Hydrology and
Earth System
Sciences

# The effect of flow and orography on the spatial distribution of the very short-term predictability of rainfall from composite radar images

L. Foresti[1,2,*] and A. Seed[2]

[1]Royal Meteorological Institute of Belgium, Brussels, Belgium
[2]Bureau of Meteorology, Centre for Australian Weather and Climate Research, Melbourne, Australia
[*]now at: Royal Meteorological Institute of Belgium, Brussels, Belgium

*Correspondence to:* L. Foresti (loris.foresti@meteo.be)

**Abstract.** The spatial distribution and scale dependence of the very short-term predictability of precipitation by Lagrangian persistence of composite radar images is studied under different flow regimes in connection with the presence of orographic features. Data from the weather radar composite of eastern Victoria, Australia, a $500 \times 500$ km$^2$ domain at 10 min temporal and $2 \times 2$ km$^2$ spatial resolutions, covering the period from February 2011 to October 2012, were used for the analyses. The scale dependence of the predictability of precipitation is considered by decomposing the radar rainfall field into an eight-level multiplicative cascade using a fast Fourier transform. The rate of temporal development of precipitation in Lagrangian coordinates is estimated at each level of the cascade under different flow regimes, which are stratified by applying a k-means clustering algorithm on the diagnosed velocity fields. The predictability of precipitation is measured by its lifetime, which is derived by integrating the Lagrangian auto-correlation function. The lifetimes were found to depend on the scale of the feature as a power law, which is known as dynamic scaling, and to vary as a function of flow regime. The lifetimes also exhibit significant spatial variability and are approximately a factor of 2 longer on the upwind compared with the downwind slopes of terrain features. The scaling exponent of the spatial power spectrum also shows interesting geographical differences. These findings provide opportunities to perform spatially inhomogeneous stochastic simulations of space–time precipitation to account for the presence of orography, which may be integrated into design storm simulations and stochastic precipitation nowcasting systems.

## 1 Introduction

The scale dependence of the predictability of the atmospheric flow was already studied by Lorenz (1969), who found that there is an intrinsic predictability limit associated to each scale of motion. Similar conclusions can also be extended to the predictability of precipitation, in particular if considering rainfall fields as emerging from multiplicative cascade processes (Schertzer and Lovejoy, 1987; Marsan et al., 1996).

The intuition that large-scale precipitation features are more predictable than small-scale features can be easily verified empirically using both Lagrangian persistence of radar precipitation patterns and outputs from numerical weather prediction (NWP) models. Zawadzki et al. (1994) found that the decorrelation time of radar precipitation patterns by Lagrangian persistence is dependent on the degree of spatial smoothing. Grecu and Krajewski (2000) also detected that the predictability depends on precipitation intensity, the most intense rain rates being less predictable. Seed (2003) studied the scale dependence of the predictability of precipitation by Lagrangian persistence using a fast Fourier transform (FFT) to decompose the radar rainfall field into a multiplicative cascade. Turner et al. (2004) employed a wavelet-based decomposition to filter out the unpredictable scales of a radar-based extrapolation technique. Wavelet decompositions were also exploited for the scale-dependent verification of NWP precipitation forecasts to account for the loss of predictability at small scales (e.g., Casati et al., 2004; Bousquet et al., 2006). Sinclair and Pegram (2005) applied an Empirical Mode Decomposition to iteratively decompose the

precipitation field into meaningful physical structures from the high to the low frequencies. Surcel et al. (2014) used a Discrete Cosine Transform to study the filtering properties of ensemble averaging and discovered that the ensemble members are completely decorrelated below a certain cutoff scale.

The multifractal and scale-dependent nature of rainfall not only complicates the study of its predictability and the verification of forecasts, but also demands more sophisticated forecasting and downscaling techniques. The Short-Term Ensemble Prediction System (STEPS; Seed, 2003; Bowler et al., 2006) is a stochastic precipitation nowcasting scheme that exploits the multifractal principle by decomposing the radar rainfall field into an eight-level multiplicative cascade with an FFT. The cascade is advected with optical flow in Lagrangian coordinates and stochastically evolves in time according to a hierarchy of auto-regressive processes of order 1 – AR(1) – or 2 – AR(2). This allows accounting for the empirical observation that the rate of temporal evolution of precipitation features is a power law of the scale of the feature, which is known as dynamic scaling (see, e.g., Venugopal et al., 1999; Mandapaka et al., 2009). STEPS estimates the rate of Lagrangian development of the cascade levels in real time, which allows adapting to the predictability of the observed sequence of radar images. This is necessary since the predictability of precipitation exhibits a strong temporal variability as shown by Seed (2003), Germann et al. (2006), and Seed et al. (2013).

Germann et al. (2006) also analyzed the geographical distribution of the predictability of precipitation over the conterminous United States and found a region of longer lifetimes extending from eastern Nebraska to Lake Michigan through Iowa, Wisconsin, and northern Illinois. Berenguer and Sempere-Torres (2013) performed a similar analysis using the European radar composite and discovered the predictability to be seasonally dependent, with higher values over the central part of the UK, central continental Europe, and the Baltic regions. However, such geographical differences are strongly affected by the inhomogeneous quality of the European radar composite between the different countries, which use different hardware, operating wavelength, scanning strategy, and signal processing (Huuskonen et al., 2014). The spatial heterogeneity of the statistical properties of rainfall also poses issues for its multifractal simulation, which traditionally assumes spatial homogeneity of the stochastic process. One way to avoid constructing complicated, spatially heterogeneous models is to separately add a spatial trend to correct a homogeneous multifractal model. This trend should account for the spatial inhomogeneity of the long-term climatological distribution of precipitation, which is often controlled by the presence of orographic features (see, e.g., Pathirana and Herath, 2002; Badas et al., 2006).

The climatology of precipitation over complex orography is strongly controlled by flow direction and air stability (Panziera and Germann, 2010), which can also be exploited

to design analogue-based nowcasting techniques (Foresti et al., 2013). The contribution of orography to the precipitation enhancement also seems to be a scale-dependent process. This can be observed by extracting features from a digital elevation model (DEM) at different spatial scales and looking at the spatial distribution of persistent precipitation cells. It appears that orographic features need a certain characteristic size (scale) in order to control the spatial distribution of precipitation patterns (e.g., Foresti et al., 2012).

The goal of this study is to analyze the spatial distribution of the scale-dependent predictability of precipitation by Lagrangian persistence of composite radar images under different flow regimes in connection with the presence of orographic features. Data from the weather radar composite of eastern Victoria, Australia, a $500 \times 500\,\mathrm{km}^2$ domain at 10 min temporal and $2 \times 2\,\mathrm{km}^2$ spatial resolutions, covering the period from February 2011 to October 2012, are used for the analyses. A k-means clustering algorithm is employed to classify the velocity fields into six main flow regimes and to stratify the evaluation of statistics.

This research is an extension of the study of Foresti and Seed (2014), who analyzed the geographical distribution of the STEPS nowcasting biases using the same radar data set in order to detect regions of systematic precipitation growth and decay. The typical areas of rainfall growth and decay due to orographic forcing should be observed also in the spatial distribution of the predictability of rainfall. The orographic forcing is expected to control the spatial distribution of the predictability of precipitation at the meso-gamma (2–20 km) and partly the meso-beta (20–200 km) scales, which are smaller than the continental scales analyzed in the literature (e.g., Germann et al., 2006; Radhakrishna et al., 2012; Berenguer and Sempere-Torres, 2013).

The dependence of the dynamic scaling relationship on flow regimes is also studied to test whether there are weather regimes that are more predictable than others. On the other hand, the geographical distribution of the spatial power spectrum is analyzed to explore the degree of spatial scaling of precipitation over the forecast domain. The findings of this study should increase our understanding of the predictability of precipitation by Lagrangian persistence of radar images, which is essential to improve its very short-term forecasting, space–time stochastic simulation, and statistical downscaling.

The paper is structured as follows. Section 2 describes the radar rainfall data set. Section 3 details the methodology. Section 4 illustrates the obtained results and interpretations, while Sect. 5 concludes the paper, and discusses potential improvements and future research perspectives.

## 2  Radar rainfall data set

Data from the weather radar composite of Eastern Victoria, Australia were used for the analyses (see Fig. 1 for the

domain and the radar locations). The composite merges data from four weather radars located at Melbourne (operating at S-band), Yarrawonga (C-band), Gippsland (C-band) and Canberra–Captains Flat (S-band). The period under analysis is from 15 February 2011 to 31 October 2012.

The operational radar data processing chain for quantitative precipitation estimation (QPE) at the Australian Bureau of Meteorology consists of the following steps:

– Ground clutter removal with Doppler filtering at the radar site.

– Additional ground clutter filtering based on a static clutter map and on the gradients of the vertical profile of reflectivity.

– Beam blockage correction using a DEM to correct for the lost power due to the interception of the radar beam with orography.

– Estimation of the vertical profile of reflectivity using data within a range of 50 km from the radar.

– Interpolation of the volumetric data into constant altitude plan position indicators (CAPPIs). CAPPIs are computed at a height of 1000 m using the 3-dimensional anisotropic Kriging technique of Seed and Pegram (2001).

– Application of a different climatological $Z$–$R$ relationship for stratiform and convective rain based on the Steiner classification (Chumchean et al., 2008).

– Compositing operation involving a linear combination of the radar measurements in the overlapping regions as a function of distance from the radar.

– Mean field bias correction with respect to rain gauge measurements using a Kalman filtering approach for its temporal update (Chumchean et al., 2006b).

The final product is a $256 \times 256$ grid with a spatial resolution of $2 \, \text{km}^2 \times 2 \, \text{km}^2$ and a temporal resolution of 10 min in a Gnomonic projection. More details on the operational QPE chain at the Australian Bureau of Meteorology are given in Chumchean et al. (2006a, b, 2008) and Seed et al. (2007).

These pre-processing steps are not sufficient to completely remove the radar measurement errors, especially over mountainous regions. The two sources of errors that are the most critical for the analysis of the precipitation predictability are the range dependence of estimated rainfall rates and the reduced visibility in the inner Victorian Alps. In addition, the compositing operation generates some discontinuities in the regions of overlapping radar measurements. Rainfall could also be slightly underestimated in a radius of $\sim$ 20–30 km around the radar due to the excessive filtering of ground clutter, which also eliminates some precipitation measurements. Precipitation is also underestimated at ranges exceeding 90–100 km due to the increasing beam width (sampling
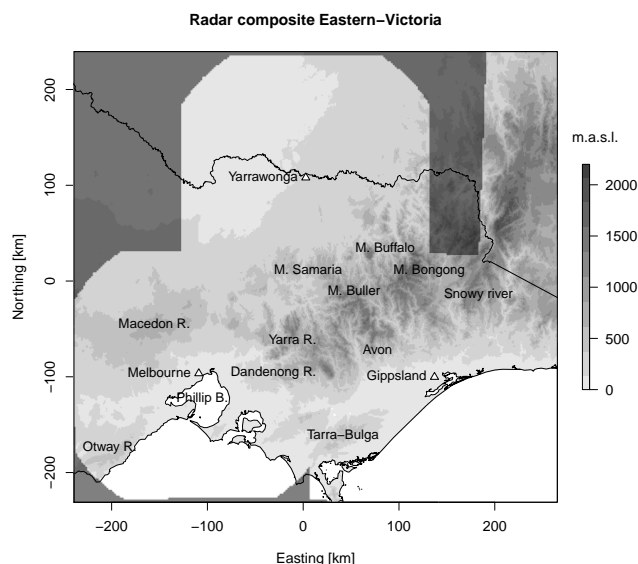


**Figure 1.** Radar composite of Eastern Victoria, Australia, overlaid on the DEM. Triangles denote the locations of the three radars at Melbourne, Yarrawonga, and Gippsland. In the top-right corner of the domain there is some contribution from the Canberra radar. White tones represent the ocean.

volume), attenuation by rainfall and blockage by orographic features. Hence, precipitation accumulations are strongly underestimated in the inner part of the Victorian Alps where the correction for the vertical profile of reflectivity is evidently not sufficient to extrapolate the higher elevation measurements to the elevation of the CAPPI.

