Dynamic gauge adjustment of high-resolution X-band radar data for convective rain storms: Model-based evaluation against measured combined sewer overflow

Borup, Morten; Grum, Morten; Linde, Jens Jørgen; Mikkelsen, Peter Steen

Published in:
Journal of Hydrology

Link to article, DOI:
10.1016/j.jhydrol.2016.05.002

Publication date:
2016

Document Version
Peer reviewed version

Dynamic Gauge Adjustment of High-Resolution X-Band Radar Data for Convective Rain Storms: Model-Based Evaluation against Measured Combined Sewer Overflow

Morten Borup[a], *Morten Grum*[b], *Jens Jørgen Linde*[b] *and Peter Steen Mikkelsen*[a]
[a)] Technical University of Denmark, Dept. of Environmental Engineering (DTU Environment), Miljoevej, Building 113, 2800 Kgs Lyngby, Denmark
[b)] Krüger A/S, Veolia Water Solutions and Technologies, Denmark

Corresponding author: Morten Borup
E-mail: morb@env.dtu.dk
Tel: +45 4525 2182

ABSTRACT
Numerous studies have shown that radar rainfall estimates need to be adjusted against rain gauge measurements in order to be useful for hydrological modelling. In the current study we investigate if adjustment can improve radar rainfall estimates to the point where they can be used for modelling overflows from urban drainage systems, and we furthermore investigate the importance of the aggregation period of the adjustment scheme. This is done by continuously adjusting X-band radar data based on the previous 5-30 minutes of rain data recorded by multiple rain gauges and propagating the rainfall estimates through a hydraulic urban drainage model. The model is built entirely from physical data, without any calibration, to avoid bias towards any specific type of rainfall estimate. The performance is assessed by comparing measured and modelled water levels at a weir downstream of a highly impermeable, well defined, 64 ha urban catchment, for nine overflow generating rain events. The dynamically adjusted radar data perform best when the aggregation period is as small as 10-20 minutes, in which case it performs much better than static adjusted radar data and data from rain gauges situated 2 – 3 km away.

Keywords: Radar; rainfall; stormwater; dynamic adjustment; distributed hydraulic model; combined sewer overflow.

1 INTRODUCTION

Accurate rainfall estimates are required in much higher temporal and spatial resolution to perform successful hydrological modelling of urban stormwater runoff than for most other uses, due to the fast hydraulic response of the urban stormwater systems. Today's detailed distributed urban drainage models can operate with thousands of sub catchments – often of sizes less than one hectare (1 ha=10000 m^2). Nonetheless the typical rainfall inputs to these models are still produced by just a few rain gauges. There have long been big expectations to the use of weather radar data for urban drainage (Einfalt et al., 2004) and it is apparent that the spatially distributed nature of weather radar data fits well with distributed runoff models, but despite the technological development within the last decade most radar rainfall estimates are still affected by significant errors that are difficult to quantify (Berne and Krajewski, 2013). A recent literature study found that the standard deviation of the error of radar rain estimates as a proportion of the rain rate typically lies in the range 0.3-0.5 for hourly data (McMillan

et al., 2012). In a thorough study of the uncertainties of radar rainfall estimates produced using the Hydro-NEXRAD algorithm it was found that "radar-rainfall uncertainty is characterized by an almost three times greater standard error at higher resolutions (15-minute and 0.5 km scale) than at lower resolutions (1-hour and 8 km)"(Seo and Krajewski, 2010). This means that it can be challenging to use radar data for urban catchments where the spatial and temporal scales are relatively small. Fortunately, the quality of weather radar rainfall estimates is continuously improved (e.g. Krämer and Verworn, 2009; Nielsen et al., 2013) and in the recent years several research projects have shown some success in using radar data for urban runoff models. Estimation of the peak rain intensities however continues to be a problem (Thorndahl and Rasmussen, 2012). A recent Belgian case study using a modern X-band weather radar and testing various calibration methods found that rain gauge data generally outperform radar data as input to the distributed urban runoff model used in the study (Goormans and Willems, 2012), despite the radar's advantage of being able to estimate the spatial distribution of rain. This shows that even the newest weather radars have difficulties in producing rainfall estimates that are suitable for urban runoff modelling. The reason for this lies in the way radars detect the rain. Radars do not measure rainfall directly but sends out a pulse of microwave radiation and measures the fraction of backscattered energy from whatever obstacles the radar pulse may hit. By using a relationship between the reflected energy and the rain rate (the Z-R relationship) it is possible to estimate the rain rate, a subject already extensively covered in literature (e.g. Atlas, 1990; Rinehart, 1991; Sauvageot, 1992; van de Beek et al., 2010). The Z-R relationship is not constant, however, and has even been observed to change dramatically several times during a rainfall event (Clemens et al., 2006). Since the Z-R relationship is heavily dependent on the drop size distribution (DSD) of the rain these changes can be explained by changes in the DSD, since this has been shown to vary drastically during events in multiple studies (Chapon et al., 2008; Cifelli et al., 2000; Smith and Krajewski, 1993; Smith et al., 2009). These changes pose a limit to how accurate the quantitative precipitation estimates from radars can be when a constant relationship between radar reflectivity and rainfall intensity is used throughout an entire event (Lee and Zawadzki, 2006). Several studies have shown that quantitative precipitation estimates from weather radars are improved by dynamically adjusting the radar data by rain gauge measurements (Cole and Moore, 2008; Creutin et al., 1997; Goudenhoofdt and Delobbe, 2009; Shrestha et al., 2013; Thorndahl et al., 2014; Wood et al., 2000). None of these, however, have focused on producing rainfall estimates suitable for online modelling of urban runoff for high intensity events.

It is rarely the absolute depth of a rain event that induces problems in urban areas such as local flooding, water in the basements and combined sewer overflow (CSO). Problems occur as soon as the mean areal rain intensity exceeds the bottle neck capacity of the sewer system for a period of time that is comparable to the response time of the system. For this reason the highest intensities are of the highest interest for the urban runoff modeller and therefore the dynamic adjustment scheme proposed in this study aims at improving especially the radar's ability to estimate the highest intensities.

In the current study radar data are provided by a DHI LAWR (Local Area Weather Radar) (Jensen, 2000). This is a small X-band weather radar of growing popularity among municipalities due to its low cost and ability to provide rainfall estimates with a

pixel size of just 100 m and a 1 minute temporal resolution. This means that the resolution of the radar data is more detailed than necessary in order to describe the temporal and spatial variability of the precipitation of importance for urban runoff modelling (Berne et al., 2004; Ochoa-Rodriguez et al., 2015). The most important disadvantages of this radar type, compared with C- and S-band radars, is its limited range (maximum range is 60 km and quantitative estimation is possible up to 20 km) and larger attenuation due to the X-band frequency. Besides that the LAWR has the same problems with quantitative precipitation estimates as other radar types and therefore needs to be adjusted using rain gauge data in order to be useful in urban hydrology (Willems et al., 2012).

The dynamics and depths of the stratiform rain events of the winter season are in general well described by gauges (Shrestha et al., 2013), while these are less good at describing the convective events of the summer season. The latter events happen to be those that most often cause problems in the urban environment due to high local intensities, but to catch the spatial variability of such a convective storm over an entire city an unrealistically high number of gauges would be required. Therefore, these are the kind of events where high resolution radar data could be useful and thus the radar rainfall estimates should be validated against this kind of events. The rain events used in the current study have been selected based on the criterion that they should have resulted in at least 100 m^3 of CSO from a specific structure. This led to nine events which are mainly of a convective nature.

Validation of radar rainfall estimates is not a simple task and using rain gauges as the "ground truth" is problematic. Rain gauges at best represent the rainfall at one specific point covering only a few hundred square centimetres, while the validation ideally should be done on areal rainfall, which is the quantity of interest for runoff modelling. Since runoff is a direct response to the areal rainfall on the catchment, the quality of different rainfall estimates are in this article assessed by comparing modelled and measured runoff from a highly impervious, well defined urban catchment. This validation method is not affected by the various kind of catching errors associated with rain gauge measurements (McMillan et al., 2012), and it focuses on a property of direct interest in urban drainage modelling, here the volume of CSO events. Instead uncertainties regarding the model setup are introduced. These are attempted minimized by building a highly detailed distributed model for an area where the system is very well known and well defined. Furthermore, a very impervious area is chosen as base for the model to minimize the big uncertainties regarding the fast runoff from permeable surfaces. To avoid bias towards any specific type of rainfall estimates, the runoff model is built purely from physical data without any calibration.

The paper has the following main aims:

- Test whether rainfall estimates from an X-band weather radar can be improved by adjusting the rainfall estimates from nearby gauge measurements, to the point where the radar data can be used for modelling urban sewer overflows.
- Explore the impact of the length of the time horizon used for the dynamic adjustment of the radar data.
- Use a well determined runoff model to assess the quality of the rainfall estimates. This is done by using the rainfall estimates as input to a detailed

distributed hydrodynamic urban runoff model and comparing the modelled and measured water levels at a downstream overflow structure. In this way the point-area sampling error and other error sources connected with rain gauge observations that often distort studies regarding radar quantitative precipitation estimation (QPE) are minimised.

The paper is structured in the following manner: In Section 2 the data is presented followed by Section 3 in which we present the various ways the final rainfall data products are produced. In Section 4 the model based validation methods are described while the results are presented and discussed in Section 5. Finally the conclusions are presented in Section 6.

