rain in the same pixels as observed (CSI = 1 corresponds to 100% of the pixels) has mean values of approx. 0.7, 0.6, and 0.4 for the 30-, 60- and 120-min nowcasts, respectively. The reason for the low CSI with regard to the 120-min nowcast is the fact that the 120-min lead time sometimes includes rain which has not yet been detected by the radar. This is primarily a problem during convective events (with fast growth and decay rates) as opposed to stratiform conditions.

Case catchment and urban drainage model

The study is completed using the small urban catchment Frejlev as case. This catchment is unique as it is equipped

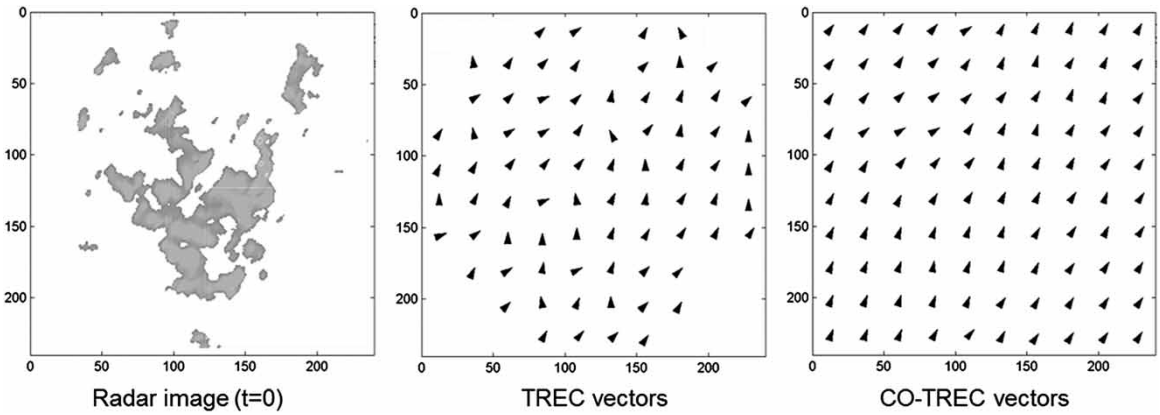

Figure 1 | Concept of the CO-TREC nowcasting method.

with two electromagnetic flow meters (Schaarup-Jensen *et al.* 1998). One flow meter has a diameter of 300 mm and measures the low flows during dry weather and light rain, while the other flow meter has a diameter of 1,000 mm and measures the flow during heavier rain events when the small flow meter is full-running. The lower limits for the flow meters are approx. 0.01 m^3/s for the 300 mm meter and 0.07 m^3/s for the 1,000 mm meter.

The urban catchment has an area of approx. 0.8 km^2 with 2,000 inhabitants and the drainage system is a partly separated and partly combined sewer system. At the location of the flow meters (the monitoring station), both storm and waste water is discharged in the same pipe. Downstream from the monitoring station is a CSO structure, which discharges to the local small stream, Hasseris, 10–15 times a year. The system is shown in Figure 2.

Two rain gauges are located in the catchment. They are a part of the national Danish rain gauge network of the Danish Water Pollution Committee. The catchment could easily be covered by one single radar pixel, however the location of the radar and the grid necessitate use of three radar pixels of 2×2 km^2. If the catchment were larger, such as presented in Thorndahl *et al.* (2012a), the runoff simulations could benefit from a distributed rainfall input, but the spatial variability between the three pixels is negligible in this case.

The drainage system is modeled with the MOUSE model (DHI 2009) with a time-area surface runoff model as well as a hydro-dynamical pipe flow model based on the 1D Saint Venant Equations. The model is fully distributed, which means that the model can also predict local

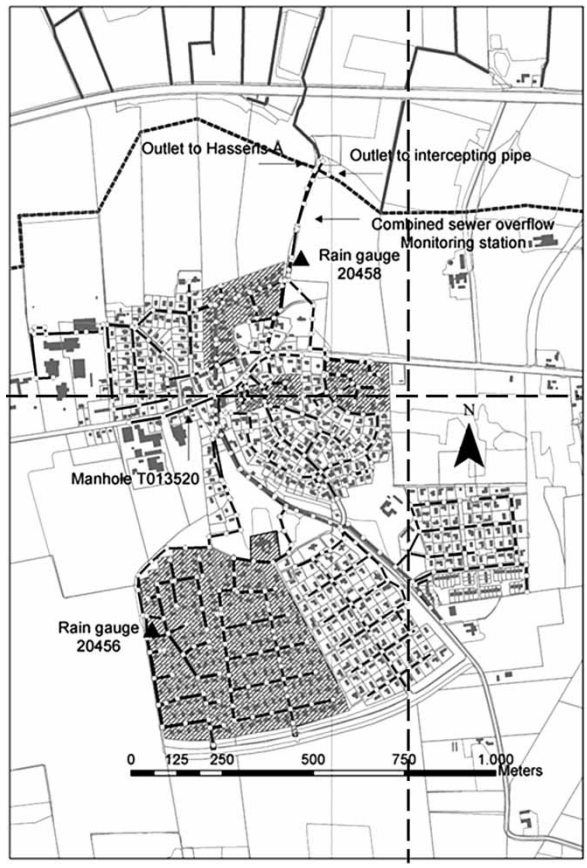

Figure 2 | The Frejlev catchment. Hatched area marks the separate sewer system. The white area is a combined system. The dashed lines correspond to the radar pixels (2×2 km grid).

flooding of manholes, storage filling, and capacity problems. The rainfall input to the model is handled individually for each sub-catchment (manhole connection). On top of the MOUSE model, a MATLAB shell has been developed in order to run the optimization and import radar data to the MOUSE model. In the real-time application, the model is executed every 10 minutes corresponding to the temporal resolution of the radar data.

The most important parameters in the MOUSE model are the ones selected for the auto-calibration procedure. They are described in detail in the following section.

The catchment and the MOUSE model set-up has previously been studied doing hindcast simulations with rain gauge data as input, e.g. Schaarup-Jensen *et al.* (2009), Thorndahl & Willems (2008) and Thorndahl *et al.* (2008). Thorndahl *et al.* (2012b) compared radar nowcast models and NWP models for short-term flow forecasting, also applying the Frejlev catchment as case (on only three selected events, however). Thorndahl *et al.* (2011) investigated performance of the radar nowcast model on the catchment using five different events.

Optimization routine and real-time operation

The urban runoff model with radar rainfall input is auto-calibrated every 30 minutes by using runoff measurements from the past 12 hours. As described in the introduction, the individual calibration of the radar rainfall is neglected and the optimization is made between the uncorrected radar rainfall data and the observed runoff. In this way, an update of the overall system to the current state is excluded.

A 4-hour time window is used to simulate the flow forecast (event though the rainfall forecast is only 2 hours – this ensures that the storm water is able to discharge from the system). It would be preferable to initiate the auto-calibration for each model run (every 10 minutes), but at this time the optimization algorithm is too computationally demanding to be implemented in real-time simulating on a single standard PC. The simulations performed in 10-min intervals therefore use the parameter set from the last available auto-calibration run.