## 3 Methodology

Section 3.1 explains the cascade decomposition framework for the analysis of the scale dependence of the predictability of precipitation. Section 3.2 details the method for estimating the Lagrangian temporal auto-correlation of precipitation, which is needed to evaluate its lifetime (Sect. 3.3). The simultaneous calculation of the Lagrangian auto-correlation at each point of the radar grid using rules for the online computation of the covariance is presented in Sect. 3.4. Section 3.5 presents a simplified approach to estimate the slope of the power spectrum from the variance of the cascade levels under the scaling hypothesis. Finally, Sect. 3.6 provides a brief summary of the k-means stratification of optical flow fields.

### 3.1 Cascade decomposition framework

The radar rainfall field is decomposed using an FFT into a multiplicative cascade of the form (Seed, 2003; Bowler et

al., 2006):

$$\mathrm{dBR}_{ij} = \sum_{k=0}^{K-1} X_{kij} \text{ for } i = 1, \dots, L \text{ and } j = 1, \dots, L, \quad (1)$$

where $L = 384$ is the size of the squared domain and $K = 8$ is the number of cascade levels. A buffer of 64 pixels is added at each side of the original $256 \times 256$ grid in an attempt to reduce the edge effects arising from the FFT transformation, thus giving a larger domain of $384 \times 384$ pixels. The cascade is multiplicative when rewritten in terms of original rain rates $R$ instead of the multiplicative decibel scale dBR. The cascade decomposition is achieved by applying a Gaussian band-pass filter to isolate a given set of spatial scales in the frequency domain (Seed, 2003; see Fig. 2). $X_k$ will be referred to as cascade level and is obtained by applying an inverse FFT to the filtered data in order to return the Fourier components back into the spatial domain. Thus, $X_k$ represents the variability of the original radar field with spatial frequencies [km$^{-1}$] in the range $q^{k-1}/L < \omega_k < q^{k+1}/L$, where $\omega_k$ is the central frequency of the Gaussian filter and $q = 2.12$ is the branching number (inverse of the scale reduction factor). Each level of the cascade is normalized to zero mean and unit variance for convenience and the normalization is kept constant in space and during the forecast period.

Figure 2 illustrates the Gaussian band-pass filters that are used to isolate the spatial scales composing the set of cascade levels. Given the size of the extended radar domain, an eight-level multiplicative cascade with the following spatial scales is obtained (see Fig. 2): 768–362, 362-*171*-81, 171-*81*-38, 81-*38*-18, 38-*18*-8, 18-*8*-4, 8-*4*-2 and 4-*2* km. The non-integer scales resulting from the non-integer branching number of 2.12 were rounded. The scales on which the Gaussian filters are centered are marked in italic. The first and last levels of the cascade will not be considered in the analyses because of not having a regular Gaussian shape. In addition, the largest scale is not able to capture the appropriate scales since the radar composite only covers a certain fraction of the $512 \times 512$ km$^2$ domain. This would lead to the underestimation of the precipitation lifetime at that scale (see Sect. 3.3).

## 3.2 Lagrangian temporal auto-correlation

The Lagrangian temporal auto-correlation is a measure for the rate of development of precipitation in storm coordinates and consequently of its predictability (Zawadzki, 1973). An efficient way to follow the rainfall evolution in storm coordinates is to estimate a velocity field using a sequence of radar rainfall fields. STEPS uses an optical flow algorithm (Bowler et al., 2004) for the estimation of the velocity field and a semi-Lagrangian backward-in-time scheme for its advection, which keeps the velocity field fixed and retrieves the rainfall values upstream by following the lines of the velocity field (e.g., Germann and Zawadzki, 2002).
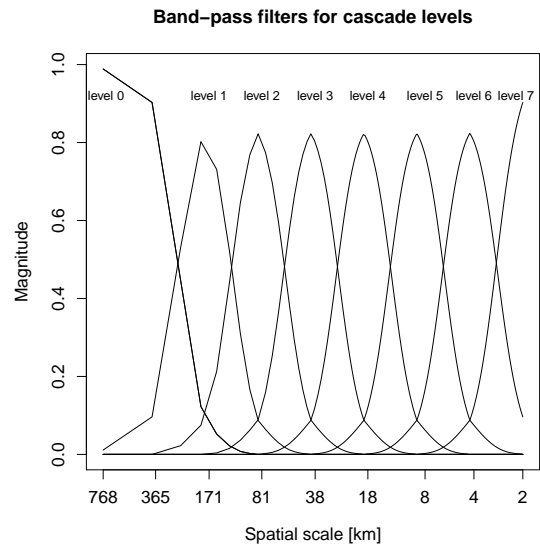


**Figure 2.** The set of eight Gaussian band-pass filters used to isolate the spatial frequencies composing the cascade levels. The total magnitude for a given spatial frequency is normalized to one.

The Lagrangian lag 1 temporal auto-correlations at each level of the cascade are estimated as follows (Bowler et al., 2006):

1. Estimate the velocity field with optical flow using rainfall fields at time $t - 1$ and $t$.

2. Decompose the radar rainfall field at time $t - 1$ using FFT into a multiplicative cascade.

3. Decompose the radar rainfall field at time $t$ using FFT into a multiplicative cascade.

4. Advect the cascade from time $t - 1$ to time $t$. Note that each level of the cascade is advected with the same velocity field computed on the original rainfall fields.

5. The lag 1 Lagrangian temporal auto-correlation is simply obtained by computing the correlation coefficient between each cascade level $k$ advected from time $t - 1$ to $t$ and the corresponding cascade level at time $t$:

$$\rho_1(k) = \frac{\frac{1}{L \cdot L} \sum_{i=1}^{L} \sum_{j=1}^{L} \left( X_{kij} - \overline{X}_k \right) \cdot \left( X_{kij}^{\mathrm{adv}} - \overline{X}_k^{\mathrm{adv}} \right)}{\sqrt{\frac{1}{L \cdot L} \sum_{i=1}^{L} \sum_{j=1}^{L} \left( X_{kij} - \overline{X}_k \right)} \sqrt{\frac{1}{L \cdot L} \sum_{i=1}^{L} \sum_{j=1}^{L} \left( X_{kij}^{\mathrm{adv}} - \overline{X}_k^{\mathrm{adv}} \right)}}$$
$$\text{for } k = 1, \dots, K - 1, \quad (2)$$

where $L = 256$ is the size of the radar domain and "adv" refers to the previous value advected forward to the current time. The smaller the correlation coefficients, the higher are the growth and decay of rainfall processes occurring in Lagrangian frame of reference. The lag 2 Lagrangian temporal auto-correlation could be estimated

as well by advecting a cascade at time $t-2$ to time $t$, but is not presented in this paper.

Equation (2) is the ordinary Pearson's correlation coefficient, which involves the subtraction of the field mean. On the other hand, Zawadzki (1973) and Germann and Zawadzki (2002) employed a correlation estimation without subtraction of the mean for estimating the decorrelation time of precipitation fields. The difference between the two approaches is not very important over continental scales, where the forecast and observed fields have similar mean values, but it may become an issue over smaller domains, where the observed mean field precipitation can be significantly different than the forecast one (see, e.g., Foresti et al., 2012). In such a case, Eq. (2) would give lower but more realistic correlation coefficients compared with Germann and Zawadzki (2002).

The Lagrangian auto-correlation estimations are also affected by the presence of different scales of motion. A multiscale optical flow estimation at each level of the cascade may be foreseen but could cause algorithm convergence issues when one is trying to correlate the small-scale features. Also, it is not yet clear how to avoid the appearance of artifacts in the final reconstructed rainfall field when advecting the cascade levels with different velocity fields over several time steps.

Note that the correlation function of Eq. (2) is obtained by integrating over space, i.e., over the total number of pixels $L \cdot L$ within a radar image. This allows the Lagrangian auto-correlation to be estimated in real time and to adapt to the predictability of the sequence of radar images. This approach, however, assumes the predictability to be homogeneous over the forecast domain. Section 3.4 will explain how to obtain estimates of the Lagrangian auto-correlation by performing the summations through time, which is a necessary step for analyzing its spatial distribution.

The hierarchy of Lagrangian temporal auto-correlations defines a hierarchy of auto-regressive processes of order 1 – AR(1). This is exploited by STEPS to stochastically simulate the rainfall growth and decay processes that occur in storm coordinates at different spatial scales to reproduce the dynamic scaling of the field (Seed, 2003; Bowler et al., 2006). The procedure consists of blending the radar cascade with a cascade of spatially and temporally correlated stochastic noise. The spatially correlated noise field is generated using a power law filter while temporal correlations are maintained by a hierarchy of auto-regressive processes. The power law filter ensures that the noise cascade has the same power spectrum of the observed radar rainfall fields. This technique was already employed to generate continuous multifractals (Schertzer and Lovejoy, 1987) and also appeared in the nowcasting system SBMcast (Berenguer et al., 2011), based on the "String of Beads" model of Pegram and Clothier (2001a). The stochastic simulations are stationary and no attempt is made to actually forecast temporal trends in growth and decay of precipitation. Indeed, trying to predict growth and de-

cay processes using as predictor the past evolution of radar precipitation does not seem to significantly improve the forecast accuracy, except for the regions characterized by systematic orographic forcing (see a review in Foresti and Seed, 2014). In addition, Radhakrishna et al. (2012) showed that the predictability of growth and decay patterns is 10 times shorter than that of precipitation fields and is limited to spatial scales of the order of $250 \times 250\,\mathrm{km}^2$, which would require continental-scale radar images to be studied properly. The stochastic simulations are not presented in this paper but only explained for completeness since they are based on the Lagrangian auto-correlation coefficients.