2 DATA BASIS

2.1 Instruments

The radar used is a DHI LAWR X-band radar located 7 km from the centre of the Danish city Odense. The processed radar data used as basis for the investigation have a spatial resolution of 100 x100 m and a temporal resolution of 1 minute. The opening angle of the radar is ±10 degrees, which implies that the radar detects rain up to an elevation of 700 m above the centre of Odense and up to 1200 m in the northern most outskirts of the city. Four RIMCO tipping bucket rain gauges with 0.2 mm resolution are used for the adjustment of the radar data. The raw gauge data are transformed into time series as described in (Jørgensen et al., 1998). The location of the rain gauges and radar can be seen in Figure 1. Note that gauge A is situated centrally in the small validation catchment and therefore this gauge is expected to represent the rainfall over the catchment much better than the other gauges.

The runoff from the validation catchment is only measured indirectly by a water level gauge situated at the downstream weir. This is used instead of discharge data simply because only water level data are available. A sketch of the overflow structure can be seen in Figure 2. The gauge is a pressure gauge situated in level 6.66 m above datum, 0.42 m above the bottom of the structure, which means that the gauge has a lower detection limit at this level. The crest of the weir is situated 6.95 m above datum and is a 2 m wide horizontal crest situated orthogonally to the flow direction and with free overflow.

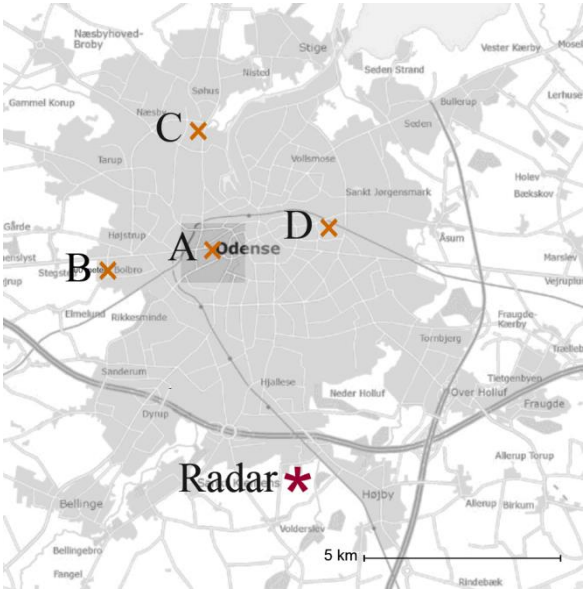


Figure 1: *Location of the four rain gauges and the radar. The validation catchment is located within the gray shaded square covering gauge A, see Figure 4.*

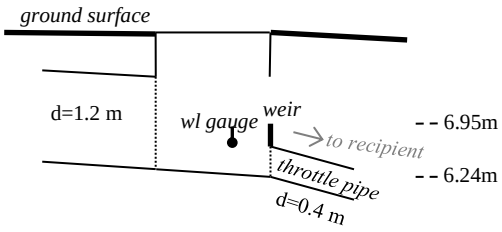


Figure 2: *Sketch of the overflow structure seen from the side.*

2.2 Rain events

The rain events used in this study are the nine events during the summer of 2008 that generated more than 100 m^3 of overflow at the weir used for the validation and for which radar data, rain gauge data and weir water level data are available. The event definition used is the time span from either of the rain gauges record the first rain until none of the gauges record rain. The individual events are separated by at least four hours without any recorded rain, which is much more than it takes for the stormwater from a rain event to drain from the catchment (the catchment properties are described in section 4.1).

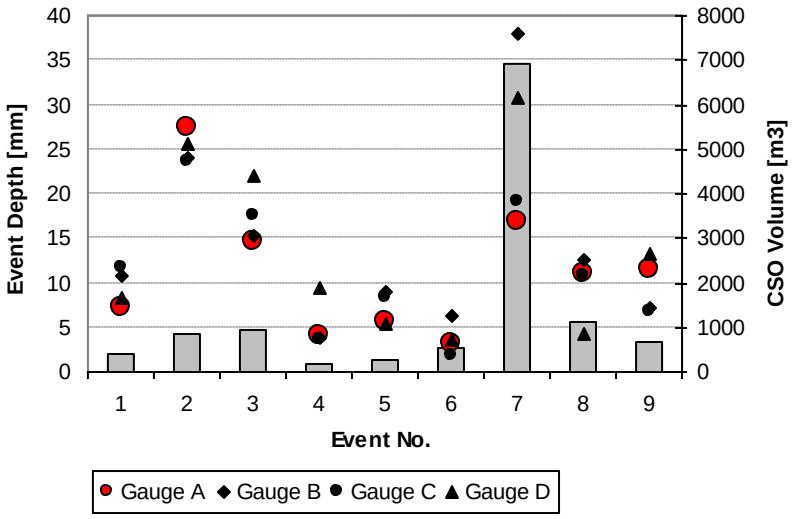


Figure 3: *Event depths measured by gauge plotted on top of the calculated overflow volumes (grey bars).*

The depths of the nine events measured by the rain gauge within the catchment range from 3 to 27 mm. Descriptive data of the events can be seen in Table 1 and Figure 3, and the rainfall hyetographs are shown in the Appendix. Note that overflow sometimes occurs for surprisingly small rainfall events, which however have short intense periods causing this CSO structure in the old city centre of Odense to overflow. The rain cell speed in Table 1 is calculated by cross-correlation analysis of the radar data (Austin and Bellon, 1974). By looking at Figure 3 it is easily seen that the overflow volume is not strongly correlated to the event depth, since event 7, which results in the largest overflow with a factor of six down to the second largest, has an event depth comparable to that of event 2 and 3. Furthermore it is noticeable that event 6, for which only 2 – 7 mm rainfall is recorded (3.2 mm at gauge A), still manages to generate a substantial overflow. This shows that the dynamics of the rainfall is just as important as the event depth for generating overflow. The reason for this is that overflow only happens when the discharge from the catchment exceeds the capacity of the throttle pipe, and this is a consequence of a period with intense rainfall which is only vaguely related to the event depth for most events.

Table 1: Date and time for the nine events, the horizontal rain cell speed acquired by correlation analysis of the radar images and the event rain depth recorded by gauge A.

Event no.	Date	Time	Rain Cell Speed [m/s]	Rain depth gauge A [mm]
1	19. maj	04:10 - 07:30	9	7.2
2	26. maj	03:50 - 16:00	5	27.4
3	4. aug.	01:30 - 09:00	10	14.6
4	11. juli	00:15 - 02:10	13	4.0
5	21. juli	04:20 - 06:50	10	5.7
6	21. aug.	15:50 - 16:50	13	3.2
7	21. aug.	21:00 - 23:10	11	16.8
8	23. aug.	06:00 - 09:10	7	11.1
9	23. aug.	15:10 - 22:00	9	11.5

The observed water level at the weir and the rainfall hyetographs for all the gauges and for all events can be seen in the appendix.

3 PROCESSING OF RADAR DATA

3.1 Preliminary processing

The internal software of the LAWR system converts the raw polar radar images into cartesian precipitation maps. This process includes a distance correction algorithm which ensures that the radar signal is corrected for the decline in returned signal as the distance to the radar increases, and an en-route correction procedure that adjust the signal for beam blockage effects. Clutter is removed before and after the final conversion into cartesian coordinates by subtracting the values from a clutter map generated on a rain-free day (Pedersen et al., 2010).

The LAWR converts the radar data into precipitation estimates by multiplying the value at each radar pixel with the ratio between accumulated rain gauge data and accumulated radar data for a period with spatially uniform rainfall. The use of this kind of empirical

relationship for conversion into precipitation estimates is possible because the LAWR has a logarithmic receiver, which gives a linear relationship between rainfall and reflectivity (Pedersen et al., 2010). The accumulated rain depth recorded by the gauge(s) is assumed to be representative for the entire area covered by the radar. If the rainfall is in fact not spatially uniform during the calibration period, the consequence will be that the LAWR systematically over- or underestimates the rainfall in areas where the actual rainfall depth deviated from the rainfall depth recorded by the rain gauges during the calibration. The calibration period used for the preliminary calibration of all the radar data in the current study lasted from early April to mid-May 2008, in which period approximately 40 mm of rainfall was recorded by the rain gauges. The maximum recorded 5 minute rainfall intensity in this period was just 7.2 mm/h.

3.2 Definition of adjustment factors

The gauge adjustment of the radar data is a final processing step that correct for errors in the already calibrated radar rain estimates. The adjustment factor α relates the rain rate estimated by the radar Rr with the actual mean areal rain rate R on an area covered by a radar pixel by assuming that:

$$R = \alpha \cdot Rr + \epsilon_r \qquad (1)$$

where ϵ_r is the radar rainfall estimation error. Likewise it is assumed that if a rain gauge is situated within a radar pixel then R can be related to the rain gauge measurements Rg by:

$$R = Rg + \epsilon_g \qquad (2)$$

where ϵ_g is the rain gauge rain estimation error including catching errors and area to point estimation errors. Combining these to equations leads to:

$$Rg = \alpha \cdot Rr + \epsilon_{total} \qquad (3)$$

where ϵ_{total} is an error term that represents mean areal rain estimation errors from both radar and rain gauge. By ignoring that the variance of ϵ_{total} most likely is bigger for larger values of Rr and therefore does not fulfil the requirement of homoscedasticity, it is possible to estimate α using the Least Squares Method for a linear regression through the origin (Gordon, 1981). By furthermore assuming that ϵ_{total} is of much larger magnitude than the spatial variations in the actual rainfall-to-radar relationship within the area covered by the gauges in the setup, the linear regression used to estimate α can be written as:

$$\hat{\alpha}(t) = \frac{\sum_{\tau=t-T_{ad}+1}^{t} \sum_{n=1}^{N} Rr(n,\tau) Rg(n,\tau)}{\sum_{\tau=t-T_{ad}+1}^{t} \sum_{n=1}^{N} Rr(n,\tau)^2} \qquad (4)$$

where N is the number of rain gauges used for the adjustment and T_{ad} is the number of one minute time steps that are used.