The current system consists of one PC downloading radar data and performing the radar forecast, one PC performing auto-calibration and optimization of parameters as well as downloading and saving flow data, and finally one PC performing the flow forecast. The auto-calibration is therefore running independently of the flow forecast.

The optimization algorithm is based on a Quasi-Newton method for non-linear systems. The concept of this optimization is to find minima on a multidimensional surface. In this case, the algorithm has been implemented using the Broyden (1970), Fletcher (1970), Goldfarb (1970), and Shanno (1970), BFGS formulation as this has been proven to converge quickly with few iterations. This method requires calculation of the first order partial derivatives, which in this case is done numerically. The optimization criterion is in this case defined by:

$$f = 1 - \exp\left(\frac{-\frac{1}{N}\sum_{j=1}^{N}\left((M_i - O_i) - \overline{(M_i - O_i)}\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2}\right) \quad (1)$$

where the numerator is the negative variance of the residuals between model (M) and observations (O), and the denominator is the variance of the observations (O). This measure was shown by Thorndahl *et al.* (2008) and Beven & Freer (2001) to ensure low peak and volume errors fitting flow time series. The best possible model performance is found by:

$$f_{\min} = \min(f) \quad (2)$$

Three different parameters are included in the auto-calibration, the *overall system scaling factor* (*SSF*), the *surface concentration time,* t_c, and the *dry weather flow* (*DWF*). Variations of the three parameters are considered to be able to cover the most zero- and first-order errors and are therefore suitable for the auto-calibration routine. Furthermore, Thorndahl *et al.* (2008) classified these parameters to be the most sensitive for the model in question.

The overall mass balance of the system is certified using the following equation:

$$Q(t) = \varphi \cdot \psi \cdot \sum_{c=1}^{N} i_c(t) \cdot A_c \cdot \beta_c \quad (3)$$

where $Q(t)$ if the flow output from the whole system as a function of time; N is the total number of sub-catchments; $i_c(t)$ is the uncorrected radar rainfall intensity in sub-catchment c as a function of time; A_c is the total area of sub-catchment c; β_c is the percentage of impervious area in sub-catchment c; φ is the *hydrological reduction factor* which in the MOUSE model (DHI 2009) determines the percentage of the impervious areas which contributes to the runoff; and ψ is the mean field radar bias.

Since no individual real-time calibration (bias adjustment) of the radar rainfall data is performed, the mean field radar bias, ψ and the hydrological reduction factor, φ are combined to one *overall SSF* which certifies the mass balance between the uncorrected radar rainfall input and the observed flow in the drainage system.

$$\mathrm{SSF} = \varphi \cdot \psi \tag{4}$$

SSF is used to correct for uncertainties in the radar calibration as well as in the runoff model and can assume values both larger and smaller than 1. Generally, this parameter ensures the correct volumes in the system. If the radar rainfall data are unbiased, SSF will correspond to the hydrological reduction factor. Thorndahl *et al.* (2006) have calibrated the hydrological reduction factor parameter to a value of 0.45 on the same catchment using local (unbiased) rain gauge data.

The *surface concentration time*, t_c, is used to ensure the optimal temporal flow variations. Other parameters related to the temporal variations could have been included as well, e.g. the pipe roughness or local head loss, but to ensure a low number of iterations in the optimization algorithm, only one parameter governing the temporal variations is applied. The default calibrated value for the concentration time is 6 minutes.

Since the system is a combined sewer with runoff of both storm and waste water, the DWF is used to correct the runoff volumes when no rain is observed by the radar. Schaarup-Jensen & Rasmussen (2004) found this parameter to be 0.135 m^3/(day inhabitant) on the Frejlev catchment. Auto-calibration on the DWF is disabled when there is a rainfall input to the model and until 6 hours after the rain has stopped (in order to make sure all the storm water has discharged). Since this publication only addresses the runoff during rain, and the DWF is not changed from the standard setting during rain, the DWF is omitted throughout the rest of the paper for clarity reasons.

For each auto-calibration run, these three parameters values are estimated and values are tracked as shown in Table 1.

Simulated and forecasted events

As mentioned in the introduction this system has been operated in real-time since November 2010. Until June 2012, a total of 52 rainfall events (larger than 3 mm according rain gauge no. 20458, Figure 2) have been simulated and forecasted. Some longer periods with outage due to hardware problems occurred and a period has been discarded due to snow and snowmelt runoff, hence the list of events is not complete for the 19 months of operation. Two events were discarded due to radar data outage, which gives the list of 50 events presented in Table 1.

For clarity reasons, it is chosen to present complete events in the paper, even though the forecast systems run in 16-hour windows (12-hour hindcast and 4-hours forecast). The events therefore have different durations and only one optimized parameter set is shown for each event.

Evaluation measures

The observed, modeled and forecasted data are compared using three different measures: The *relative volume error* between observed and modeled volumes; the *relative peak error* between observed and modeled peaks; and the *Nash Sutcliffe Efficiency* (*NSE*).

The *relative volume error* is negative if the model underestimates the volume and positive if the model overestimates the volume compared with the observations. Values are calculated in percent.

The *relative peak error* between observed and modeled peaks (maximum values for each event) is calculated in two locations corresponding to the two flow meters. Values are calculated in percent.