## 3.3 Estimation of the precipitation lifetime

By knowing the lag 1 auto-correlation coefficient, the AR(1) auto-correlation function (ACF) can be recursively derived as follows:

$$\rho(t) = \rho_1^t \text{ for } t = 1, \ldots, T, \tag{3}$$

where $\rho_1$ is the lag 1 Lagrangian auto-correlation coefficient computed with Eq. (2). Note that this simplification indirectly assumes that the diagnosed velocity field does not change during the forecast period. In fact, it extrapolates the whole ACF knowing only the lag 1 auto-correlation. This assumption is reasonable up to 2–3 h (Germann et al., 2005) and 3–4 h lead times (Bowler et al., 2006), since using the correct velocity does not reduce the forecast errors much. A complete study of the Lagrangian predictability of precipitation including the non-stationarity of the velocity field, would involve the direct calculation of the correlation coefficients at each forecast lead time by comparing the forecasts to the observations (see Germann and Zawadzki, 2002). The basic principle of STEPS is to actually estimate the Lagrangian ACF in real time and allow it to adapt to the predictability of the situation. It would be computationally intensive to estimate the complete ACF using a few hours of radar fields before the analysis time. Eventually, the predictability of the field would be representative of the previous hours and not of the last two or three rainfall fields. The adaptability of the system is particularly important, for example, when the field is rapidly evolving from a convective to a stratiform situation or in the early stages of a new rainfall event.

Finally, the lifetime of precipitation (decorrelation time) can be evaluated by integrating the ACF over time (Zawadzki, 1973; Germann and Zawadzki, 2002):

$$LT = \int_0^\infty \rho(t)\mathrm{d}t. \tag{4}$$

For an exponentially decaying ACF the lifetime is defined as the time at which the ACF falls below the value $1/e = 0.37$ (Zawadzki, 1973). Note that with an exponentially decaying

function, the integral of Eq. (4) can be analytically derived and is equal to $-\frac{1}{\ln(\rho_0)}$.

In order to generalize the methodology to different ACFs and for the comparison of observed and forecast field at all lead times, Eq. (4) was numerically integrated using the extended Simpson's rule (Press et al., 2007).

## 3.4 Online collection of rainfall statistics

Instead of analyzing the temporal distribution of the Lagrangian auto-correlation by integrating the data over space, we want to analyze its spatial distribution by integrating over time. More precisely, the summations of Eq. (2) need to be done over the number of radar images in the archive, not the number of pixels within a radar image. A joint evaluation of the summations at each pixel in a radar field is intractable as it would require loading the whole archive of rainfall fields into the computer memory to compute the correlations in a single pass. An efficient way to overcome this issue is to exploit rules for the online computation of the mean, the variance and the covariance (Knuth, 1998). The online estimation of the mean is obtained as follows:

$$\overline{x}_{t+1} = \overline{x}_t + \delta/N, \tag{5}$$

where $t$ is the iteration, $\overline{x}_{t+1}$ is the new mean, $\delta = x_{t+1} - \overline{x}_t$ is the residual contribution of the new sample $x_{t+1}$ to the old mean $\overline{x}_t$ and $N$ is the number of samples.

The online estimation of the variance is obtained similarly as follows:

$$q_{t+1} = q_t + \delta\left(x_{t+1} - \overline{x}_{t+1}\right), \tag{6}$$

where $q$ is the squared sum of the differences of $x$ from its mean and $\delta = x_{t+1} - \overline{x}_t$. The variance is obtained offline by dividing $q$ by the number of samples $N$.

The online computation of the AR(1) Lagrangian temporal auto-correlation is evaluated by keeping track of the sum of squared residuals:

$$s_{t+1} = s_t + \left(x_{t+1}^{adv} - \overline{x}_t^{adv}\right) \cdot \left(x_{t+1} - \overline{x}_{t+1}\right). \tag{7}$$

The Lagrangian auto-correlation is obtained offline as:

$$\rho\left(x, x^{adv}\right) = \frac{s}{N\sqrt{\mathrm{Var}(x) \cdot \mathrm{Var}\left(x^{adv}\right)}}. \tag{8}$$

The online computation of statistics using these rules gradually converges towards a stable value as the time progresses. In order to admit temporal fluctuations of the statistics and local smoothing, one can introduce a weight in the recursive equations similarly to the technique of recursive least squares computation.

The technical implementation of the online update of the field statistics is performed by keeping binary files containing the arrays of interim statistics. For each new radar field,

the old file is read, updated and rewritten with the new statistics. The statistics are only updated when the rainfall fraction exceeds 5 % over the radar composite and when the four radars are jointly operating. With this criterion we obtained 9578 valid rainfall fields, which roughly correspond to 1600 h of precipitation over the period spanning from February 2011 to October 2012.

## 3.5 Offline spectral slope estimation

A precipitation field that is scale-invariant (also referred to as *scaling*) typically exhibits a power spectrum of the form:

$$P(f) \propto f^{-\beta}, \tag{9}$$

where $f$ is the spatial frequency (km$^{-1}$) and $\beta$ is the scaling exponent (the slope of the power spectrum). The power law behavior of rainfall fields usually appears as a straight line on a graph of the logarithm of the power against the logarithm of the spatial frequency. The slope of the line measures the degree of scaling of the field and is equal to 0 for an unstructured white noise field. The scaling exponent of a 2-dimensional rainfall field is often greater than 2, which complicates its multifractal simulation (see, e.g., Schertzer and Lovejoy, 1987). One possibility to simulate stochastic rainfall fields to obtain $\beta > 2$ is to apply a power law filter to a field of white noise as briefly mentioned in Sect. 3.2.

Radar rainfall fields often deviate from the theoretical framework of perfect scale-invariance and typically show a scaling break at frequencies of 15–20 km$^{-1}$ (see, e.g., Gires et al., 2011; Seed et al., 2013). On the other hand, precipitation fields computed by NWP models have a break around 40–50 km (e.g., Gires et al., 2011). The scaling break is observed as an increase in the spectral slope at the smaller convective scales. This seems to have a physical origin and could be attributed to different scaling regimes of the large-scale stratiform rainfall and the smaller convective scales. However, recent analyses explain this phase transition with the presence of zeros in the field, which also affects the estimation of universal multifractal parameters (Gires et al., 2012). The presence of a scaling break requires using two spectral slopes $\beta_1$ and $\beta_2$ for the study and parameterization of the power spectrum. $\beta_1$ accounts for wavelengths that are larger than 15–40 km and $\beta_2$ for wavelengths that are lower than 15–40 km.

In this research we analyze the spatial distribution of the spectral slopes $\beta_1$ and $\beta_2$. A complete analysis would require visiting each radar pixel and performing a local spatial FFT decomposition in its neighborhood, which is very computationally demanding if one wants to repeat the analysis over a long period of time. Instead, the two spectral slopes are derived offline from the spatial distribution of the variance at each level of the cascade. This can be achieved by assuming scaling of the variance of the cascade levels (see Menabde et al., 1997). It consists of evaluating the average slope increments between successive levels of the cascade level standard

**Table 1.** Characteristics of the six flow cluster centers that are used to stratify the statistics. Wm and Ws refer to moderate and strong westerly flows respectively. The detailed average velocity maps can be found in Foresti and Seed (2014).

| Cluster label | 0-SE | 1-Wm | 2-N | 3-SW | 4-NW | 5-Ws |
|---|---|---|---|---|---|---|
| Average flow direction | Southeast | West | North | Southwest | Northwest | West |
| Average flow magnitude ($\mathrm{km\,h}^{-1}$) | 8.2 | 17.2 | 21.3 | 21.9 | 40.0 | 37.5 |

deviations:

$$H = -\frac{1}{K-1} \sum_{k=1}^{K-1} \frac{\log_{10}\left(\frac{\mathrm{sd}(X_{kij})}{\mathrm{sd}(X_{(k+1)ij})}\right)}{\log_{10}(q)}$$

$$\text{and } \beta = 2H + E, \tag{10}$$

where $\mathrm{sd}(X_{kij})$ is the standard deviation of cascade level $k$ at pixel $ij$ and $E = 2$ is the dimension of the space. $\beta_1$ is estimated using levels 1 to 3 (scales of 171, 81 and 38 km, $K = 3$ in Eq. 10) while $\beta_2$ using levels 3 to 6 (scales of 38, 18, 8 and 4 km; $K = 4$ in Eq. 10). A scaling break of 40 km instead of 20 km was chosen to obtain smoother fields of the spectral exponent $\beta_2$, which is consequently slightly underestimated.

Note that this approach is different than estimating power spectra on rainfall time series and analyzing the spatial distribution of the spectral exponents. The approach proposed in this paper should give insights into the spatial heterogeneity of the degree of spatial scaling of rainfall fields.

### 3.6 K-means stratification of optical flow fields

To analyze the dependence of rainfall statistics on flow regimes, the optical flow fields were stratified using the k-means clustering algorithm. The details on the preparation of the archive of optical flow fields and the clustering algorithm can be found in Foresti and Seed (2014).

Table 1 summarizes the statistics of the six cluster centers obtained after running the k-means algorithm on the archive of flow fields. The cluster centers mainly differ in terms of flow direction and magnitude, while the spatial variability of the velocity vectors within a field is only marginal (see Foresti and Seed, 2014). The number of clusters was empirically chosen to represent a sufficient number of flow regimes and to have enough samples per cluster to compute significant verification statistics. The cluster 0 is characterized by weak southeasterly winds, the cluster 1 by moderate westerly winds, the cluster 2 by moderate northerlies, the cluster 3 by moderate southwesterly winds, the cluster 4 by strong northwesterly winds, and the cluster 5 by strong westerlies. It is understood that winds refer to the apparent motion of radar images derived with optical flow and not to real wind fields.

The online update of the statistics (Sect. 3.4) is performed by keeping a set of six binary files containing the interim fields of the rainfall mean, variance, Lagrangian autocorrelations and number of samples. The files are read, updated and rewritten according to the cluster membership of a given field.

## 4 Results

Section 4.1 illustrates the scale-dependent geographical distribution of the precipitation lifetime without stratification into flow regimes. On the other hand, Sect. 4.2 shows the flow dependence of the dynamic and spatial scaling relationships by averaging the results over space. Finally, Sect. 4.3 analyzes the spatial distribution of the precipitation lifetime under different flow regimes to understand the effect of orographic forcing.

### 4.1 Geographical distribution of precipitation predictability and spatial scaling

Figure 3 illustrates the spatial distribution of the precipitation lifetime for the cascade levels 2, 3, 4, and 5 without stratification into flow regimes. Refer to Fig. 1 for geographical details. The level 0 is not presented since the FFT filter does not have a Gaussian shape (see Fig. 2). On the other hand, the level 1 is too influenced by the edge effects that propagate from the borders of the radar composite towards the interior regions. The levels 6 and 7 are too noisy and exhibit lifetimes that are below the temporal resolution of the radar composite (10 min).