This method possesses the desired property that data pairs where *Rr* is relatively large have a relative high impact on the result, while data pairs where *Rr* is small or zero have little or no impact at all. This is an advantage for two reasons. When *Rr* is large, *Rg* will tend to be large as well, which decreases the uncertainty induced by the coarse 0.2 mm discretisation of the tipping bucket rain gauge, and secondly it is more important to estimate the adjustment factor correctly for large *Rr* due to the potential larger impact of high radar rainfall estimates on the runoff model.

The *Rg* and *Rr* values used to estimate α are calculated for each minute as an average of the last five minutes of data. This is done to ensure some sort of physical connection between gauge data and radar data, since it takes time for drops reflecting a radar signal to reach the ground and be detected by a point gauge. 5 minutes was chosen since this is sufficient time for raindrops larger than 1 mm, falling with a velocity of minimum 4 m/s (Beard, 1976), to travel the distance between the maximum elevation of the radar beam anywhere over the city to the ground. An optimal averaging period cannot, however, be computed simply by looking at the drops vertical fall velocity, since rain drops fall with different velocities depending on drop sizes and metrological conditions (Beard, 1976) and seldom fall vertically as pointed out by Dai and Han (2014). Note that due to the 5 minute averaging period α is actually calculated based on the previous T_{ad} + 4 minutes of data.

In the following equation 4 is used for calculating the adjustment factor regardless of the size of the time period and number of gauges. That is, the adjustment factor is calculated the same way whether it is based on 15 minutes of data from a single gauge or data from all nine events from all four gauges. The same α is used for the entire radar image.

It was also tested to calculate the adjustment factors as in mean field bias correction of Smith and Krajewski (1991) as the ratio between the sum of rain gauge data and the sum of radar data, but this method led to worse results when based on small periods of time (single events or less) and no improvement for long periods of time. For the readability of the paper these results have been omitted in the results section.

3.3 Static Adjustment

Static adjustment aims to improve the quality of radar rainfall estimates by correcting for the long term bias between gauges and radar. This is done using a single constant-in-time adjustment factor, which is calculated using the linear regression from equation 4 on all data for the period for which the factor is used. In this way the static adjustment is assumed to perform better than realistically achievable by static adjustment in a real time application where only prior data is known.

Three different Event based Static Adjustments (ESA) are tested: ESA_A, ESA_4 and ESA_3. These are either based on rain gauge A, on all four rain gauges or on the three gauges left when omitting gauge A, respectively. ESA_A is used to test the effect of adjusting the radar data against a gauge that represents the local rainfall well (ESA_A). ESA_3 is used to test how well adjusted radar data works for an area without a rain gauge, which is the typical situation for most small urban catchments, while ESA_4 is used to see how good a rainfall estimate can be obtained when combining all available

data. A static adjustment that is constant for all nine events and based on the data from all four gauges is also tested. This is referred to as CSA_4 (Constant Static Adjustment).

Note that the static adjustments differ conceptually from the preliminary conversion into precipitation estimates described in section 3.1. The preliminary conversion assumes that rain gauge data is representative for the rainfall for the entire area covered by the radar, while the static adjustments assume that rain gauge data is representative for the specific radar pixel covering the gauge only and let the spatial distribution of the rainfall be determined by the radar data.

3.4 Dynamic Adjustment

The purpose of using a dynamic adjustment scheme for radar data of such a high spatial and temporal resolution as used in this paper and for such a small area is to be able to adjust for the changes in the Z-R relationship, that happen during an event. To be able to detect and react upon these changes the period used to calculate the adjustment factor should be as small as possible. On the other hand, the uncertainty of the calculated adjustment factor will increase with the decreasing data basis for the calculation. This can be counteracted by using more gauges for the calculation. If these gauges are located far apart, it will result in some smoothing of the calculated adjustment factor due to the spatial and temporal correlation of the structure of the rain. The typical horizontal speed of the rain cells was 10 m/s, see Table 1, which means that it will take a rain cell 10 minutes to travel the 6 km between the two gauges that are the furthest apart (gauge B and D). This implies that changes in Z-R relations that are faster than 10 minutes will be smoothened by the adjustment scheme when multiple gauges are used.

To allow changes to take effect quickly α is calculated for each one minute time step, based only on the data available up to, and including, the current time step. This makes the method suitable for online situations in which rapid response is required, as opposed to the mean field bias adjustment of Smith and Krajewski (1991) that adjusts radar estimates in time intervals based on prior knowledge of the accumulated radar and rain gauge volumes for the entire interval.

A conservative measure has been added to the method to avoid excessive variations in the alpha value when there are only a few valid data pairs (both radar and rain gauge differs from zero) available, which is always the case at the beginning of an event. When actual valid data pairs account for less than half of the total number of data pairs, α is calculated as a weighted mean between the actual estimated α calculated from (4) and a default alpha value, as:

$$\alpha = \alpha_{actual} \cdot \frac{N_{actual}}{N_{total}} + \alpha_{default} \cdot \left(1 - \frac{N_{actual}}{N_{total}}\right) \qquad (5)$$

where N_{actual} is the number of valid data pairs within the interval, α_{actual} is the result of using (4) on the valid data pairs, $\alpha_{default}$ is a default alpha value, and N_{total} is the total number of potentially valid data pairs, i.e. the number of gauges multiplied with T_{ad}. $\alpha_{default}$ could in principle be changed from event to event based on weather type but when using multiple gauges the parameter proved to be of minor importance and thus a

value of one was chosen, which corresponds to using the pre-processed radar data unadjusted.

The following dynamic adjustment schemes are tested:

$DA_A(T_{ad})$: Dynamic Adjustment based solely on gauge A
$DA_4(T_{ad})$: Dynamic Adjustment based on all four gauges
$DA_3(T_{ad})$: Dynamic Adjustment based on all gauges but gauge A

, where T_{ad} is varied between 5 and 30 minutes.

4 VALIDATION USING MODELLED RUNOFF

The runoff from an impermeable surface is the closest one can get to estimate how much rain is actually hitting the ground for areas of a size of hydrological interest. In the following modelled runoff from an urban catchment is compared with measured runoff by comparing water levels at a weir downstream the validation catchment, see Figure 4. By using a distributed model the radar's ability to describe the spatial variability of the rainfall is included in the evaluation.

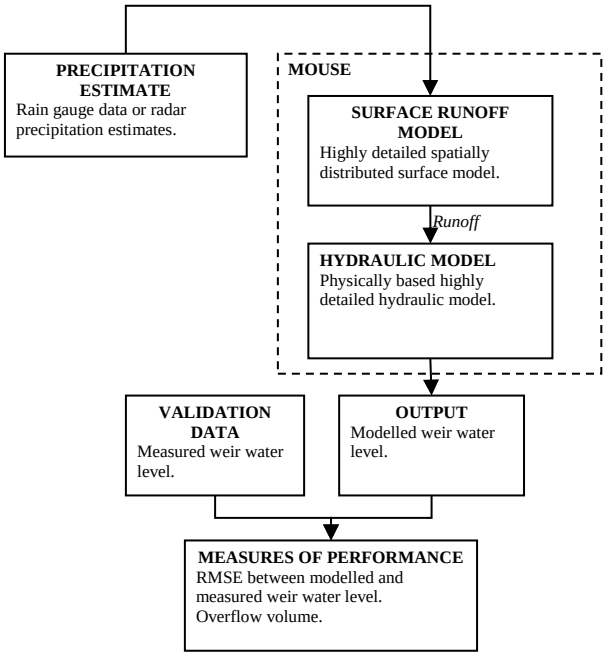


***Figure 4**: Flowchart of the method for validating by runoff.*

4.1 The runoff model

The runoff is simulated using the MOUSE (Model of Urban SEwers) software (DHI, 2007). The hydrodynamic module of MOUSE uses a finite difference scheme to solve the full one-dimensional St.-Venant equations. The parameters of a MOUSE hydrodynamic model are physical attributes such as dimensions of pipes, basins and weirs, etc. This makes it possible to create a decent model without any calibration, if data of high quality describing the sewer system is available. As input to the hydrodynamic model a simple distributed time-area surface runoff model is used, which determines the runoff from each individual sub catchment from the concentration time, a time-area curve and the impervious area. For a rectangular catchment the time-area curve will be a straight line, which implies that the runoff at any point in time is calculated as the average rainfall for the time of concentration multiplied with the impervious area.

Figure 5*: Map showing the detailed grid of sub catchments in the validation catchment and the location of rain gauge A. The orange part drains directly to the validation weir (red triangle) and the rest can interact with the orange part in case of heavy backwater effects. The small round dots are manholes and the lines connecting manholes are pipes.*

A map of the validation catchment is shown in Figure 5. The catchment is highly impermeable (81%) which reduces the model uncertainties since the fast runoff from permeable surfaces is very difficult to model. It is assumed that there is no runoff from the permeable surfaces. This is a fair assumption for small events but it might lead to a slight underestimation of the runoff for large rain events since the permeable areas in

these cases might start to contribute with runoff due to saturation of the top soil. The impervious areas are mostly asphalt roads, roofs and parking lots. The boundary of the catchment is well defined by a railroad, a river and a cemetery working as efficient water divides. The total catchment size is 64 ha and the model's mean sub-catchment size is less than 0.5 ha. The shapes and impermeable areas of the sub-catchments come from processed aerial photos and the description of the runoff system is based entirely on physical data from the municipality – no calibration has been done. The concentration time for each individual sub-catchment is set to 7 minutes. That this is not a very sensitive parameter can be seen in section 5.1. Included in the model are 266 sub-catchments defined so that the surface subdivision matches the level of detail in the pipe system model (302 manholes and 321 pipes). The exact same model is used for routing of both radar data and rain gauge data through the catchment model. The response time scale of the flow to the weir, i.e. the time scale at which the pattern of the time-averaged radar rainfall hyetograph is the most similar to the modelled runoff hydrograph (Morin et al., 2001), is approximately 25 minutes. That is, the time averaged rainfall is calculated for time *t* as the average of the previous *Tc* minutes of rainfall. The estimated response time is the *Tc* that makes the time averaged rainfall look the most like the modelled runoff in terms of shape and location of the peaks. For a rectangular catchment *Tc* correspond to the time of concentration.