Table 1 | Summary of the auto-calibration routine for the 50 events

Event no.	Simulation start time	Simulation end time	Relative volume error (%)	Relative peak error, Flow gauge ø300 (%)	Relative peak error, Flow gauge ø1,000 (%)	Optimization parameter, f (–)	Optimized overall system scaling factor, SSF (mm/mm)	Optimized conc. time, t_c (min)	Rainfall depth (mm)
1	12.11.2010 23:00	13.11.2010 19:00	0.2	−1.7	−14.6	0.39	0.31	24	23.2
2	03.02.2011 14:00	04.02.2011 10:00	0.9	−0.3		0.30	1.24	25	7.6
3	04.02.2011 08:00	05.02.2011 04:00	0.0	30.1		0.33	0.42	40	3.2
4	07.02.2011 04:00	08.02.2011 00:00	2.3	−6.2		0.16	0.70	0	8.8
5	25.02.2011 19:00	26.02.2011 15:00	0.0	0.0		0.51	0.86	40	7.2
6	31.03.2011 10:00	01.04.2011 06:00	13.7	0.5	36.4	0.64	0.55	0	8.8
7	05.04.2011 01:00	05.04.2011 21:00	−0.7	14.1	−7.0	0.63	0.68	0	7.6
8	08.07.2011 00:00	08.07.2011 20:00	43.0	11.8	4.5	0.73	1.23	40	9.6
9	14.07.2011 06:00	15.07.2011 02:00	11.3	2.9	−8.3	0.81	1.58	22	6.6
10	17.07.2011 00:00	17.07.2011 20:00	75.0	−12.3		0.61	0.64	40	4.0
11	21.07.2011 21:00	22.07.2011 17:00	20.8	−6.0	−49.0	0.77	0.92	39	7.8
12	22.07.2011 07:00	23.07.2011 03:00	33.8	1.5	60.0	0.90	2.24	40	3.0
13	14.08.2011 06:00	15.08.2011 02:00	−4.6	2.4		0.43	1.68	1	5.0
14	18.08.2011 18:00	19.08.2011 14:00	−2.5	6.7	−4.4	0.55	1.97	1	14.8
15	23.08.2011 10:00	24.08.2011 06:00	34.9	−6.0		0.46	0.39	1	5.0
16	26.08.2011 12:00	27.08.2011 08:00	71.2	−9.1		0.63	0.37	1	3.8
17	29.08.2011 00:00	29.08.2011 20:00	12.5	−2.1	−49.1	0.76	0.86	39	7.4
18	05.09.2011 00:00	05.09.2011 20:00	3.2	−12.4		0.65	0.39	1	3.6
19	06.09.2011 09:00	07.09.2011 05:00	43.7	2.3	−38.8	0.66	1.55	1	17.2
20	06.09.2011 21:00	07.09.2011 17:00	21.8	−2.5	−41.9	0.65	1.00	17	5.0
22	12.09.2011 04:00	13.09.2011 00:00	37.1	6.2	−16.9	0.67	3.17	2	8.8
23	18.09.2011 01:00	18.09.2011 21:00	23.5	−6.0	−51.3	0.69	0.97	1	9.8
24	20.09.2011 17:00	21.09.2011 13:00	16.4	−3.8	−44.8	0.49	1.42	1	17.6
25	26.09.2011 02:00	26.09.2011 22:00	10.6	−9.5		0.58	1.31	0	4.4
26	09.10.2011 23:00	10.10.2011 19:00	7.8	−4.7		0.53	2.19	1	3.8
27	17.10.2011 20:00	18.10.2011 16:00	4.9	4.3		0.63	3.95	1	10.2
28	18.10.2011 15:00	19.10.2011 11:00	4.7	−5.1		0.63	1.77	1	8.2
29	26.11.2011 20:00	27.11.2011 16:00	0.0	−5.7		0.37	0.77	2	3.8
30	02.12.2011 22:00	03.12.2011 18:00	0.0	−1.8	−54.2	0.46	0.79	22	7.0
31	31.12.2011 18:00	01.01.2012 14:00	4.9	1.9		0.25	0.75	0	3.6
32	03.01.2012 00:00	03.01.2012 20:00	13.1	−6.6		0.63	0.58	1	5.4
33	03.01.2012 13:00	04.01.2012 09:00	15.2	−5.9		0.75	0.54	40	9.0
35	11.01.2012 19:00	12.01.2012 15:00	−2.8	−0.9		0.60	0.80	28	4.8
36	18.01.2012 03:00	18.01.2012 23:00	21.7	−1.1		0.37	0.38	0	4.4
37	05.04.2012 17:00	06.04.2012 13:00	2.4	−58.5		0.43	0.20	40	3.2
38	08.04.2012 18:00	09.04.2012 14:00	3.6	−10.3		0.46	0.20	40	3.4
39	09.04.2012 19:00	10.04.2012 15:00	8.8	−44.4		0.50	0.20	40	3.2
40	21.04.2012 18:00	22.04.2012 14:00	43.2	3.0	−49.4	0.63	0.40	1	5.4

(*continued*)

Table 1 | continued

Event no.	Simulation start time	Simulation end time	Relative volume error (%)	Relative peak error, Flow gauge ø300 (%)	Relative peak error, Flow gauge ø1,000 (%)	Optimization parameter, f (–)	Optimized overall system scaling factor, SSF (mm/mm)	Optimized conc. time, t_c (min)	Rainfall depth (mm)
41	26.04.2012 00:00	26.04.2012 20:00	31.8	8.0	6.8	0.63	0.37	1	3.2
42	15.05.2012 19:00	16.05.2012 15:00	−0.2	1.5		0.63	0.76	1	5.0
43	16.05.2012 06:00	17.05.2012 02:00	14.8	0.4	65.6	0.63	0.73	1	6.0
44	31.05.2012 18:00	01.06.2012 14:00	3.4	−9.4		0.70	0.81	1	3.4
45	01.06.2012 17:00	02.06.2012 13:00	49.7	−1.8	20.8	0.64	1.17	1	9.2
46	08.06.2012 01:00	08.06.2012 21:00	12.7	−6.7	−42.2	0.92	0.68	40	13.4
47	10.06.2012 01:00	10.06.2012 21:00	26.4	0.7	44.8	0.63	2.73	1	3.2
48	15.06.2012 23:00	16.06.2012 19:00	9.4	−36.5		0.63	0.60	0	3.2
49	22.06.2012 00:00	22.06.2012 20:00	−5.1	−2.4	−25.8	0.65	0.83	1	9.6
50	24.06.2012 05:00	25.06.2012 01:00	−0.1	−20.4		0.52	0.27	40	3.8
51	24.06.2012 23:00	25.06.2012 19:00	19.2	−6.5	−39.3	0.63	4.06	1	3.6
52	30.06.2012 18:00	01.07.2012 14:00	50.0	12.2	18.3	0.43	3.99	27	12.6
Mean			16.1	−3.9	−11.6	0.58	1.14	14.1	7.0

The NSE (Nash & Sutcliffe 1970) is a measure of how well the modeled data (M) fits the observed data (O):

$$\mathrm{NSE} = 1 - \frac{\frac{1}{N}\sum_{i=1}^{n}(O_i - M_i)^2}{\frac{1}{N}\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad (5)$$

The NSE can range from $-\infty$ to 1, where NSE = 1 is the best possible fit, NSE = 0 indicates that the model is as accurate as the mean of the observations, and NSE < 0 indicates worse performance than the mean of the observations. In this case, NSE is applied to compare both flows, volumes, and peaks. The NSE is somewhat similar to the optimization measure, f, (Equation (1)), and 1-NSE could have been applied as the optimization measure. Beven & Freer (2001) indicated that the exponential part of Equation (1) would perform better in estimating peaks as well as be able to secure low mass balance errors.

RESULTS OF AUTO-CALIBRATION

Table 1 presents results of the auto-calibration routine for the 50 events. The top left panel plot of Figures 3 and 4 shows the observed vs. modeled volumes and observed and modeled peaks for the 50 events, respectively. In Table 1, the general tendency is a slight overestimation of the modeled runoff volume compared with the observed volumes. It is difficult to reduce the error of the mass balances completely however, due to the fact that the flow gauges have minimum thresholds for measuring the flow (approx. 0.01 m^3/s for the 300 mm flow gauge and approx. 0.07 m^3/s for the 1,000 mm flow gauge). This is obvious examining the example presented in the middle panel of Figure 5 in which a number of small local observed peaks are present in the 1,000 mm flow gauge.

The optimized values of the overall SSF (mean of 1.14 for all events) are approximately a factor of 2.5 larger than other investigations (Thorndahl *et al.* 2006, 2008). This is due to the fact that neither calibration nor bias correction has been performed on the radar data. The concept of the auto-calibration routine is that all volume errors can be handled by the *overall SSF* and that no *a priori* radar data calibration are necessary. Since the radar data seem to have a negative bias for the modeled period, the overall SSF will take on larger values than the default value of hydrological reduction factor.