The cascade level 2 (171-*81*-38 km, Fig. 3a) has lifetimes comprised between 5 and 10 h but still highlights the presence of some edge effects. The long lifetimes obtained may still be a consequence of assuming the diagnosed velocity field to be temporally stationary. An important part of the spatial variability at this scale is affected by the shape of the radar composite, and long lifetimes tend to be located in its central parts. The other cascade levels (Fig. 3b–d) are less affected by the edge effects, which remain limited to a small region close to the borders of the radar composite. All of them have the longest lifetimes over the flat regions surrounding the Yarrawonga radar. In this region the lifetimes are up to 3.5 h, and 70 and 30 min for the 81-*38*-18, 38-*18*-8 and 18-*8*-4 km scales, respectively. These long lifetimes can be explained by a higher Lagrangian predictability over flat continental areas (see for instance Germann et al., 2006). The lifetimes around the Macedon ranges are lower on their southeast flanks in the direction of the Melbourne radar compared with their northwest flanks. Despite being
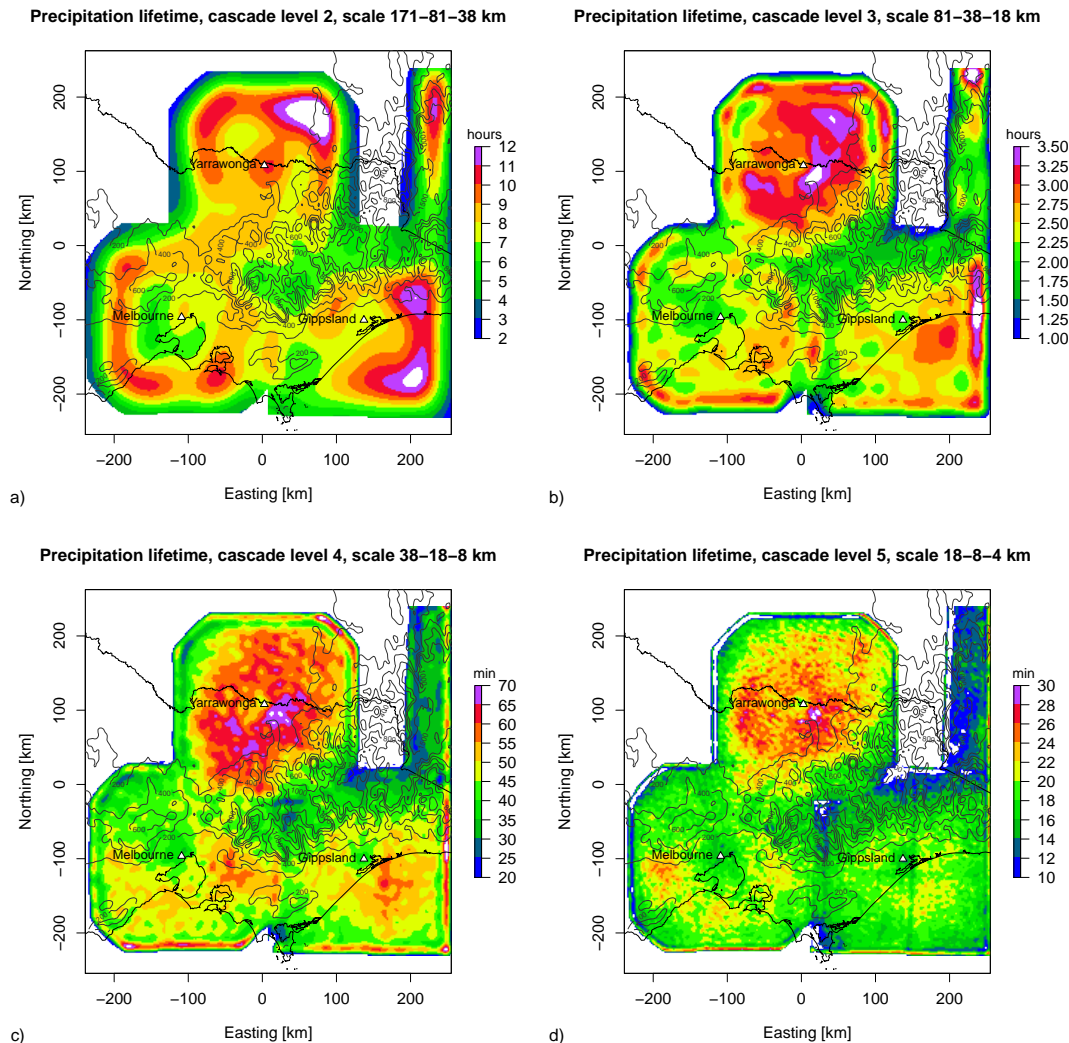
**Figure 3.** Spatial distribution of the precipitation lifetimes for the four middle cascade levels. **(a)** 171-*81*-38 km, **(b)** 81-*38*-18 km, **(c)** 38-*18*-8 and **(d)** 18-*8*-4 km. White tones are used for regions outside the radar domain or presenting values that exceed the range of the color scale.

less pronounced, this pattern was already observed by Foresti and Seed (2014) and is a consequence of the prevailing westerly flows, which cause systematic rainfall decay on the leeward side of the Macedon ranges and orographic enhancement on their windward side. This effect is also the origin of the long lifetimes observed on the Dandenong ranges as they are located upwind relative to the prevailing westerlies. The lifetimes surrounding the Gippsland radar tend to be longer over the ocean, which is particularly visible in Fig. 3b and c. Finally, the shorter lifetimes on the inner parts of the Victorian Alps are probably due to the reduced accuracy of the radar measurements (see Sect. 2). In particular, the blockage of radar beams, the rainfall attenuation and overshooting reduce the accuracy of the optical flow estimations, which consequently affects the lifetimes derived from the Lagrangian auto-correlation. In addition, it seems that there is a proportional effect between the precipitation lifetime and the cli-

matological precipitation amount: the lifetimes are generally lower in the places where the radar measures less precipitation (see e.g., Berenguer and Sempere-Torres, 2013).

Figure 4 illustrates the spatial distribution of the spectral slopes $\beta_1$ and $\beta_2$ derived from the standard deviation of the cascade levels (see Sect. 3.5). $\beta_1$ represents the degree of scaling above the 40 km scale; $\beta_2$ represents that below it and they therefore account respectively for the large-scale precipitation structures and convective features. Except for the regions close to the radar domain edges, $\beta_1$ is generally larger than 2 with the highest values in the range 2.2–2.4 and centered on the three radars. These are the regions where the spatial scaling of rainfall can be measured more efficiently and is the highest. At first sight, these inhomogeneities can only be explained by the shape of the radar composite and not by the presence of different atmospheric processes. However, the spectral slopes are higher on the southern slope of
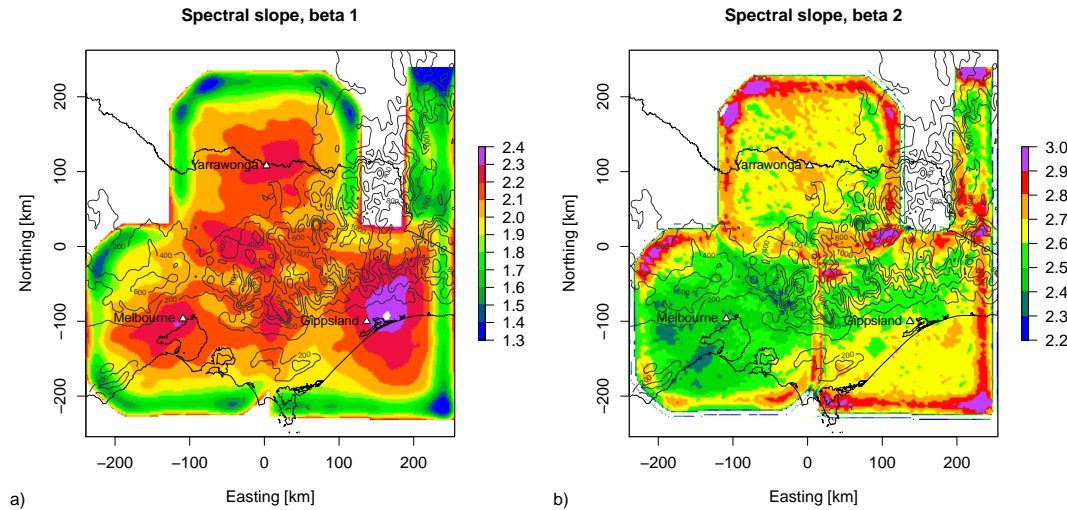
**Figure 4.** Spatial distribution of the spectral slopes **(a)** $\beta_1$ and **(b)** $\beta_2$ derived by assuming the scaling of the standard deviation of the cascade levels.

the Alps located northeast of the Gippsland radar. This depicts a region characterized by rainfall fields that are highly organized in space with convection embedded into stratiform rainfall, which is typical of orographic rainfall (see Fig. 6a). It would be interesting to perform a similar analysis using outputs from NWP models to eliminate the heterogeneities introduced by the inhomogeneous quality of radar measurements. As expected, the spectral slopes at the small scales ($\beta_2$, Fig. 4b) are systematically higher than the ones at the large scales ($\beta_1$, Fig. 4a), with values in the range 2.3–2.8. However, the spectral slope $\beta_2$ is lower in the surroundings of the Melbourne radar (S-band) compared with the other two (C-band). Both the C- and S-band radars have a 1° azimuth and 250 m range resolution (see for instance, Rennie, 2012). Notwithstanding the same resolution, the rainfall field exhibits more power in the last cascade level in the surroundings of the Melbourne radar, which can explain the lower spectral exponent $\beta_2$ (Fig. 4b). The patterns observed in Fig. 4b are hard to explain in terms of different precipitation regimes and seem to be more associated to the type of radar or data processing chain. Despite these differences, the spectral exponents $\beta_2$ tend to be lower upwind than upstream of the mountain ranges, in particular over the Yarra and Dandenong ranges, the southern slopes of the Alps between Avon and the Snowy River, and on the northern slopes of the Alps located southeast of the Yarrawonga radar. This depicts that strong convection is more likely to occur over flat regions than over complex orography, where it is less intense and often embedded into stratiform rainfall.

## 4.2 Flow dependence of the dynamic and spatial scaling relationships

Table 2 and Fig. 5 illustrate the dynamic scaling relationship between the spatial scale and the precipitation lifetime

for each flow regime. As already explained in Sect. 3.1, only the cascade levels 1 to 6 are shown. The values are obtained by spatial averaging of the lifetimes within the radar composite. Figure 5 demonstrates the presence of dynamic scaling, which is observed as a clear power law relationship between the spatial scale of precipitation features and its estimated lifetime (Venugopal at al., 1999; Seed, 2003). It is worth mentioning that Venugopal at al. (1999) employed another statistical quantity to account for the temporal evolution of rainfall and the obtained dynamic scaling exponents cannot be directly compared. The figure also shows significant variability of the lifetimes as a function of flow regime. The clusters NW and Ws are characterized by the shortest lifetimes, and the cluster SE by the longest. These differences are in part due to the type of rainfall, which is more convective under northerly than southerly flows. In fact, the organized convective activity mostly occurs when the warm continental northwesterly flows meet the colder maritime air. On the other hand, it is not clear whether the faster translational speed of convective rain relative to stratiform rain affects the estimation of the predictability by Lagrangian persistence. These lifetime estimations are a bit higher than the original ones of Seed (2003), who used a single motion vector to advect the radar rainfall field. Similar issues were encountered by Pegram and Clothier (2001b) because of using a single displacement vector and as a consequence of the high level of noise at the pixel scale. This demonstrates the added value of the optical flow algorithm of Bowler et al. (2004), which better defines the differential motion within a rainfall field, as well as the analysis of the Lagrangian predictability at larger scales using the Fourier-based scale decomposition. The estimations also compare well with the results of Germann et al. (2006), who reported lifetimes of 0.1–0.2 h on the 4–8 km scale using a wavelet decomposition of the rainfall