4.2 Applying radar data to the model

The radar rainfall estimates are applied to the model by letting the values from each radar pixel be represented as a separate virtual rain gauge located in the centre of the pixel, so that the model is covered with a virtual grid of 294 virtual rain gauges with 100 m in-between. By using a standard feature in MOUSE, each catchment is assigned the rainfall input from the gauge that is nearest to the centre of the catchment. This rather simple method of distributing the radar data is feasible due to the small size of the sub-catchments. If the sub-catchments had been much bigger than the radar pixels, it would have been necessary, in order to fully utilize the information in the radar data, to calculate the radar rainfall on each sub-catchment by averaging the data from several pixels depending on how big a percentage of the catchment area each pixel was covering.

4.3 Measures of performance

Root Mean Squared Error

To assess the quality of a given rainfall input to the model, a measure of performance is defined as the Root Mean Squared Error (RMSE) between observed (O) and modelled (M) water levels at the weir downstream of the validation catchment:

$$RMSE = \sqrt{\frac{\sum_{n=1}^{Nobs}(O_n - M_n)^2}{Nobs}} \qquad (6)$$

Nobs is the total number of data points used for the calculation. Only data points where the measured water level is above the lower detection limit of the level meter are included. This measure of performance is very sensitive to temporal differences which makes it suitable to evaluate the dynamics of the runoff.

Overflow volume

The total overflow volume for each event is furthermore used as an indicator of the quality of the rainfall input. The total overflow volume does not take temporal differences into account but it does say something about the rainfall estimates´ ability to describe the periods with the highest intensities correctly, since these are the ones that cause overflows. Furthermore, overflow is of paramount importance when managing urban drainage systems and therefore this measure is interesting by itself.

A calibrated weir discharge coefficient was not available for the weir. Instead the standard overflow formulae for an orthogonal weir implemented in the hydrodynamic computations of MOUSE is used (DHI, 2007):

if $h > h_{crest}$

$$Q_w(h) = B \cdot \sqrt{g} \cdot \left(\frac{2}{2 + K_c} \cdot (h - h_{crest}) \right)^{3/2} \quad (7)$$

else

$$Q_w(h) = 0 \quad (8)$$

Qw is the discharge over the weir, B is the crest width, g is the gravitational constant, K_c is the energy loss coefficient (0.5 for a sharp-edged outlet), h is the water level and h_{crest} is the level of the crest of the weir.

In the following the exact same formula (equations 7+8) is used to compute the observed and the modelled overflow.

There is a big difference in the amount of overflow from the events. To be able to compare model performance for different events, the overflow based on the modelled water level is divided with the overflow based on the measured water level to create the Relative Overflow Volume (ROV). That is, ROV equal to one is the best score. ROV is calculated as:

$$ROV = \frac{\sum Q_w(\text{modelled h})}{\sum Q_w(\text{observed h})} \quad (9)$$

Notice that the parameters from equation (7) will cancel each other out in the calculation of ROV (9). The parameters do, however, play an important role when the model calculates the water level at the weir, and the parameters are therefore still important. Overflow volume computations are rather uncertain when not based on a calibrated weir discharge coefficient and therefore too much emphasis should not be put on the ROV from individual events but rather on the distribution of the ROV from all events, which preferably should have a median close to one and a small spread.

5 RESULTS AND DISCUSSION

5.1 Uncertainties

Model errors and errors in the water level observations will affect both measures of performance defined above. Model and observation errors can affect both modelled and observed overflow in a systematic way that could change the conclusions based on ROV. If e.g. the observed overflow is positively biased then the optimal ROV is a value smaller than one.

Table 2 shows how sensitive the modelled and observed overflow volume is to parameter and observation uncertainties for event 9. Each parameter in the table is affected with a large, equally likely, change that could correspond to two times the standard deviation of the parameter uncertainty. The table shows that the roughness of the throttle pipe is the most sensitive parameter, while the roughness of the rest of the pipes in the system and the concentration time of the sub catchments is without importance. Even changing the weir formulae to a side weir does not have a big effect on the modelled overflow either, since the water that arrives at the CSO structure will go over the crest of the weir if the throttle pipe capacity is reached, and the parameters of the weir formulae will only have an impact on the water level at the weir while the overflow happens.

Table 2: Effect of parameter and observation uncertainties on modelled and observed accumulated overflow volume for event 9.

Parameter change	**Change in overflow volume**
Changing throttle pipe material from smooth to normal concreate (13% increase in Manning's n)	+29%
Changing pipe material for all pipes in the model except for the throttle pipe from smooth to normal concreate	-2%
Increase all impervious areas with 5%	+15%
Decrease time of concentration for sub-catchments with 30%	+1%
Change weir formula to side weir	-2%
Overflow observation error	
Change in calculated overflow volume due to bias in water level observations (2 cm).	+20%

Furthermore, the use of a simple Q/h relationship to calculate overflows can even in the case with an optimal Q/h relationship lead to event volume errors of 14% (Isel et al., 2014). Short term errors in the water level observations might occur due to e.g. debris being stuck on the gauge for a while. This sort of error is impossible to detect with only one gauge and the frequency and magnitude of this sort of error is unknown. While the latter two errors are random, all errors in Table 2 are systematic. The random errors should even each other out when quantifying the performance over multiple events, and therefore the median is a more robust estimate than the extremes. Based on the data in Table 2 the total uncertainty of ROV is estimated by assuming that the errors are uncorrelated and normally distributed and by assuming that the changes in ROV caused by changing the various parameters are independent of each other. This enables for

calculating the total ROV uncertainty as the square root of the sum of squares of the data in the right column of Table 2, which leads to a total uncertainty of 40%. The 40% is used as an indication of a 95% confidence interval for the ROV in the following, but keeping in mind the rough assumptions used to produce the number.

The uncertainty of the RMSE has not been assessed explicitly. All the rainfall estimates are affected by the same model and observational errors and these will to a large extent determine the minimum possible level of the RMSE for the individual events, but the errors are not likely to change the relative performance of the individual rainfall estimates. This means that a low RMSE can be seen as an indication of a representative rainfall input to the model and it furthermore implies that it is not strictly required to know the uncertainty of the RMSE to interpret the results. This is, of course, not necessarily true for individual events, and therefore care should be taken in reading too much into single event RMSE.

Since the model is distributed in space and uses the full dynamic wave equations for the hydraulic calculations, both modelled and actual runoff will to some extent be sensitive to most rainfall variations in both space and time. The validation method will, non the less, be less sensitive to temporal rainfall variations at time scales significantly smaller than the response time of the catchment (25 minutes). Therefore a smaller validation catchment should be used if the aim is to assess the quality of radar data on 10 minute time scale or less. In the same way the validation method is less sensitive to spatial variations in the rainfall when the spatial dimension goes below a certain limit. This limit is not easily quantified since it makes a difference in which direction the variation is: it makes a big difference on the runoff hydrograph if the variations are along the flow direction towards the weir and a smaller difference if the variations are perpendicular to this flow direction. These effects are artefacts of validating on catchment scale runoff.

5.2 Adjustment towards gauge A

Since gauge A is situated centrally in the small validation catchment it is expected to represent the rainfall much better than the other gauges, and it is therefore treated separately in this article.

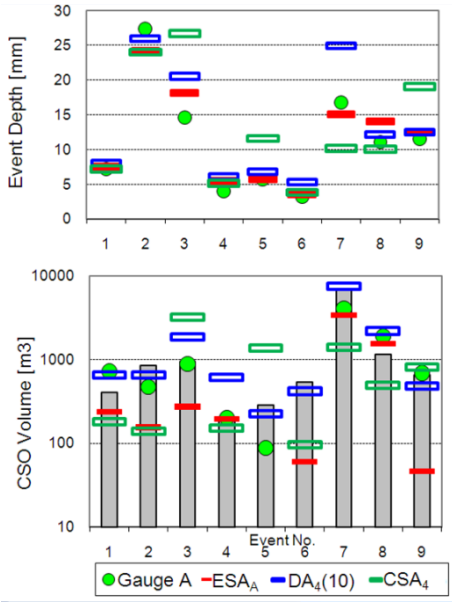


Figure 6: *(Upper figure) Event depth measured by gauge A or by radar for the pixel covering gauge A. (Lower figure) Modelled CSO dependent on which rainfall input is used as model input. This is plotted on top of the measured CSO (grey bars).*

Event CSO volumes and rainfall depths are shown in Figure 6 for selected rainfall products. The figure shows that the event depths as recorded by gauge A are very similar to those of ESA_A for the radar pixel covering gauge A, which might not come as a surprise since the radar data in this case is adjusted towards the very same gauge on event basis. The lower panel in Figure 6 reveals, however, that ESA_A for eight of the nine events leads to less overflow than gauge A, and for five of these events ESA_A leads to less than one third of the volume of overflow found when using gauge A. The gauge data leads to a CSO volume closer to the observed for seven of the nine events which suggests that the gauge data is the most representative for the rainfall over the catchment of the two rainfall estimates. Since the event volumes for the two rainfall estimates are similar, the worse performance of the adjusted radar data must be due to either a poor representation by the radar of the spatial distribution of the rainfall or a poor representation of the dynamics of the rainfall.