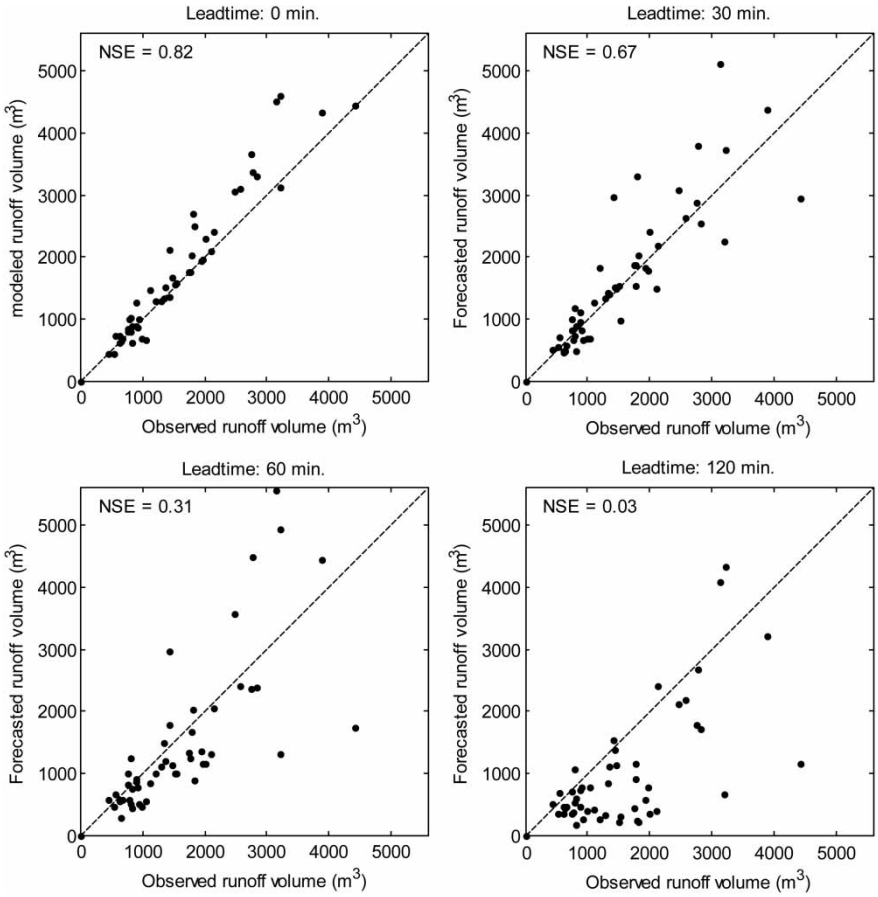

Figure 3 | Top left panel: Observed vs. modeled (optimized) runoff volumes. Other panels: Observed vs. forecasted runoff volumes with different lead times.

With regard to the peak errors (Table 1 and Figure 4 top left panel), there is a small tendency towards underestimation of the peak flow in the model, especially on the very high peaks, this is not considered as a significant problem however. (See examples in Figures 6 and 7). Since the auto-calibration is performed with only two parameters, it is not possible to obtain a better fit of the modeled and observed time series. In Thorndahl *et al.* (2008), stochastic calibration was performed with seven parameters and thus a better fit of the time series was obtained. Currently, however, it is not possible to perform this type of calibration in real-time and thus slightly larger peak and volume errors must be accepted. In some cases, there is also the possibility that the optimization routine has found a local minimum instead of the global minimum due to non-linearities in the model. This can be improved by adjusting the steps which are used to find the first order partial derivatives numerically.

For several of the events, the optimized value of the concentration time corresponds to 0 or 1 min (Table 1). This indicates that the concentration time might not be sufficient in order to optimize the temporal variations in the auto-calibration routine.

RESULTS OF SHORT-TERM FLOW FORECASTING

Figure 3 presents the volume errors for the 30-, 60-, and 120-min forecasts. The 30- and 60-min forecasts have no particular bias (with NSE >0.3), but to some extent have increasing scatter for increasing lead time. The 120-min lead time shows large underestimations of the

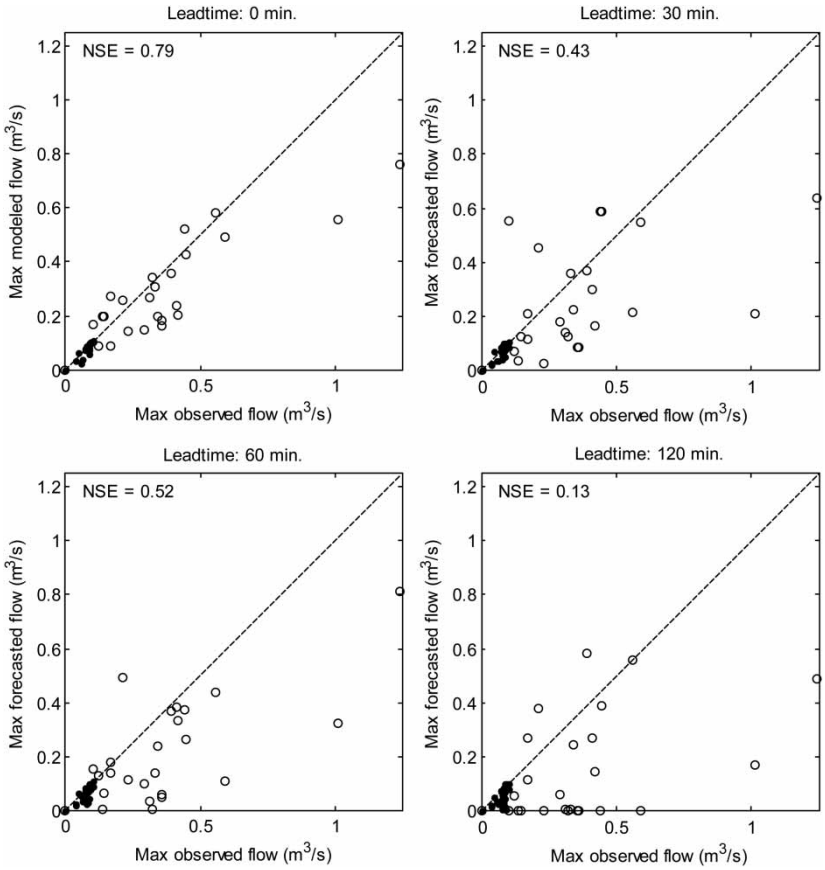

Figure 4 | Top left panel: Observed vs. modeled (optimized) peak flows. Other panels: Observed vs. forecasted peak flows with different lead times. Solid dots mark the flow in the small (ø300 mm flow meter) and the hollow dots mark the large (ø1,000 mm) flow meter.

volumes (NSE close to 0). It is expected that the longest lead time will be associated with the largest uncertainty. Especially convective rainfall events have proven difficult to forecast by extrapolation of radar images with lead times over 1 hour. They often have a fast build-up, a smaller spatial extent compared with stratiform precipitation, and relatively short durations. The same tendencies are present for the peak flow estimates (Figure 4).