**Table 2.** Precipitation lifetimes for each spatial scale and flow regime averaged over the radar composite. Levels 0–3 are expressed in hours and 4–7 in minutes. The power law extrapolation of lifetimes for smaller spatial scales is given in seconds. Ext.: estimation of the lifetimes at smaller spatial scales by extrapolating the power law. The extrapolation uses the original non-integer scales for increased precision. The scales on which the Gaussian filters are centered are marked in italic.

| Level | Spatial scales [km] | Cluster 0 SE | Cluster 1 Ww | Cluster 2 N | Cluster 3 SW | Cluster 4 NW | Cluster 5 Ws | Weighted average |
|---|---|---|---|---|---|---|---|---|
| 0 | 768-362 | 27.7 | 25.9 | 28.5 | 25.4 | 19.3 | 20.2 | 24.1 h |
| 1 | 362-*171*-81 | 24.9 | 21.4 | 18.6 | 22.5 | 14.6 | 16.0 | 19.2 h |
| 2 | 171-*81*-38 | 13.7 | 8.9 | 7.9 | 9.6 | 5.6 | 6.3 | 8.3 h |
| 3 | 81-*38*-18 | 4.6 | 2.8 | 2.5 | 3.1 | 1.8 | 2.0 | 2.7 h |
| 4 | 38-*18*-8 | 79.6 | 53.9 | 48.9 | 56.3 | 34.4 | 38.7 | 49.8 min |
| 5 | 18-*8*-4 | 26.4 | 21.0 | 19.2 | 20.5 | 14.0 | 15.7 | 19.0 min |
| 6 | 8-*4*-2 | 10.5 | 9.2 | 8.5 | 8.7 | 6.4 | 7.31 | 8.3 min |
| 7 | 4-*2* | 5.9 | 5.0 | 5.2 | 5.4 | 4.7 | 5.0 | 5.1 min |
| ext. 8 | 1.89-*0.89*-0.42 | 78 | 65 | 62 | 58 | 43 | 49 | 58.0 s |
| ext. 9 | 0.89-*0.42*-0.20 | 28 | 23 | 23 | 16 | 16 | 18 | 20.5 s |
| ext. 10 | 0.42-*0.20*-0.09 | 10 | 8 | 8 | 5 | 6 | 7 | 7.4 s |
| | No. of fields | 1095 | 2390 | 1449 | 1112 | 2058 | 1474 | 9578 |

field over the continental United States. In Fig. 5 the 4–8 km scales roughly correspond to the 8-*4*-2 km and 18-*8*-4 scales, which exhibit lifetimes of 0.1–0.4 h.

To obtain an order of magnitude for the predictability at smaller spatial scales, power law relationships were fitted using the method of least squares per each flow cluster. The extrapolation of the fitted power laws towards smaller spatial scales could give an idea of the minimum temporal resolution that is required to reliably measure the Lagrangian auto-correlation of precipitation, which is very important for stochastic precipitation nowcasting at urban scales (e.g., Goormans and Willems, 2013; Ruzanski and Chandrasekar, 2012). The bottom of Table 2 shows the results of such extrapolations for scales of 1.89-*0.89*-0.42, 0.89-*0.42*-0.20 and 0.42-*0.20*-0.09 km. Because of working on a logarithmic scale such estimations are quite uncertain and to a certain degree pessimistic, in particular because the dynamic scaling relationship does not perfectly follow a power law. The imperfect dynamic scaling could also be due to using the lifetime instead of the temporal rainfall changes as a measure for the rainfall evolution (see Venugopal et al., 1999). It must also be considered that the optical flow is representative of the scales measured by the C- and S-band radars and cannot capture the motion at smaller scales. From this simple extrapolation, the kilometric scale (1.89-*0.89*-0.42 km) only displays a predictability of 40–80 s. It would be interesting to study whether the temporal resolution of X-band radars is sufficient to reliably measure the Lagrangian auto-correlation of the very small scale precipitation features. Using such high resolution data will also pose the computational challenge of generating the nowcasts before the predictability limits have been exceeded to avoid the forecasts becoming obsolete. Ruzanski and Chandrasekar (2012) re-

ported a predictability of 20 min using data from a network of X-band radars and the CASA nowcasting system. The scale dependence was analyzed by upscaling the forecasts and the values are not directly comparable to the ones obtained by scale separation within STEPS. At these temporal scales, the quality of the nowcasts is still strongly affected by the accuracy of the input radar observations. Therefore, it becomes necessary to complement nowcasting systems with heuristic models of the radar measurement uncertainty, for example to account for stochastic sampling errors (Jordan et al., 2003).

Table 3 illustrates the spectral slopes $\beta_1$ and $\beta_2$ of the spatial power spectrum stratified by flow regime. $\beta_1$ typically oscillates around the dimension of the field with the smallest values occurring under the flows SE–SW (1.88–1.91) and the largest under the flows Wm, N, and NW (2.01–2.03). The values are slightly smaller than the ones found in the literature (e.g., Seed et al., 2013), which is explained again by the presence of edge effects that locally reduce the spectral exponents (see Fig. 4a). This may have consequences on the power law filtering performed by STEPS to generate the noise cascade needed to update the hierarchy of autoregressive processes. In fact, the filtering uses the spatial power spectrum of rainfall as target distribution, which does not account for the spatial heterogeneities within the forecast domain.

The values of $\beta_2$ are significantly higher and oscillate between 2.45 and 2.8. The clusters NW and Ws have the highest $\beta_2$ (2.68–2.79), which can be attributed to a higher convective activity occurring under these flow conditions. The cluster NW also has a high $\beta_1$ and is the one having the most organized rainfall structures from the large down to the small convective scales.

**Table 3.** Average spatial spectral exponents stratified by flow regime. The standard deviation over space is given in brackets.

| Cluster label | 0-SE | 1-Wm | 2-N | 3-SW | 4-NW | 5-Ws |
|---|---|---|---|---|---|---|
| $\beta_1$ | 1.88 (0.29) | 2.02 (0.24) | 2.03 (0.25) | 1.91 (0.22) | 2.03 (0.24) | 1.96 (0.23) |
| $\beta_2$ | 2.46 (0.20) | 2.55 (0.15) | 2.61 (0.18) | 2.46 (0.19) | 2.79 (0.16) | 2.68 (0.18) |
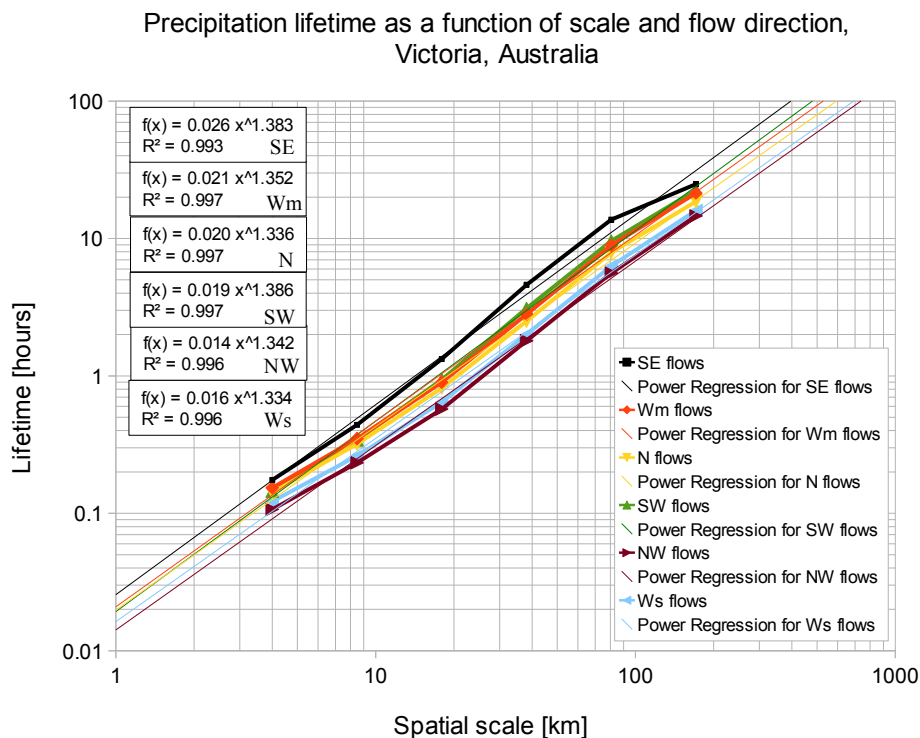


**Figure 5.** Dynamic scaling relationship between the spatial scale and precipitation lifetime stratified by flow regime. The equations of the power law fits are shown in the upper left corner.

## 4.3 Effect of orography on the predictability of precipitation

According to the results of Berenguer and Sempere-Torres (2013), the regions of long predictability seem to be correlated with the regions with the highest rainfall accumulations. In fact, the regions that are often affected by organized large-scale precipitation systems are more likely to exhibit higher predictability than the ones with infrequent isolated convection. It is therefore important to analyze the climatology of precipitation to study the spatial distribution of its predictability.

Figure 6 shows the conditional mean 10 min rainfall accumulations stratified by flow regime. It clearly illustrates the flow dependence of the spatial distribution of precipitation, which is mostly located on the windward side of mountain ranges. Most of the precipitation occurring under southeasterly flows is located along the upwind side of the Victorian Alps in a region going from Avon to the Snowy River (Fig. 6a). The spatial distribution of rainfall under

moderate westerly flows presents maxima on the Dandenong and Macedon ranges, but also on the northern side of the Alps around Mount Buffalo (Fig. 6b). The enhancement on the Northwest flank of the Alps is much more pronounced with northerly and northwesterly flows, which approach the mountain range more perpendicularly (Fig. 6c and e, respectively). Southwesterly flows lead to high accumulations on the Yarra and Dandenong ranges as well as the southern side of the Alps around the Gippsland radar (Fig. 6d). It is interesting to note that northwesterly flows also give high accumulations on the leeside of the Alps (Fig. 6e), which could be caused by the lower air stability of these conditions (refer to Foresti and Seed, 2014, for a more detailed interpretation). Finally, strong westerly flows lead again to high accumulations on the Dandenong and Macedon ranges, but also on the West of the Gippsland radar (Fig. 6f). A clear rainfall shadow effect on the leeside of the Macedon ranges is noticed for the clusters Wm, NW, and Ws.