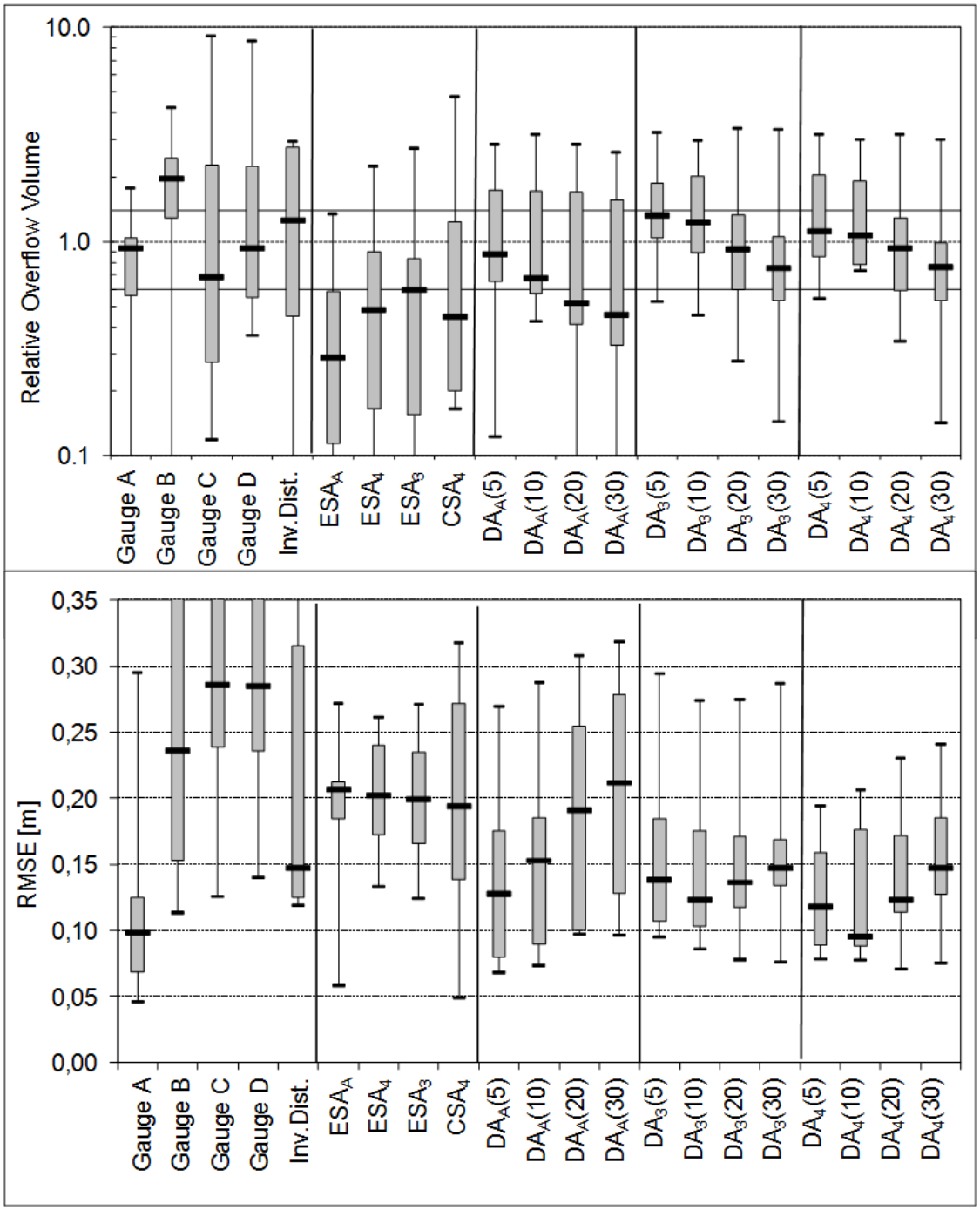


Figure 7: *Box plot showing the relative overflow volume and the RMSE for water levels for the nine events. The plots show the minimum, the 1st quartile, the median (thick line), the 3rd quartile and the maximum values. Where the plot of Relative Overflow Volume goes below 0.1 is due to values of zero (No modelled overflow).The horizontal lines on the ROV plot marks one and the 40% confidence interval around one based on systematic errors only.*

Figure 7 shows descriptive statistics of the results from the CSO comparisons for all events and rainfall products. The lower panel in Figure 7 shows the RMSE between modelled and observed water levels. This performance measure also shows that gauge

A provides a better rainfall estimate than ESA_A. When the radar is adjusted against gauge A on a much shorter time frame the results become much better, which can be seen in Figure 7 when looking at the performance of $DA_A(5)$. The median of the ROV for this rainfall estimate is close to one and it is closer to one than ESA_A for all quantiles. Likewise the RMSE is lower for the 1st to the 3rd quantiles. This shows that the poor performance of ESA_A is not due to a poor spatial representation by the radar data, since this would be the same for $DA_A(5)$. This leaves a poor representation of the rainfall dynamics as the main cause of error. This notion is further supported by the performance of DA_A as T_{ad} grows from 5 towards 30 minutes. The RMSE steadily increases with the increase in T_{ad}, while the median of the ROV falls out of the confidence interval as T_{ad} goes from 10 to 20 minutes. For a T_{ad} of 30 minutes, the DA_A adjustment is not unambiguously better than ESA_A since the RMSE for most of the quantiles is worse for the dynamically adjusted data, while the ROV continues to show better results for DA_A.

The DA methods aim to adjust the radar to better reflect the current conditions, such as e.g. changes in the drop size distribution, at any given point in time. The current conditions over the validation catchments might not be represented the best by the T_{ad} previous minutes above gauge A, since these conditions might be moving away from the catchment with the speed of the rain cell. Therefore current conditions might be better represented by the previous T_{ad} minutes above a larger area as sampled via the three distant rain gauges. The results for DA using only gauge A indicate that it is not a good idea to adjust radar data towards the nearest rain gauge only.

5.3 Adjustment using multiple gauges

From Figure 7 it can be seen that gauge A clearly outperforms the other point gauges in terms of RMSE. The median ROV for gauge C and D are within the confidence interval around ROV of one, which indicates that the statistical properties of the data from these gauges correspond to the rainfall over the validation catchment. The corresponding large spread of the ROV and the very high RMSE values show, however, that the same gauges are not very good at producing representative rainfall estimates for the catchment for the individual events. When a gauge is not present in a catchment a modeller would have to use data from the nearest gauge available or produce an interpolated rainfall estimate from several gauges. Therefore a traditional Inverse Distance (ID) rainfall estimate (Teegavarapu et Al, 2009) based on the three distant gauges (gauge B, C and D) has been included to be able to assess whether it is worthwhile to use rain gauge data to adjust radar data instead of just creating artificial rainfall data from the rain gauges. Figure 7 shows that the ID estimate gives a much better RMSE than the individual gauges situated outside the catchment while there is no clear improvement in the ROV. Generally the ID estimate performs better in terms of ROV than the static adjusted radar data. The comparison is less clear when looking at the RMSE since opposite conclusions can be drawn depending on which quantile is in focus.

The conclusions are clear when comparing the ID with the dynamic adjustment methods based on the same three gauges as the ID estimate, since the radar data in this case performs the best for all quintiles for both ROV and RMSE when DA_3 is based on 20 minutes or less. Furthermore both gauge A and ID only lead to eight overflows while all the multi-gauge DA estimates yield overflows for all events. This shows that it is indeed

worthwhile to combine radar and rain gauge data when estimating rainfall for ungauged locations. When comparing the results for DA_3 and DA_4 it seems that including gauge A into the multi-gauge adjustment scheme only has a minor impact on the performance, except for the maximum RMSE which is improved from the range 0.27-0.29 to 0.19-0.24.

It is worth noting that each of the gauges B, C and D perform poorly by themselves, which can be explained by the spatial variability of the rainfall. When the same gauges are used to estimate the radar to rainfall relationship, as it is done in DA_3, the performance is much better. This shows that the spatial variability of the radar to rainfall relationship is smaller than the variability of the rainfall.

Among all the rainfall estimates $DA_4(10)$ is the most robust in the sense that it provides the lowest spread in the ROV from the minimum to the maximum value, meanwhile having the lowest maximum RMSE and the lowest median RMSE. If looking at any other quantile of the RMSE, however, gauge A performs the best. The gauge is situated within 300 m of the major part of the impervious area that drains towards the validation weir, and within such a short distance the spatial variability of even most convective rainfall is limited. Still the rain gauge completely failed to represent event 6 since the gauge data for this event did not lead to any modelled overflow. That this is not the case for $DA_4(10)$ illustrates the benefit of using multiple gauges for dynamically adjusting radar data, since a single mis-representative gauge in this case does not have a big effect on the final rainfall estimate.

5.4 Importance of adjustment time scale

All the static adjusted radar data result in significant underestimation of the overflow volumes. The typical (median) overflow volume when using ESA_A is only 30 % of the measured volume. The third quantile of the ROV for CSA_4 is actually above one, but as shown in Figure 6, the few large CSO's when using CSA_4 is a result of large overestimations of the total rainfall and thus not a result of a good description of the rain. This shows that the radar data need to be adjusted on time scales shorter than the events used in the current study.

For all the dynamic adjustment schemes the median ROV decreases as T_{ad} increases, and the minimum ROV decreases rapidly when T_{ad} becomes larger than 10 minutes. This indicates that the rainfall peaks are smoothened too much when the DA is based on intervals of 20 minutes or more. Since the three ESA methods are closely related to the three DA methods for long T_{ad}, the performance of the DA methods are expected to converge towards the corresponding ESA values for increasing T_{ad}.