Besides increasing errors of volume and peak for increasing lead time, some differences in the temporal variations can be identified between the observations, the optimized model, and the three forecasts (see examples in Figures 5–7). The temporal errors increase for increasing lead time, but no significant positive or negative bias is present. In Figure 5 and Figure 7 for example, the peaks for the longer lead times arrive somewhat later than they actually happened, which is not the case for the event exemplified in Figure 6.

Figure 8 presents a scatterplot of the modeled CSO volumes versus the forecasted CSO volumes. In this case, it is obviously not possible to compare the forecasts with observed data, as the CSO volumes are not measured. Again, the errors increase for increasing lead times. According to the optimized model, an overflow will occur in 23 of the 50 events. The 30- and 60-min forecasts predict 21 and 17 overflow events, respectively. However, using the 120-min forecast, it is only possible to predict nine overflow events. One of these overflow events is a false alarm, indicating that no overflow has been predicted by the optimized model, but the 120-min forecast predicts an overflow.

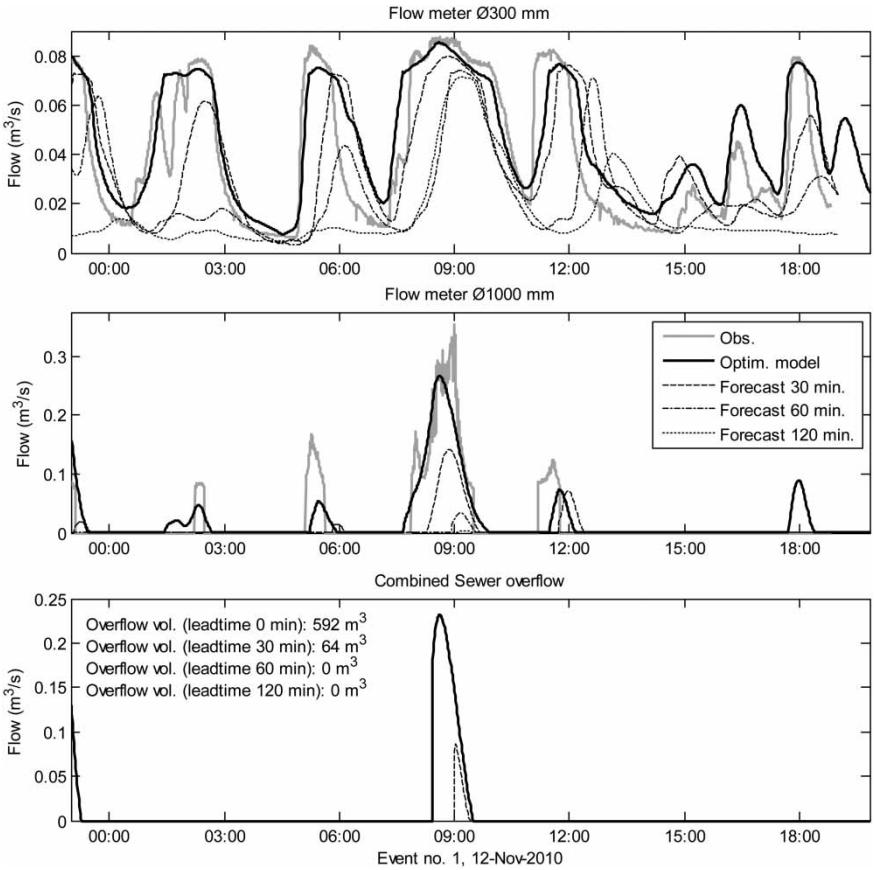

Figure 5 | Example of observed, modeled (optimized), and forecasted flows with different lead times corresponding to the location of the small flow meter (top panel), the large flow meter (middle panel), and the combined sewer overflow (bottom panel). Event 1: 12 Nov 2010.

DISCUSSION

The radar nowcast procedure presented in this paper is rather simple as it is solely an extrapolation of the observed radar rainfall. Since only one single radar is applied, it is impossible to do a reliable forecast beyond a lead time of 2 hours. Even if a mosaic of several radars was applied, it might be difficult to extend lead times further due to the rapid nature of storms (especially convective storms). Several attempts have been made to extend the extrapolation with build-up and decay of storms, but have proven very difficult to work in real-time with convincing consistency (e.g. Mueller & Megenhardt 2001; Roberts & Rutledge 2003). Therefore, it was chosen to hold on to the very simple extrapolation and accept a decrease in the forecast quality as function of increasing lead time.

The difficulties of short-term forecasting at small spatial scales as well as forecasting high rainfall intensities has recently been addressed by Liguori & Rico-Ramirez (2012) who conclude the poorest nowcast performance as the smallest spatial scale (2 km) compared with larger spatial scales (4, 10, 20, 40, and 100 km). Similar conclusions were supported by Germann & Zawadzki (2002).

Another possibility in terms of rainfall forecasting would obviously be to go towards NWP models as presented in Thorndahl *et al.* (2012a, 2012b). In these papers, a maximum lead time of 24 hours was applied, however it should be possible to extend lead times even further (although with increasing uncertainty). A weather model has the advantage that it simulates the complete meteorological conditions of the atmosphere and can thus include the development of storms, convections, orographic effects, etc. The spatial