Figure 7 shows the spatial distribution of the precipitation lifetime at the convective scale (38-18-8 km) stratified
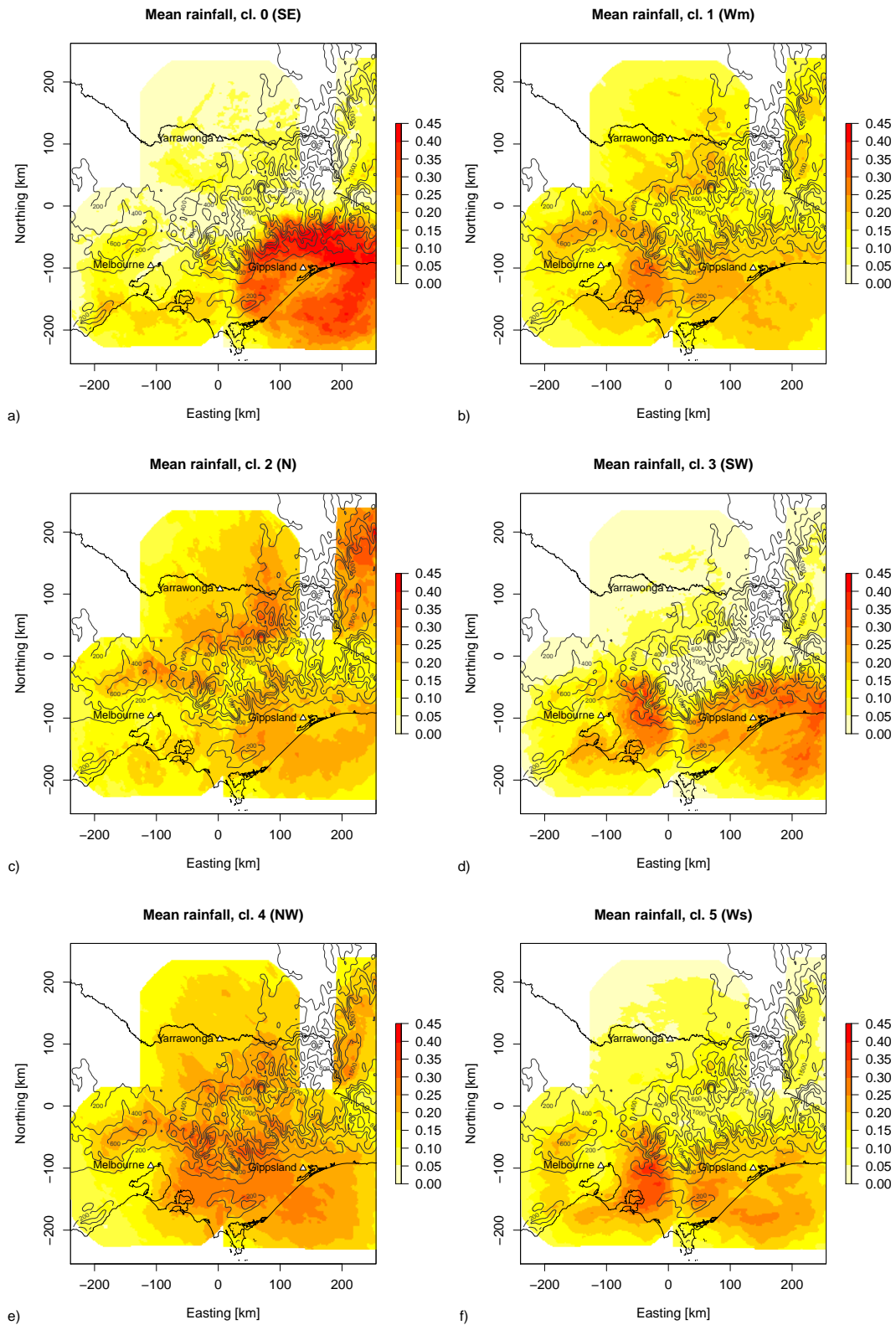
**Figure 6.** Conditional mean 10 min rainfall accumulations for flow regimes **(a)** SE, **(b)** Wm, **(c)** N, **(d)** SW, **(e)** NW and **(f)** Ws.

by flow cluster. Despite some variability arising from periodic features of the Fourier transform, it is possible to notice that lifetimes are higher on the upwind side and lower on the downwind side of terrain features. An illustrative example can be observed under southwesterly conditions (Fig. 7d). The lifetime of precipitation upstream of the Dandenong ranges is about 20–40 min; it increases to 50–70 min on the upwind side and falls again to 20–30 min when moving into the Alps. Similar patterns can be observed under the flow regime Ws (Fig. 7f). On the other hand, under NW flows short lifetimes are located on the leeward side of the Macedon ranges (Fig. 7e). Note that with reversed flow conditions (SE, Fig. 7a), this region exhibits lifetimes of 80–100 min and the shortest ones are located on top of the Macedon ranges with values oscillating between 40 and 80 min. The region located South and Southeast of the Yarrawonga radar is also interesting to analyze in particular for the clusters N and NW. In fact, the location of the longest lifetimes upstream of the Alps is different depending on flow direction (Fig. 7c and e). The plains surrounding the Yarrawonga radar also show very long lifetimes under flow conditions SE, Wm, and SW. However, this effect could be an artefact of the low rainfall accumulations over these regions (see Fig. 6a, b, and d).

These findings corroborate the results of Harris et al. (1996), who demonstrated that the precipitation intermittency is higher upstream compared with the top of the mountain ridge, with intermediate values on the upwind flank. From Fig. 7 it seems that the decreased intermittency of rainfall upwind of orographic features has a positive impact on its predictability by Lagrangian persistence. It is worth mentioning that leeside precipitation enhancement is also possible due to leeside flow convergence, flow perturbations by mountain gravity waves, or the presence of cold air pools that force the unstable air to rise. Such processes are not very frequent and would require stratifying the statistics using more complex criteria based on moist air stability indices among others.

The relationship between the precipitation lifetime and orography is less pronounced than that of nowcast biases presented in Foresti and Seed (2014). This is mostly due to the increased difficulty in computing higher order statistics, which require many more samples than a simple linear or multiplicative bias. Also, the cascade decomposition framework still needs some improvements to reduce the edge effects and to better interpret the intricate statistical dependencies between consecutive cascade levels (see Seed et al., 2013).

## 5 Conclusions

The geographical distribution of the scale-dependent predictability of precipitation by Lagrangian extrapolation of radar images was analyzed under different flow regimes in connection with the presence of orographic features. Data from the Victorian radar composite, Australia, a $500 \times 500\,\mathrm{km}^2$ domain covering the period from February 2011 to October 2012, were used for the analyses. The scale dependence of the predictability of precipitation was considered by decomposing the radar rainfall field into a multiplicative cascade using an FFT (Bowler et al., 2006). The lifetime of precipitation features was found to be a power law function of the scale of the features and to depend on flow direction, which confirms the presence of dynamic scaling (Venugopal et al., 1999; Mandapaka et al., 2009). The precipitation lifetime was found to be up to a factor of 2 higher on the upwind compared with the downwind slopes of orographic features and to be strongly flow-dependent. The degree of spatial scaling of the rainfall field was also shown to be spatially inhomogeneous. These spatial heterogeneities due to orographic forcing can be exploited to locally adapt the space–time stochastic simulation of precipitation, which is needed for very short-term forecasting (e.g., Seed et al., 2013), generating radar ensembles (Germann et al., 2009), design storm studies (e.g., Paschalis et al., 2013), and precipitation downscaling (e.g., Pathirana and Herath, 2002).

The study raised several methodological questions, in particular because the quality of radar data is much more homogeneous over time than over space. This has to be accounted for when interpreting the maps of the predictability of precipitation. Some patterns could be simply due to the geographical biases that affect the radar measurements, for example due to beam blockage, signal attenuation, or increasing sampling volume with range. Nevertheless, in the regions close to the radar, it was possible to detect a clear signal in the distribution of the precipitation lifetime, which was attributed to orographic forcing.

The predictability estimates presented in this paper are affected by other sources of uncertainty. The first is related to the assumption of the temporal stationarity of the diagnosed velocity field, which leads to over-optimistic estimates of the precipitation lifetimes, especially at the large scales. The second arises from the uncertainty in the estimation of the velocity field with optical flow. In fact, precipitation fields often show differential motion at different spatial scales. An illustrative example occurs when stationary orographic rainfall contains fast moving cellular convection (e.g., Foresti et al., 2013). Better estimates of the Lagrangian predictability would require the optical flow to be estimated on each spatial scale separately.

Finally, it is not yet clear whether the spatial variability of precipitation lifetime is more significant than its temporal variability, and how to account for both aspects in the real-time nowcasting of precipitation using stochastic simulation approaches such as STEPS. The natural solution would be to allow the predictability to vary through time in a first stage and to gradually add some spatial heterogeneity when more and more radar data are collected. This goal could be achieved by exploiting the online computation of statistics,
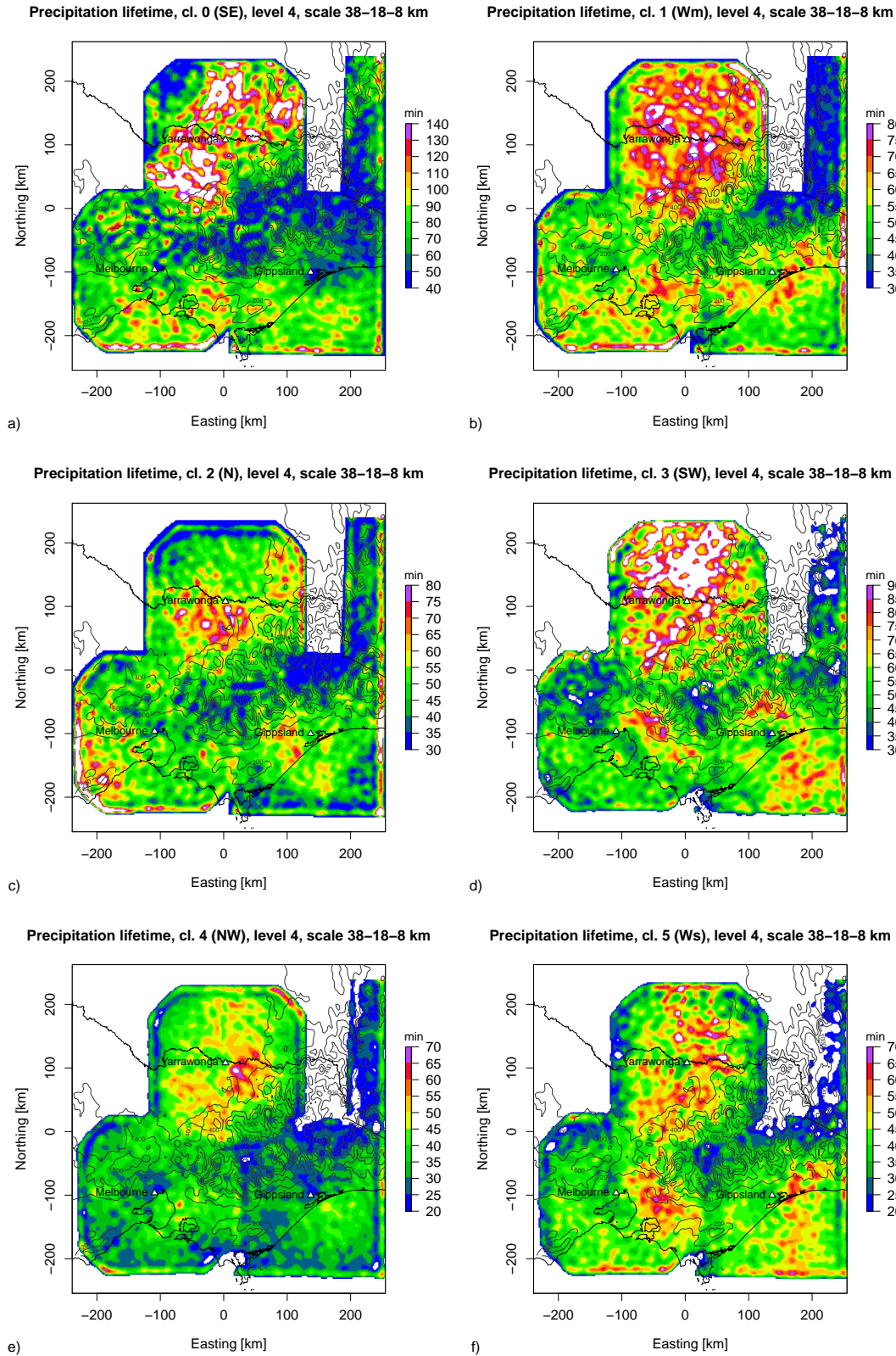
**Figure 7.** Spatial distribution of the precipitation lifetimes at cascade level 4 (38-*18*-8 km) stratified by flow regime. **(a)** SE, **(b)** Wm, **(c)** N, **(d)** SW, **(e)** NW, **(f)** Ws.

which would enable the nowcasting system to learn about the spatial distribution of predictability as more and more radar data are collected and analyzed.