By comparing the performance of DA_3 and DA_4 it appears that the benefit of using four instead of three gauges decreases as T_{ad} increases to 30 minutes. For a T_{ad} of 5 minutes adding the forth gauge improves the RMSE for all quantiles, while for a T_{ad} of 30 minutes only the highest RMSE value is substantially improved and the 3rd quantile even gets worse. Depending on which quantile one is looking at the dynamic adjustment schemes with multiple gauges work best when T_{ad} equals 20 minutes or less. This implies that the method is not suitable for creating nowcasts, since it is not reasonable to assume that the adjustment factor is constant in the forecast period. Just how rapid the factor can change is demonstrated in the next section.

5.5 Focus event

The evening event on the 23rd of August (Event No. 9) was a series of showers and it lasted in all about 7 hours. During that time there where two short periods with overflow, lasting for 30 minutes and 23 minutes, respectively, leading to a total CSO volume of 658 m^3. The recorded event depth is almost exactly the same for gauge A or for ESA_A and $DA_4(10)$ adjusted radar data, see Figure 6, whereas there are big differences in the modelled CSO. Gauge A and $DA_4(10)$ lead to a CSO estimate in the right order of magnitude but ESA_A leads to almost no CSO at all. The reason for this can be seen in Figure 8. The ESA_A adjusted radar data leads to an overestimation of the water level for the first part of the event, meaning that the radar data have overestimated the rainfall, while the rainfall is underestimated in the rest of the event. A higher static adjustment factor could not have improved the picture, since the first part is too high already. $DA_4(10)$ seems to result in a good description of the runoff. The $DA_4(10)$ adjustment factor varies with a factor of more than 6 during the event, see Figure 9, making it possible for the radar data to emulate both peaks and lows better.

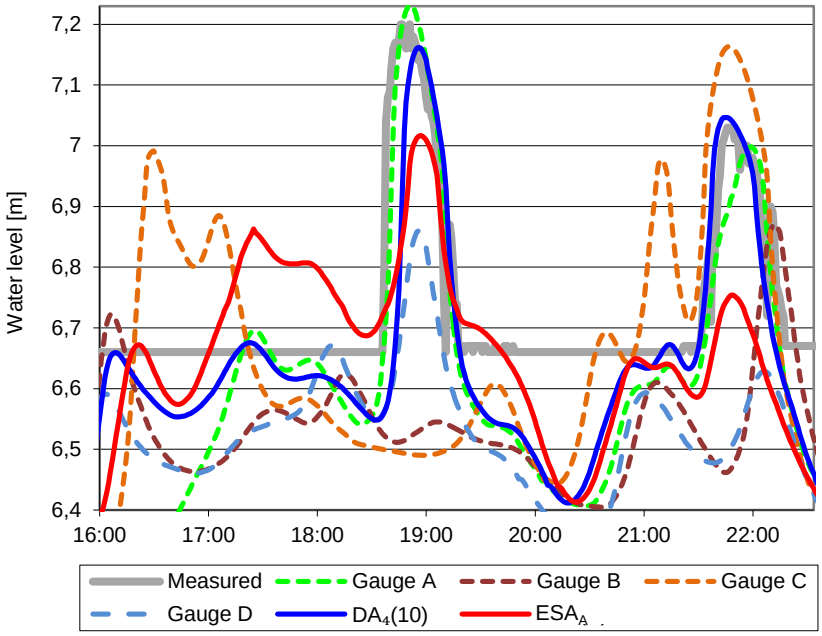


Figure 8: *Measured (grey line) and modelled water level at the downstream weir. The horizontal dotted line represents the level of the weir, and the horizontal grey curve segment represents the lower range of the level gauge signal.*

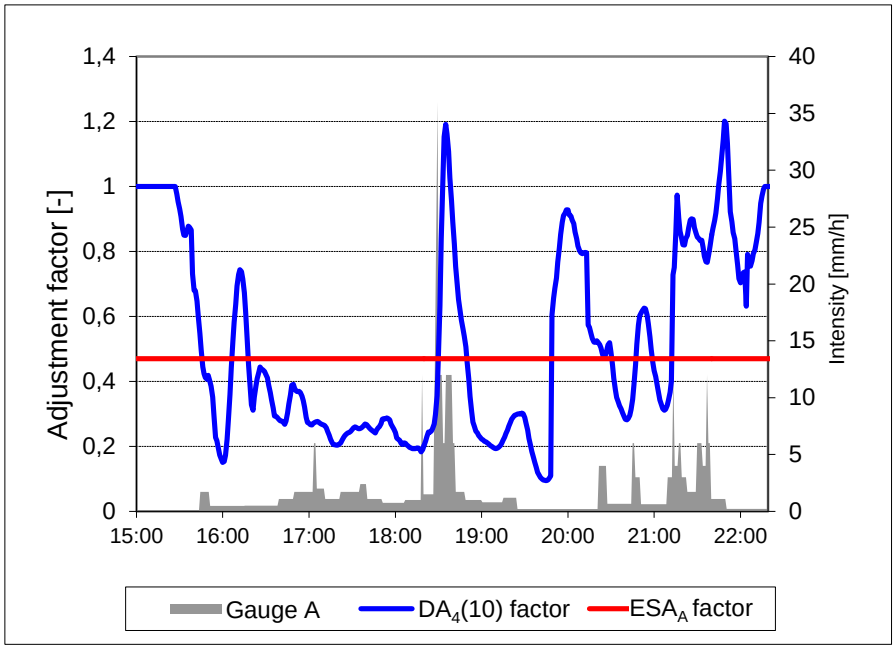


Figure 9: *Adjustment factors for Event 9 and the intensities measured by gauge A.*

6 CONCLUSIONS

In this study, static and dynamic methods for adjusting X-band radar data based on rain gauges are tested for their ability to produce accurate rainfall estimates to a distributed urban runoff model used to model combined sewer overflows. The methods are all linear and constant in space (the value of the adjustment factor is independent of the location). The static methods apply an adjustment factor to the radar rainfall estimates that is constant for an entire event, while the dynamic adjustment methods recalculate the adjustment factor for each minute based on the prior T_{ad} minutes of radar and rain gauge data pairs. The static methods are attempted favoured by allowing them to be based on all data from an event – knowledge that is not available in real time. The primary validation of the rainfall estimates is performed by comparing modelled and measured water levels at a combined sewer overflow (CSO) structure located in a 64 ha catchment with approximately 25 minutes concentration time. The model is constructed purely from physical data without any calibration to avoid favouring the type of rainfall estimate used for the calibration. Data from one rain gauge placed in the centre of the catchment and 3 gauges situated outside the catchment are used for the assessments. Only rain events that lead to significant overflow volumes are included.

The following conclusions can be drawn from the current study:

- Radar rainfall estimates can be used as basis for successful urban runoff modelling required that the radar data are dynamically adjusted against rain gauge observations. The adjustment method works well with only 3 or 4 gauges. The dynamically adjusted radar data performed indisputably better than the three distant gauges, situated 2 – 3 km from the catchment, no matter if these were used individually or all together for estimating areal rainfall using the inverse distance interpolation method.
- The adjustment works best when based on the past 10 – 20 minutes of rainfall data only. If longer time periods are used, the dynamics of the rainfall will not

be adequately described by the rainfall estimate which results in e.g. lower modelled CSO volumes. The short time frame implies that the dynamic adjustment scheme is not suitable for producing nowcasts/forecasts.

- The rain gauge within the validation catchment provided the best rainfall product if disregarding the single event where data from this rain gauge did not lead to a modelled CSO at all. This shows that it is still worthwhile to set up rain gauges in areas of special interest.

Even in cities that are well equipped with rain gauges, the majority of the areas will not be within 300 m of a rain gauge. For these areas the current study indicate that radar rainfall estimates can be improved to the point where they can be used for quantitative CSO modelling by dynamically adjusting the radar data against rain gauges, when the adjustment period is as short as 10 to 20 minutes.

ACKNOWLEDGEMENT

The work with this article was supported by the Danish Council for Strategic Research, Program Commission on Sustainable Energy and Environment through the Storm and Wastewater Informatics project (http://www.swi.env.dtu.dk). VCS Denmark is acknowledged for providing the radar data as well as the excellent physical data for the MOUSE model.

7 REFERENCES

Atlas, D., 1990. Radar in meteorology. Boston, MA, American Meteorological Society, 1990, 822.

Austin, G.L., Bellon, A., 1974. The use of digital weather radar records for short-term precipitation forecasting. Quarterly Journal of the Royal Meteorological Society 100, 658–664.

Beard, K. V, 1976. Terminal velocity and shape of cloud and precipitation drops aloft. J. Atmos. Sci 33, 851–864.

Berne, A., Delrieu, G., Creutin, J.-D., Obled, C., 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. Journal of Hydrology 299, 166–179.

Berne, A., Krajewski, W.F., 2013. Radar for hydrology: Unfulfilled promise or unrecognized potential? Advances in Water Resources 51, 357–366.

Chapon, B., Delrieu, G., Gosset, M., Boudevillain, B., 2008. Variability of rain drop size distribution and its effect on the Z–R relationship: A case study for intense Mediterranean rainfall. Atmospheric Research 87, 52–65.

Cifelli, R., Williams, C.R., Rajopadhyaya, D.K., Avery, S.K., Gage, K.S., May, P.T., 2000. Drop-size distribution characteristics in tropical mesoscale convective systems. Journal of Applied Meteorology 39, 760–777.

Clemens, M., Peters, G., Seltmann, J., Winkler, P., 2006. Time-height evolution of measured raindrop size distributions, in: Proceeding of ERAD 2006, Barcelona.

Cole, S.J., Moore, R.J., 2008. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. Journal of Hydrology 358, 159–181.

Creutin, J.., Andrieu, H., Faure, D., 1997. Use of a weather radar for the hydrology of a mountainous area. Part II: radar measurement validation. Journal of Hydrology 193, 26–44.