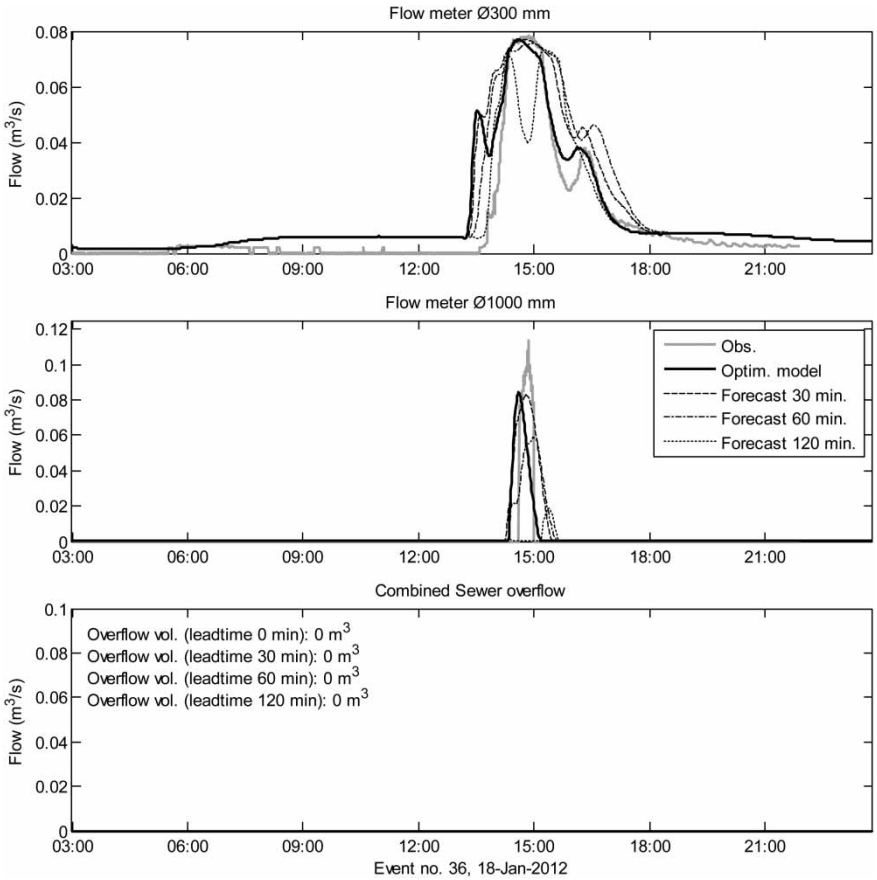

Figure 6 | Example of observed, modeled (optimized), and forecasted flows with different lead times corresponding to the location of the small flow meter (top panel), the large flow meter (middle panel), and the combined sewer overflow (bottom panel). Event 36: 18 Jan. 2012.

and temporal resolutions are, however, coarser compared with the radar nowcast. Furthermore, Thorndahl *et al.* (2012a, 2012b) showed that the weather model DMI-HIRLAM-S05 had a poor performance in lead times from 0 to 6 hours due to the implementation of the boundary and initial conditions of the model and succeeding spin-up effects. Future work includes assimilation of radar rainfall data into the numerical weather models in order to improve the very short-term forecasts. In the case where this is successful, the enhanced weather model will be applied in real-time operation on the Frejlev catchment in combination with the system presented in this paper.

The auto-calibration concept which has been presented ensures consensus between the state of the model and the observations prior to initialization of the flow forecast. This facilitates the best possible flow forecast performance relying on the accuracy of rainfall forecast. It is arguable whether the auto-calibration should be performed for every initialization of the model, every third initialization as executed here, even rarer, or not performed in real-time at all. Since, neither calibration of the quantitative radar rainfall estimates nor the radar nowcast are performed, the authors believe that this auto-calibration is necessary. The concept requires that it is possible to adjust the uncertainties related to the rainfall input by adjustment of hydrological parameters in the drainage model. In this case, in which the *overall SSF* is calibrated to ensure the mass balance of the system, the auto-calibration will in fact correspond to a mean field bias correction of the radar rainfall estimates, however based on integrated flow measurements rather than point rain gauge measurements. Consequently, *a priori* calibration of the radar rainfall estimates is redundant.

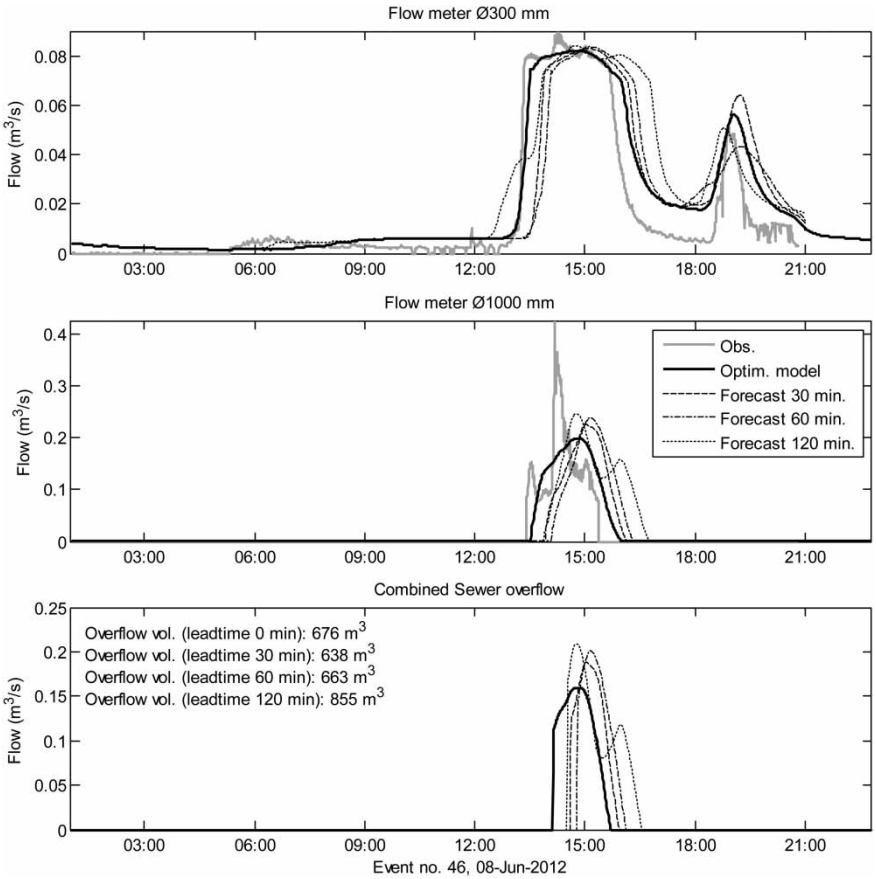

Figure 7 | Example of observed, modeled (optimized), and forecasted flows with different lead times corresponding to the location of the small flow meter (top panel), the large flow meter (middle panel), and the combined sewer overflow (bottom panel) Event 46: 8 June 2010.

This concept is also easier to implement in practice compared with a system in which the radar rainfall estimates are calibrated against online rain gauges in the catchment. The same concept has been presented by Ahm *et al*. (2012).

It is chosen to use the optimization criteria based on the exponential likelihood measure from Thorndahl *et al*. (2008). In the initial phase of this project, different criteria were tested, for example the NSE, a simple mass balance criterion, and a criterion based on a combined log-likelihood function taking into account both observed values and an *a priori* distribution of the parameters (Thorndahl *et al*. 2012a). In this case, it was estimated that the exponential likelihood measure performed the best. As presented in the Results of Auto-Calibration section for some of the events, it was difficult for the auto-calibration routine to optimize the temporal flow variations due to the fact that only the surface concentration time was used to calibrate the model temporally. The model clearly includes a number of non-linearities which makes the optimization process difficult. In order to obtain a better temporal fit, other parameters such as pipe roughness or head losses in manholes could have been implemented in the auto-calibration as well, at the expense of a more complex auto-calibration with ambiguous solutions, however. This problem was discussed by Beven (2006) who addressed it as 'the equifinality thesis'. This would increase the computation time of the auto-calibration routine significantly.