Edited by: H. Leijnse

# References

Badas, M. G., Deidda, R., and Piga, E.: Modulation of homogeneous space-time rainfall cascades to account for orographic influences, Nat. Hazards Earth Syst. Sci., 6, 427–437, doi:10.5194/nhess-6-427-2006, 2006.

Berenguer, M., and Sempere-Torres, D.: Radar-based rainfall nowcasting at European scale: long-term evaluation and performance assessment, Proc. of the 36th AMS Conf. on Radar Meteorology, Breckenridge, Colorado, USA, 2013.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G. S.: SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, J. Hydrol., 404, 226–240, 2011.

Bousquet, O., Lin, C. A., and Zawadzki, I.: Analysis of scale dependence of quantitative precipitation forecast verification: a case-study over the Mackenzie river basin, Q. J. Roy. Meteorol. Soc., 132, 2107–2125, 2006.

Bowler, N. E. H., Pierce, C. E., and Seed, A. W.: Development of a precipitation nowcasting algorithm based upon optical flow techniques, J. Hydrol., 288, 74–91, 2004.

Bowler, N. E. H., Pierce, C. E., and Seed, A. W.: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Q. J. Roy. Meteorol. Soc., 132, 2127–2155, 2006.

Casati, B., Ross, G., and Stephenson, D. B.: A new intensity-scale approach for the verification of spatial precipitation forecasts, Meteorol. Appl., 11, 141–154, 2004.

Chumchean, S., Sharma, A., and Seed, A.: An integrated approach to error correction for real-time radar-rainfall estimation, J. Atmos. Ocean. Tech., 23, 67–79, 2006a.

Chumchean, S., Seed, A., and Sharma, A.: Correcting of real-time radar rainfall bias using a Kalman filtering approach, J. Hydrol., 317, 123–137, 2006b.

Chumchean, S., Seed, A., and Sharma, A.: An operational approach for classifying storms in real-time radar rainfall estimation, J. Hydrol., 363, 1–17, 2008.

Foresti, L. and Seed, A.: On the spatial distribution of rainfall nowcasting errors due to orographic forcing. Meteorol. Appl., doi:10.1002/met.1440, in press, 2014.

Foresti, L., Kanevski, M., and Pozdnoukhov, A.: Kernel-based mapping of orographic rainfall enhancement in the Swiss Alps as detected by weather radar, IEEE T. Geosci. Remote, 50, 2954–2967, 2012.

Foresti, L., Panziera, L., Mandapaka, P. V., Germann, U., and Seed, A.: Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall, Meteoforol. Appl., doi:10.1002/met.1416, in press, 2013.

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

Germann, U., Zawadzki, I., and Turner, B.: Scale-dependence of the predictability of precipitation from continental radar images, Part IV: Limits to Prediction, J. Atmos. Sci., 63, 2092–2108, 2006.

Germann, U., Berenguer, M., Sempere-Torres, D., and Zappa, M.: REAL – Ensemble radar precipitation estimation for hydrology in a mountainous region, Q. J. Roy. Meteorol. Soc., 135, 445–456, 2009.

Gires, A., Tchiguirinskaia, I., Schertzer, D., and Lovejoy, S.: Multifractal and spatio-temporal analysis of the rainfall output of the Meso-NH model and radar data, Hydrolog. Sci. J., 55, 380–396, 2011.

Gires, A., Tchiguirinskaia, I., Schertzer, D., and Lovejoy, S.: Influence of the zero-rainfall on the assessment of the multifractal parameters, Adv. Water Resour., 45, 13–25, 2012.

Goormans, T. and Willems, P.: Using local weather radar data for sewer system modeling: case study in Flanders, Belgium, J. Hydrol. Eng., 18, 269–278, 2013.

Grecu, M. and Krajewski, W. F.: A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting, J. Hydrol., 239, 69–84, 2000.

Harris, D., Menabde, M., Seed, A., and Austin, G.: Multifractal characterization of rain fields with a strong orographic influence, J. Geophys. Res., 101, 26405–26414, 1996.

Huuskonen, A., Saltikoff, E., and Holleman, I.: The operational radar network in Europe, B. Am. Meteorol. Soc., 95, 897–907, doi:10.1175/BAMS-D-12-00216.1, 2014.

Jordan, P., Seed, A. W., and Weinnman, P. E.: A stochastic model of radar measurement errors in rainfall accumulations at catchment scale, J. Hydrometeorol., 4, 841–855, 2003.

Knuth, D. E.: The Art of Computer Programming, in: volume 2: Seminumerical Algorithms, 3rd Edn, Addison-Wesley, Boston, 1998.

Lorenz, E. N.: Predictability of flow which possesses many scales of motion, Tellus, 21, 289–307, 1969.

Mandapaka, P. V., Lewandowski, P., Eichinger, W. E., and Krajewski, W. F.: Multiscaling analysis of high resolution space-time lidar-rainfall, Nonlin. Processes Geophys., 16, 579–586, doi:10.5194/npg-16-579-2009, 2009.

Marsan, D., Schertzer, D., and Lovejoy, S.: Causal space-time multifractal processes: Predictability and forecasting of rain fields, J. Geophys. Res., 101, 26333–26346, 1996.

Menabde, M., Harris, D., Seed, A., Austin, G., and Stow, D.: Multiscaling properties of rainfall and bounded random cascades, Water Resour. Res., 33, 2823–2830, 1997.

Panziera, L. and Germann, U.: The relation between airflow and orographic precipitation on the southern side of the Alps as revealed by weather radar, Q. J. Roy. Meteorol. Soc., 136, 222–238, 2010.

Paschalis, A., Molnar, P., Fatichi, S., and Burlando, P.: A stochastic model for high-resolution space-time precipitation simulation, Water Resour. Res., 49, 8400–8417, doi:10.1002/2013WR014437, 2013.

Pathirana, A. and Herath, S.: Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity, Hydrol. Earth Syst. Sci., 6, 695–708, doi:10.5194/hess-6-695-2002, 2002.

Pegram, G. G. S. and Clothier, A. N.: High resolution space-time modelling of rainfall: the "String of Beads" model, J. Hydrol., 241, 26–41, 2001a.

Pegram, G. G. S. and Clothier, A. N.: Downscaling rainfields in space and time, using the String of Beads model in time series mode, Hydrol. Earth Syst. Sci., 5, 175–186, doi:10.5194/hess-5-175-2001, 2001b.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes: The Art of Scientific Computing, 3rd Edn., Cambridge University Press, 2007.

Radhakrishna, B., Zawadzki, I., and Fabry, F.: Predictability of precipitation from continental radar images, Part V: growth and decay, J. Atmos. Sci., 69, 3336–3349, doi:10.1175/JAS-D-12-029.1, 2012.

Rennie, S. J.: Doppler weather radar in Australia, CAWCR technical report, no. 055, Centre for Australian Weather and Climate Research, Melbourne, Australia, 2012.

Ruzanski, E. and Chandrasekar, V.: An investigation of the short-term predictability of precipitation using high-resolution composite radar observations, J. Appl. Meteorol., 51, 912–925, 2012.

Schertzer, D. and Lovejoy, S.: Physical modelling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res., 92, 9696–9714, 1987.

Seed, A.: A dynamic and spatial scaling approach to advection forecasting, J. Appl. Meteorol., 42, 381–388, 2003.

Seed, A. and Pegram, G. G. S.: Using Kriging to infill gaps in radar data due to ground clutter in real-time. Proceedings of the 5th Int. Symp. on Hydrological Applications of Weather Radar, Kyoto, Japan, 73–78, 2001.

Seed, A., Duthie, E., and Chumchean, S.: Rainfields: the Australian Bureau of Meteorology system for quantitative precipitation estimation, Proc. of the 33rd Conf. on Radar Meteorology, Cairns, Australia, 2007.

Seed, A. W., Pierce, C. E., and Norman, K.: Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme, Water Resour. Res., 49, 6624–6641, 2013.

Sinclair, S. and Pegram, G. G. S.: Empirical Mode Decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting, Hydrol. Earth Syst. Sci., 9, 127–137, doi:10.5194/hess-9-127-2005, 2005.

Surcel, M., Zawadzki, I., and Yau, M. K.: On the filtering properties of ensemble averaging for storm-scale precipitation forecasts, Mon. Weather Rev., 142, 1093–1105, 2014.

Turner, B. J., Zawadzki, I., and Germann, U.: Predictability of precipitation from continental radar images, Part III: operational nowcasting implementation (MAPLE), J. Appl. Meteorol., 43, 231–248, 2004.

Venugopal, V., Foufoula-Georgiou, E., and Sapozhnikov, V.: Evidence of dynamic scaling in space-time rainfall, J. Geophys. Res., 104, 31599–31610, 1999.

Zawadzki, I. I.: Statistical properties of precipitation patterns, J. Appl. Meteorol., 12, 459–472, 1973.

Zawadzki, I., Morneau, J., and Laprise, R.: Predictability of precipitation patterns – an operational approach, J. Appl. Meteorol., 33, 1562–1571, 1994.

Hydrology and
Earth System
Sciences

# Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications

**L.-P. Wang**[1], **S. Ochoa-Rodríguez**[2], **C. Onof**[2], **and P. Willems**[1]

[1]Hydraulics Laboratory, Katholieke Universiteit Leuven, 3001 Heverlee (Leuven), Belgium
[2]Department of Civil and Environmental Engineering, Imperial College London, London, SW7 2AZ, UK

*Correspondence to:* L.-P. Wang (lipen.wang@bwk.kuleuven.be)

**Abstract.** Gauge-based radar rainfall adjustment techniques have been widely used to improve the applicability of radar rainfall estimates to large-scale hydrological modelling. However, their use for urban hydrological applications is limited as they were mostly developed based upon Gaussian approximations and therefore tend to smooth off so-called "singularities" (features of a non-Gaussian field) that can be observed in the fine-scale rainfall structure. Overlooking the singularities could be critical, given that their distribution is highly consistent with that of local extreme magnitudes. This deficiency may cause large errors in the subsequent urban hydrological modelling. To address this limitation and improve the applicability of adjustment techniques at urban scales, a method is proposed herein which incorporates a local singularity analysis into existing adjustment techniques and allows the preservation of the singularity structures throughout the adjustment process. In this paper the proposed singularity analysis is incorporated into the Bayesian merging technique and the performance of the resulting singularity-sensitive method is compared with that of the original Bayesian (non singularity-sensitive) technique and the commonly used mean field bias adjustment. This test is conducted using as case study four storm events observed in the Portobello catchment ($53\,\mathrm{km}^2$) (Edinburgh, UK) during 2011 and for which radar estimates, dense rain gauge and sewer flow records, as well as a recently calibrated urban drainage model were available. The results suggest that, in general, the proposed singularity-sensitive method can effectively preserve the non-normality in local rainfall structure, while retaining the ability of the original adjustment techniques to generate nearly unbiased estimates. Moreover, the ability of the singularity-sensitive technique to preserve the non-normality in rainfall estimates often leads to better reproduction of the urban drainage system's dynamics, particularly of peak runoff flows.