Dai, Q., Han, D., 2014. Exploration of discrepancy between radar and gauge rainfall estimates driven by wind fields. Water Resources Research 50, 8571–8588.

DHI, 2007. MOUSE PIPE FLOW Reference Manual. DHI Software, Horsholm, Denmark.

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G., Vieux, B., 2004. Towards a roadmap for use of radar rainfall data in urban drainage. Journal of Hydrology 299, 186–202.

Goormans, T., Willems, P., 2012. Using local weather radar data for sewer system modelling: a case study in Flanders, Belgium. Journal of Hydrologic Engineering. Journal of Hydrologic Engineering 18.2, 269-278.

Goudenhoofdt, E., Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrology and Earth System Sciences 13, 195–203.

Isel, S., Dufresne, M., Fischer, M., Vazquez, J., 2014. Assessment of the overflow discharge in complex CSO chambers with water level measurements – On-site validation of a CFD-based methodology. Flow Measurement and Instrumentation. 35, 39–43.

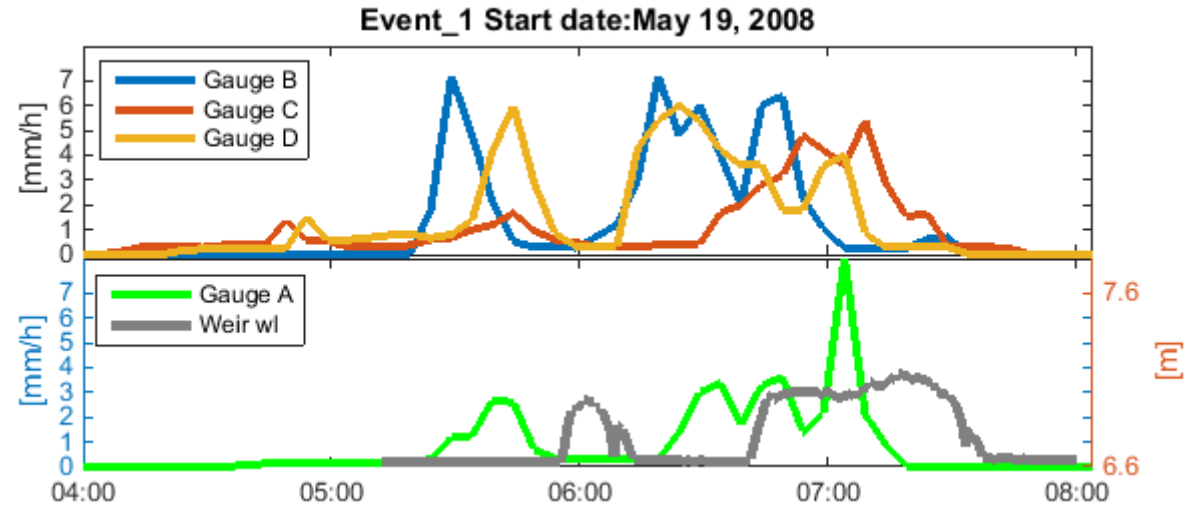
Jensen, N.E., 2000. Local Area Weather Radar Documentation v. 3.0 (2004), DHI. Institute for the water environment.

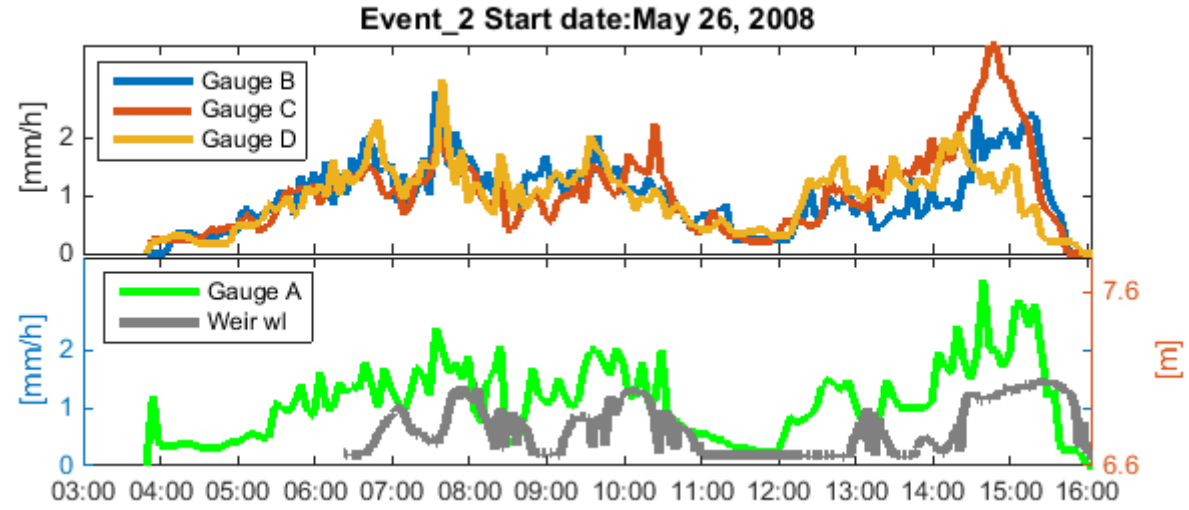
Jørgensen, H.K.H., Rosenørn, S., Madsen, H., Mikkelsen, P.S., 1998. Quality control of rain data used for urban runoff systems. Water Science and Technology 37, 113–120.
Krämer, S., Verworn, H.R., 2009. Improved radar data processing algorithms for quantitative rainfall estimation in real time. Water science and technology 60, 175–84.
Lee, G., Zawadzki, I., 2006. Radar calibration by gage, disdrometer, and polarimetry: Theoretical limit caused by the variability of drop size distribution and application to fast scanning operational radar data. Journal of Hydrology 328, 83–97.
McMillan, H., Krueger, T., Freer, J., 2012. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. Hydrological Processes 26, 4078–4111.
Morin, E., Enzel, Y., Shamir, U., Garti, R., 2001. The characteristic time scale for basin hydrological response using radar data. Journal of Hydrology. 252 (1–4), 85–99.
Nielsen, J.E., Jensen, N.E., Rasmussen, M.R., 2013. Calibrating LAWR weather radar using laser disdrometers. Atmospheric Research 122, 165–173.
Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R.D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., ten Veldhuis, M.-C., 2015. Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. Journal of Hydrology 531, 389–407.
Pedersen, L., Jensen, N.E., Madsen, H., 2010. Calibration of Local Area Weather Radar—Identifying significant factors affecting the calibration. Atmospheric Research 97, 129–143.
Rinehart, R.E., 1991. Radar for meteorologists. Department of Atmospheric Sciences, Center for Aerospace Sciences, University of North Dakota.
Sauvageot, H., 1992. Radar meteorology. Artech House Publishers.
Seo, B.-C., Krajewski, W.F., 2010. Scale Dependence of Radar Rainfall Uncertainty: Initial Evaluation of NEXRAD's New Super-Resolution Data for Hydrologic Applications. Journal of Hydrometeorology 11, 1191–1198.
Shrestha, N., Goormans, T., Willems, P., 2013. Evaluating the accuracy of C and X band weather radars and their application for stream flow simulation. Journal of Hydroinformatics 15, 1121–1136.
Smith, J. a., Hui, E., Steiner, M., Baeck, M.L., Krajewski, W.F., Ntelekos, A. a., 2009. Variability of rainfall rate and raindrop size distributions in heavy rain. Water Resources Research 45.
Smith, J. a., Krajewski, W.F., 1993. A modeling study of rainfall rate-reflectivity relationships. Water Resources Research 29, 2505–2514.
Smith, J.A., Krajewski, W.F., 1991. Estimation of the mean field bias of radar rainfall estimates. Journal of Applied Meteorology 30, 397–412.
Teegavarapu, R. S., Tufail, M., & Ormsbee, L., 2009. Optimal functional forms for estimation of missing precipitation data. Journal of Hydrology 374, 106-115.
Thorndahl, S., Nielsen, J.E., Rasmussen, M.R., 2014. Bias adjustment and advection interpolation of long-term high resolution radar rainfall series. Journal of Hydrology 508, 214–226.
Thorndahl, S., Rasmussen, M.R., 2012. Marine X-band weather radar data calibration. Atmospheric Research 103, 33–44.

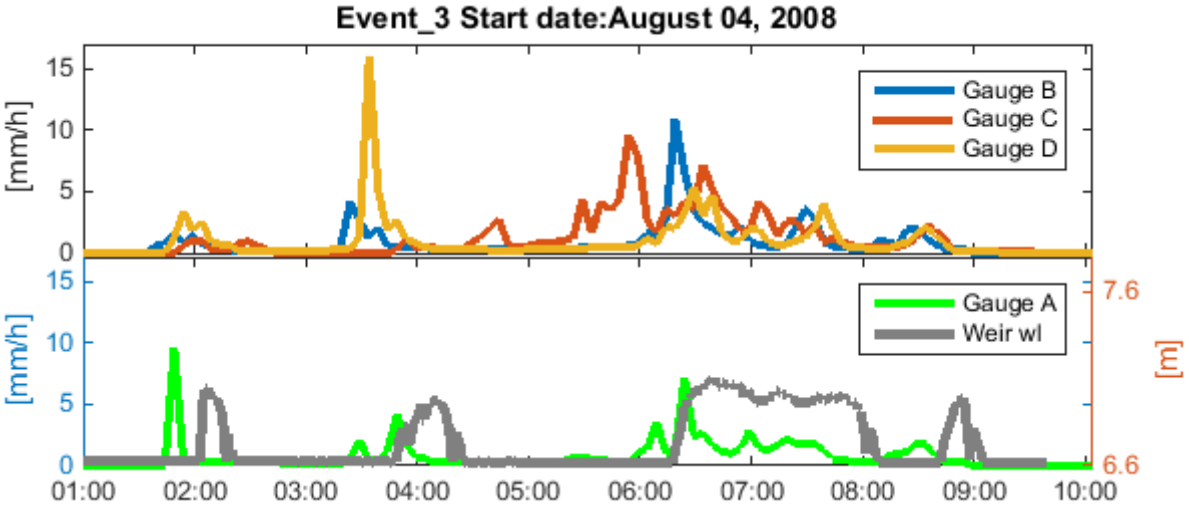
Van de Beek, C.Z., Leijnse, H., Stricker, J.N.M., Uijlenhoet, R., Russchenberg, H.W.J., 2010. Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. Hydrology and Earth System Sciences 14, 205–221.
Willems, P., Molnar, P., Einfalt, T., Arnbjerg-Nielsen, K., Onof, C., Nguyen, V.-T.-V., Burlando, P., 2012. Rainfall in the urban context: Forecasting, risk and climate change. Atmospheric Research 103, 1–3.
Wood, S.J., Jones, D.A., Moore, R.J., 2000. Static and dynamic calibration of radar data for hydrological use. Hydrology and Earth System Sciences Discussions 4, 545–554.