The fact that only two parameters are used for the auto-calibration during rainfall clearly introduces a number of simplifications and suppresses some of the variability in the model output. An increase of calibration parameters and calibration points would increase the degrees of

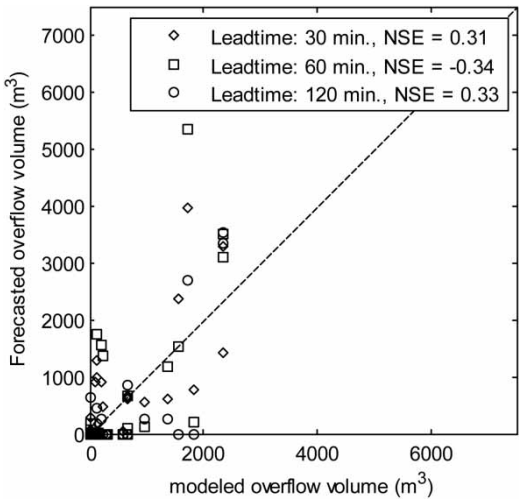

Figure 8 | Modeled (optimized) vs. forecasted combined sewer overflow volumes.

freedom considerably, and make the optimization even more difficult, and therefore it was chosen to settle for the two most sensitive parameters.

There is always a large focus in hydrological modeling on obtaining the perfect fit between observed and modeled hydrographs. Obviously, complete perfect fits have not been presented in this paper, but in terms of the concept of short-term forecasting, acceptable fits have been obtained. In urban hydrological forecasting, it might not be necessary to predict very accurate flows, but to determine whether the flow will exceed a certain threshold (e.g. in an inlet to a waste water treatment plant), or if CSO or local flooding will occur. If urban hydrological forecasting were to be applied in combination with real-time control of drainage systems, the objective would also be to predict when and where the flow or water level would exceed a certain threshold, for example, in order to utilize spare capacity of the drainage system to prevent or minimize CSO or local flooding.

As stated in the introduction, the modeling concept of this paper has been to apply a fully distributed urban drainage model (as opposed to a simplified lumped model). This fully distributed model is the same kind of model used to analyze the system with respect to frequencies of CSOs or return periods of local surcharge or flooding in present and future scenarios. The fully distributed model has the possibility to predict local variability and thus potential local failures of the drainage system, as well as specific loads on basins, pumps, overflow structures, etc. which can be applied in terms of real-time control. The drawback is the computation time. A simplified (non-distributed) model, such as the linear reservoir model, WaterAspects (Grum *et al.* 2004) applied in Thorndahl *et al.* (2012a) has the advantage of a very fast computation and thus, it is possible to run the auto-calibration routine for every time step. The model is only forecasting the flow in one single point of the drainage system however (the inlet to the waste water treatment plant) and no upstream information is available.

In order to develop the concept further, the future work based on this system will be to implement a probabilistic forecast by including uncertainties related to the forecasting system. Thorndahl *et al.* (2008) concluded that the largest uncertainties of a similar set-up of an urban drainage model were related to the rainfall input. Therefore, besides quantifying the uncertainties related to the hydraulics of the drainage model itself, both uncertainties related to radar data calibration as well as the extrapolation of radar images need to be addressed and quantified. By quantifying the output variability of the drainage model (e.g. overflow discharges or manhole water levels), it should be possible to associate risk of failure with the system (e.g. CSO or local flooding). With today's computer power, the simple model approaches as discussed above, might have better potential for implementing probabilistic flow forecasts.

CONCLUSION

This paper has presented an online urban drainage flow forecasting system which makes use of radar nowcasts for generating prognoses of rainfall inputs. The system has been running online since November 2010. A total of 50 rainfall events has been analyzed in the paper.

It is shown that a better quality of forecast is obtained when implementing the auto-calibration procedure in order to condition the model on observations before the initiation of the forecast. This requires that data are present in the period prior to the forecast, otherwise the model will run with un-optimized standard parameter values.

The demonstrated system shows great potential in forecasting urban runoff for real-time control applications. The catchment used as the case study might be too small to really benefit from real-time control, but it is clear that larger systems with larger storage volumes could have an advantage in utilizing the full capacity of the system. Furthermore, the system could be applied for simulating the inlet flow to waste water treatment plants in order to optimize treatment processes during rain. In this case, the runoff time of the whole system can be added to the lead time of the rain, predicting flow or water levels with a larger lead time than 2 hours. In this case, a fully distributed drainage model such as MOUSE might be too detailed and therefore too computationally demanding. A simplified lumped model might work better for these purposes.

ACKNOWLEDGEMENTS

Some of the work presented in this paper has been developed as part of the Storm and Water Informatics (SWI) project funded by The Danish Council for Strategic Research.

The authors would like to acknowledge the Danish Meteorological Institute (DMI) for free usage of radar data and DHI group for free use of the MOUSE software for research purposes.

REFERENCES

Achleitner, S., Fach, S., Einfalt, T. & Rauch, W. 2009 Nowcasting of rainfall and of combined sewage flow in urban drainage systems. *Water Sci. Technol.* **59** (6), 1145–1151.

Ahm, M., Thorndahl, S., Rasmussen, M. R. & Bassø, L. 2012 Estimating runoff coefficients using weather radars. *Water Sci. Technol.* **December**, 1–9.

Beven, K. 2006 A manifesto for the equifinality thesis. *J. Hydrol.* **320** (1–2), 18–36.

Beven, K. J. & Freer, J. 2001 Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **249** (1–4), 11–29.

Bøvith, T. 2008 Detection of Weather Radar Clutter. PhD Thesis, Technical University of Denmark, IMM-PHD-2008-201.

Bøvith, T., Nielsen, A. A., Hansen, L. K., Overgaard, S. & Gill, R. S. 2006 Detecting weather radar clutter by information fusion with satellite images and numerical weather prediction model output. International Geoscience and Remote Sensing Symposium (IGARSS), Denver, Colorado, USA, 31 July–4 August 2006, art. no. 4241283, pp. 511–514.

Bowler, N. E., Pierce, C. E. & Seed, A. W. 2006 STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Q. J. R. Meteorolog. Soc.* **132** (620), 2127–2155.

Broyden, C. G. 1970 The convergence of a class of double-rank minimization algorithms. *J. Inst. Math. Appl.* **6**, 76–90.

Collier, C. G. 2007 Flash flood forecasting: what are the limits of predictability? *Q. J. R. Meteorolog. Soc.* **133** (622), 3–23.

DHI 2009 *Mouse Reference Manual*. DHI. Available from www.dhigroup.com.

Fletcher, R. 1970 A new approach to variable metric algorithms. *Comput. J.* **13**, 317–322.

Germann, U. & Zawadzki, I. 2002 Scale-dependence of the predictability of precipitation from continental radar images. Part I: Methodology. *Mon. Weather Rev.* **130**, 2859–2873.

Gill, R. S. 2007 Sea clutter removal using radar elevation dependent second order texture parameters. Danish Meteorological Institute, Scientific Report 07-01, available at www.dmi.dk/dmi/sr07-01, ISBN: 978-87-7478-544-6.

Gill, R. S., Overgaard, S. & Bøvith, T. 2006 The Danish weather radar network. *Proceedings of Fourth European Conference on Radar in Meteorology and Hydrology*, Barcelona, Spain, 18–22 September 2006, Available from www.crahi.upc.edu/ERAD2006/proceedingsMask/00107.pdf.

Goldfarb, D. 1970 A family of variable metric updates derived by variational means. *Math. Comput.* **24**, 23–26.

Golding, B. W. 1998 Nimrod: a system for generating automated very short range forecasts. *Meteorol. Appl.* **5**, 1–16.

Grum, M., Longin, E. & Linde, J. J. 2004 A flexible and extensible open source tool for urban drainage modelling. *Proceedings of the 6th International Conference on Urban Drainage Modelling*, Dresden, Germany, 15–17 September 2004. www.WaterAspects.org.

Krajewski, W. F. & Smith, J. A. 2002 Radar hydrology–rainfall estimation. *Adv. Water Resour.* **25**, 1387–1394.

Krajewski, W. F., Kruger, A., Smith, J. A., Lawrence, R., Gunyon, C., Goska, R., Seo, B.-C., Domaszczynski, P., Baeck, M. L., Ramamurthy, M. K., Weber, J., Bradley, A. A., DelGreco, S. A. & Steiner, M. 2011 Towards better utilization of NEXRAD data in hydrology: an overview of Hydro-NEXRAD. *J. Hydroinf.* **13** (2), 255–266.

Leitão, J. P., Simões, N. E., Maksimović, Č., Ferreira, F., Prodanović, D., Matos, J. S. & Sá Marques, A. 2010 Real-time forecasting urban drainage models: full or simplified networks? *Water Sci. Technol.* **62** (9), 2106–2114.

Li, L., Schmid, W. & Joss, J. 1995 Nowcasting of motion and growth of precipitation with radar over a complex orography. *J. Appl. Meteorol.* **34** (6), 1286–1300.

Liguori, S. & Rico-Ramirez, M. A. 2012 Quantitative assessment of short-term rainfall forecasts from radar nowcasts and MM5 forecasts. *Hydrol. Process.* **26** (25), 3842–3857.

Liguori, S., Rico-Ramirez, M. A., Schellart, A. N. A. & Saul, A. J. 2012 Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. *Atmos. Res.* **103**, 80–95.

Marshall, J. S. & Palmer, W. M. 1948 The distribution of raindrops with size. *J. Atmos. Sci.* **5**, 165–166.

Mecklenburg, S., Joss, J. & Schmid, W. 2000 Improving the nowcasting of precipitation in an Alpine region with an enhanced radar echo tracking algorithm. *J. Hydrol.* **239** (1–4), 48–68.

Mueller, C. K. & Megenhardt, D. 2001 Short-term (0–2 hr) automated growth forecasts of multi-scale convective systems associated with large scale, daytime forcing. In: *Preprints, 30th International Conf. On Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc., pp. 231–233.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *J. Hydrol.* **10** (3), 282–290.

Rinehart, R. E. & Garvey, E. T. 1978 Three-dimensional storm motion detection by conventional weather radar. *Nature* **273**, 287–289.

Roberts, R. & Rutledge, S. 2003 Nowcasting storm initiation and growth using GOES-8 and WSR-88D data. *Weather Forecast.* **18** (4), 562–584.

Schaarup-Jensen, K. & Rasmussen, M. R. 2004 The characteristics of waste water flow in a Danish combined sewer. *Proceedings of the 4th International Conference on Sewer Processes and Networks*, Funchal, Madeira, 22–24 November 2004.

Schaarup-Jensen, K., Hvitved-Jacobsen, T., Jütte, B., Jensen, B. & Pedersen, T. 1998 A Danish sewer research and monitoring station. *Water Sci. Technol.* **37** (1), 197–204.

Schaarup-Jensen, K., Rasmussen, M. R. & Thorndahl, S. 2009 To what extent does variability of historical rainfall series influence extreme event statistics of sewer system surcharge and overflows? *Water Sci. Technol.* **60** (1), 87–95.

Schellart, A. N. A., Shepherd, W. J. & Saul, A. J. 2012 Influence of rainfall estimation error and spatial variability on sewer flow prediction at a small urban scale. *Adv. Water Resour.* **45**, 65–75.

Seo, D.-J., Breidenbach, J. P. & Johnson, E. R. 1999 Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol.* **223**, 131–147.

Seo, B.-C., Krajewski, W. F., Kruger, A., Domaszczynski, P., Smith, J. A. & Steiner, M. 2011 Radar-rainfall estimation algorithms of Hydro-NEXRAD. *J. Hydroinf.* **13** (2), 277–291.

Shanno, D. F. 1970 Conditioning of quasi-Newton methods for function minimization. *Math. Comput.* **24**, 647–656.

Thorndahl, S. & Willems, P. 2008 Probabilistic modelling of overflow, surcharge and flooding in urban drainage using the first-order reliability method and parameterization of local rain series. *Water Res.* **42** (1–2), 455–466.

Thorndahl, S. & Rasmussen, M. R. 2012 Marine X-band weather radar data calibration. *Atmos. Res.* **103**, 33–44.

Thorndahl, S., Beven, K. J., Jensen, J. B. & Schaarup-Jensen, K. 2008 Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. *J. Hydrol.* **357** (3–4), 421–437.

Thorndahl, S., Bøvith, T., Rasmussen, M. R. & Gill, R. S. 2012b On comparing NWP and radar nowcast models for forecasting of urban runoff. *IAHS Proc. Rep.* **351**, 620–625.

Thorndahl, S., Grum, M., Rasmussen, M. R. & Schaarup-Jensen, K. 2011 Flow forecasting in drainage systems with extrapolated radar rainfall data and auto calibration on flow observations. *Proceedings of the 12th International Conference on Urban Drainage*, Porto Alegre, Brazil, September 2012.

Thorndahl, S., Johansen, C. & Schaarup-Jensen, K. 2006 Assessment of runoff contributing catchment areas in rainfall runoff modelling. *Water Sci. Technol.* **54** (6–7), 49–56.

Thorndahl, S., Poulsen, T. S., Bøvith, T., Borup, M., Ahm, M., Nielsen, J. E., Grum, M., Rasmussen, M. R., Gill, R. & Mikkelsen, P. S. 2012a Comparison of short term rainfall forecasts for model based flow prediction in urban drainage systems. *Water Sci. Technol.* **December**, 1–9.

Thorndahl, S., Rasmussen, M. R., Grum, M. & Neve, S. L. 2009 Radar based flow and water level forecasting in sewer systems – a Danish case study. *Proceedings of the 8th International workshop on precipitation in urban areas*, 10–13 December, 2009, St. Moritz, Switzerland.

Thorndahl, S., Rasmussen, M. R., Nielsen, J. E. & Larsen, J. B. 2010 Uncertainty in Nowcasting of radar rainfall: a case study of the GLUE methodology. *Proceedings: Advances in Radar Technology. Sixth European Conference on Radar in Meteorology and Hydrology*, Sibiu, Romania.

Wang, L. P., Ochoa, S., Simões, N., Onof, C. & Maksimović, C. 2012 Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. *Proceedings of the 9th International Conference on Urban Drainage Modelling*, Belgrade, 2012.

First received 12 October 2012; accepted in revised form 14 January 2013. Available online 11 February 2013