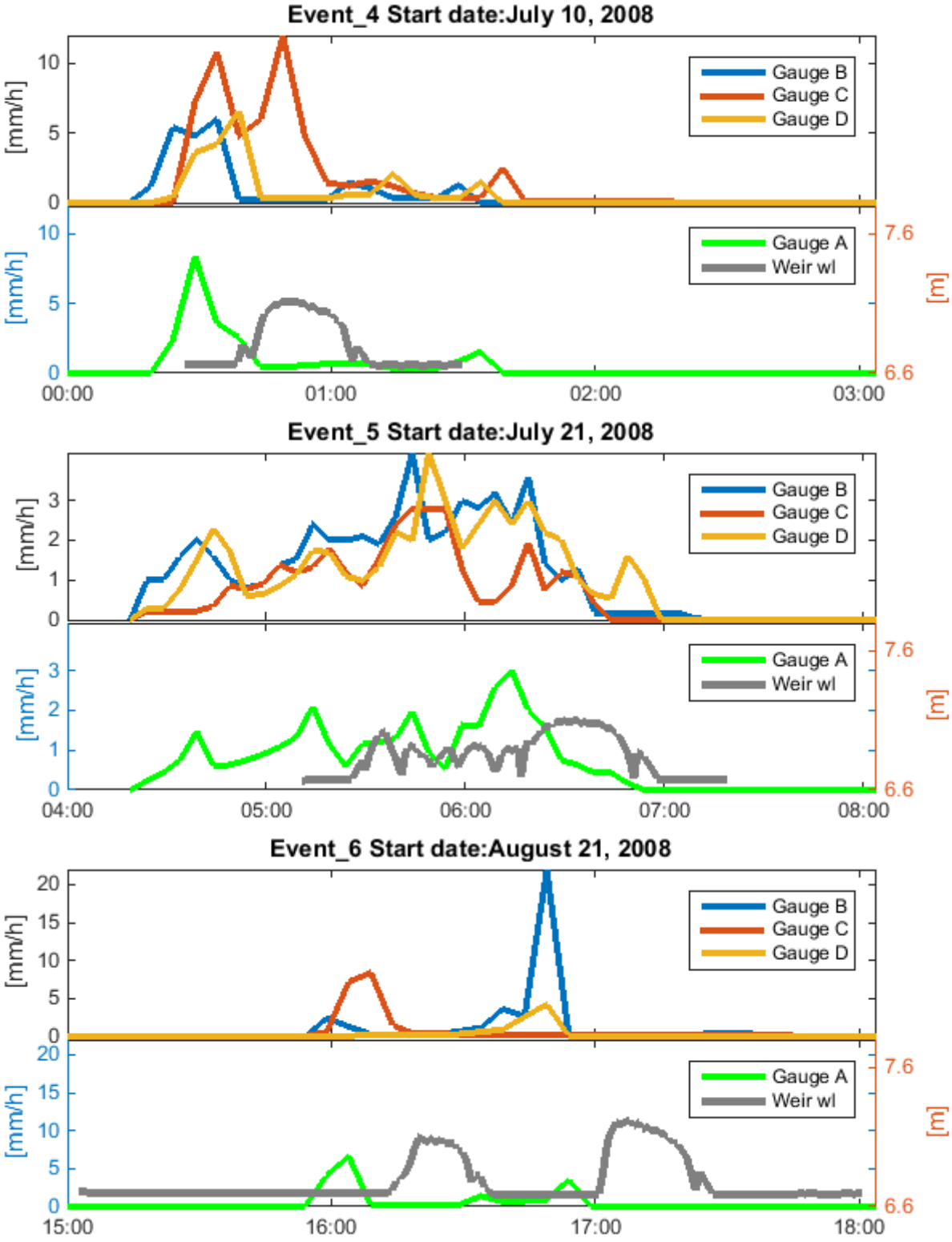
Appendix: Rain gauge and weir water level data

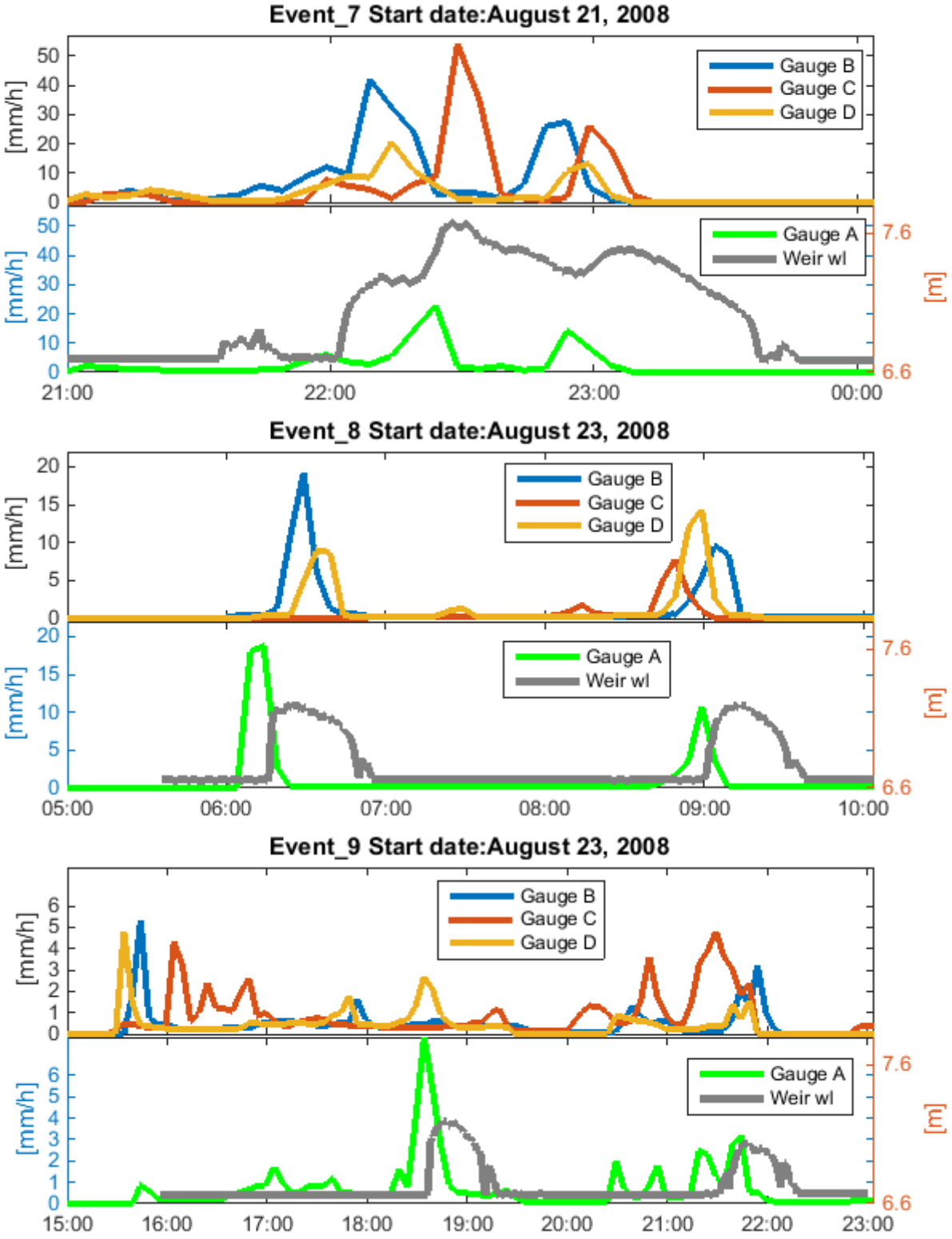
The rain gauge data are aggregated to 5 minute averages for better readability. The data for the rain gauge situated within the catchment (gauge A) is plotted in the same figure as the observed weir water level.








Event_4 Start date:July 10, 2008
[mm/h]
Gauge B
Gauge C
Gauge D
Gauge A
Weir wl
[m]
7.6
6.6
00:00
01:00
02:00
03:00
Event_5 Start date:July 21, 2008
04:00
05:00
06:00
07:00
08:00
Event_6 Start date:August 21, 2008
15:00
16:00
17:00
18:00


Event_7 Start date:August 21, 2008
[mm/h]
Gauge B
Gauge C
Gauge D
Gauge A
Weir wl
[m]
7.6
6.6
21:00
22:00
23:00
00:00
Event_8 Start date:August 23, 2008
05:00
06:00
07:00
08:00
09:00
10:00
Event_9 Start date:August 23, 2008
15:00
16:00
17:00
18:00
19:00
20:00
21:00
22:00
23:00