## 1 Introduction

Traditionally, urban hydrological applications have relied mainly upon rain gauge data as input. While rain gauges generally provide accurate point rainfall estimates near the ground surface, they cannot properly capture the spatial variability of rainfall, which has a significant impact on the urban hydrological system and thus on the modelling of urban runoff (Gires et al., 2012; Schellart et al., 2012). Thanks to the development of radar technology, weather radar data have been playing an increasingly important role in urban hydrology (Krämer et al., 2007; Liguori et al., 2011). Radars can survey large areas and better capture the spatial variability of the rainfall, thus improving the short-term predictability of rainfall and flooding. However, the accuracy of radar measurements is in general insufficient, particularly in the case of extreme rainfall magnitudes (Einfalt et al., 2004, 2005). This is due to the fact that, instead of being a direct measurement, radar rainfall intensity is derived indirectly from measured radar reflectivity. As a result, both radar reflectivity measurements and the reflectivity–intensity conversion process are subject to multiple sources of error.

First, errors in radar reflectivity measurements may arise from blockage of the radar beam, attenuation, ground clutter, anomalous propagation of the signal, among other sources (Collier, 1996; Einfalt et al., 2004; Harrison et al., 2000). Although the radar reflectivity measurements undergo a number

of corrections before they are converted into rainfall intensity, it is virtually impossible to have error-free reflectivity measurements. Second, the conversion between radar reflectivity ($Z$) and rainfall rate ($R$) uses the so-called $Z$–$R$ relationship, $Z = a\,R^b$ (Marshall and Palmer, 1948), where $a$ and $b$ are variables generally deduced by physical approximation or empirical calibration. This can be theoretically linked to the rain drop size distribution, which varies for different rainfall types. Operationally, a number of "static" $Z$–$R$ relations are usually derived to generate radar rainfall rates for different rainfall types (e.g. stratiform, convective and tropical storms), and the associated $a$ and $b$ variables are calibrated based upon long-term comparisons (Collier, 1986; Krajewski and Smith, 2002). However, in reality, it is almost impossible to classify a single storm purely under a specific rainfall type. Consequently, it is not entirely appropriate to use a static $Z$–$R$ relation to derive rainfall intensities, even for a single storm event. This has been confirmed by several studies, which indicate that rain drop size distribution is a highly dynamic process and may significantly or suddenly change within a storm event (Smith et al., 2009; Ulbrich, 1983). Because of this, the $Z$–$R$ derived rainfall intensity cannot effectively reflect the short-term dynamics of true rainfall intensities and may statistically compromise with intermediate rainfall intensities.

In order to overcome these drawbacks of radar rainfall estimates while preserving their spatial description of rainfall fields, it is possible to dynamically adjust them using rain gauge measurements. Many studies on this subject have been carried out over the last few years, though most of them focus on hydrological applications at large scales (Anagnostou and Krajewski, 1999; Fulton et al., 1998; Germann et al., 2006; Goudenhoofdt and Delobbe, 2009; Harrison et al., 2000, 2009; Seo and Smith, 1991; Thorndahl et al., 2014). Only few studies have examined the applicability of these adjustment techniques to urban-scale hydrological applications and concluded that they can effectively reduce rainfall bias, thus leading to improvements in the reproduction of hydrological outputs (Smith et al., 2007; Vieux and Bedient, 2004; Villarini et al., 2010; Wang et al., 2013). However, despite the improvements achieved with the current adjustment techniques, underestimation of storm peaks can still be seen after adjustment and this is particularly evident in the case of small drainage areas (such as those of urban catchments) and extreme rainfall magnitudes (Wang et al., 2012; Ochoa-Rodríguez et al., 2013). This may be due to the fact that the underlying adjustment techniques, mainly based upon Gaussian (first and/or second statistical moments) approximations (Goudenhoofdt and Delobbe, 2009; Krajewski, 1987; Todini, 2001), cannot properly cope with the so-called "singularities" (which imply non-normality and often correspond to local extreme magnitudes) observed at small scales (Schertzer et al., 2013; Tchiguirinskaia et al., 2012). In fact, it is often the case that the radar image captures the spatial structure of striking local extremes (albeit the actual rainfall

depth/intensity may be inaccurate), but these structures are lost or smoothened throughout the merging process. This deficiency may cause large errors in the subsequent urban hydrological modelling.

To address this limitation and improve the applicability of adjustment techniques at urban scales, a method is proposed herein which incorporates a local singularity (identification) analysis (Cheng et al., 1994; Schertzer and Lovejoy, 1987; Wang et al., 2012) into existing adjustment techniques and enables the preservation of the singularity structures throughout the adjustment process. The singularity-sensitive method is particularly intended to improve geostatistical-based merging techniques (e.g. co-kriging, Krajewski, 1987; Bayesian merging, Todini, 2001; kriging with external drift, Wackernagel, 2003 and conditional merging, Sinclair and Pegram, 2005), which seek to represent the spatial covariance structure of the rainfall field or its errors by making use of the semi-variogram (Goudenhoofdt and Delobbe, 2009). Being a second-order tool, the semi-variogram cannot adequately capture higher-order features of the rainfall field, thus causing these to be lost in the merging process.

The proposed singularity-sensitive method was initially developed and preliminarily tested in the reconstruction of a storm event which led to reported flooding in the Maida Vale area, Central London, in June 2009 (Wang and Onof, 2013; Wang et al., 2014). The radar rainfall product for this event showed strong and localised singularity structures, but the accuracy of the actual estimates was poor. A dynamic gauge-based adjustment was conducted using the Bayesian data merging method (Todini, 2001), which in previous studies had been shown to outperform other adjustment methods (Mazzetti and Todini, 2004; Wang et al., 2013). Nonetheless, for this particular event records from only a few rain gauge sites were available and these were located away from the area of interest and at points where less intense radar rainfall was observed. Under these circumstances, the aforementioned shortcomings associated with existing adjustment techniques became evident. The Bayesian data merging method proved inadequate as it smoothed out the singularity structures, which had the effect of considerably reducing the peak rainfall intensities. It was then that the singularity-sensitive method was devised and effectively incorporated into the Bayesian merging technique. The resulting singularity-sensitive Bayesian merging method led to rainfall fields which better preserved the spatial structure as captured by the radar and better reproduced peak rainfall intensities.

In the present paper the formulation of the proposed singularity-sensitive method is explained in detail and new numerical strategies aimed at improving the use of singularity information are introduced. Moreover, the method is further tested using as case study four storm events observed in the Portobello catchment ($53\,\mathrm{km}^2$) (Edinburgh, UK) during 2011 and for which radar estimates, a spatially dense net-
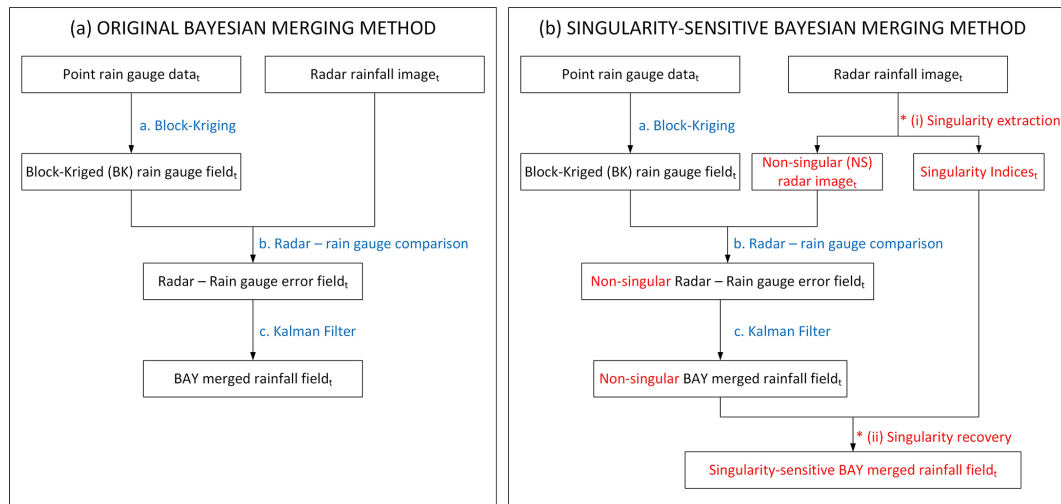
**Figure 1.** Schematics of **(a)** original Bayesian merging method (adapted from Fig. 2 in Mazzetti (2012)) and **(b)** singularity-sensitive Bayesian merging method.

work of rain and flow gauges, as well as a recently calibrated urban drainage model were available.

The paper is organised as follows. In Sect. 2 a detailed explanation is provided of the theoretical development of the singularity-sensitive method, including the newly implemented numerical strategies. In Sect. 3 we present the case study, including a description of the study area and data set, the performance criteria used to evaluate the proposed methodology, and the results of the testing. Lastly, in Sect. 4 the main conclusions are presented and future work is discussed.

## 2 Formulation of the singularity-sensitive Bayesian data merging method

Firstly, a description is provided of the two key techniques used in this paper: the Bayesian data merging method and the local singularity analysis. Afterwards, the proposed method for integrating these two techniques is explained. Intermediate results of each of the steps described in this section, which help illustrate the main features of the proposed methodology, can be found in the Supplement.

### 2.1 Bayesian radar–rain gauge data merging method

The Bayesian data merging method (BAY) is a dynamic adjustment method (applied independently at each time step) intended for real-time applications (Todini, 2001). The underlying idea of this method is to analyse and quantify the uncertainty of rainfall estimates (in terms of error co-variance) from multiple data sources – in this case, radar and rain gauge sensors – and then combine these estimates in such a way that the overall (estimation) uncertainty is minimised. The BAY

merging method consists of the following steps (illustrated in Fig. 1a):

a. For each time step $t$, the point rain gauge (RG) measurements are interpolated into a synthetic rainfall field using the block kriging (BK) technique. The result of this step is an interpolated rain gauge rainfall field, with areal estimates at each radar grid location ($y_t^{RG}$), and which are accompanied by the associated estimation error co-variance function ($V_{\varepsilon_t}^{RG}$), representing the uncertainty of rain gauge estimates.

b. The interpolated rain gauge rainfall field is compared against the radar field ($y_t^{RD}$), based upon which a field of errors (estimated as the bias at each radar grid location: $\varepsilon_t^{RD} = y_t^{RD} - y_t^{RG}$) is obtained empirically. Assuming that areal rain gauge estimates are unbiased, the expectation value ($\mu_{\varepsilon_t^{RD}}$) and the co-variance function ($V_{\varepsilon_t}^{RD}$) of this radar–rain gauge error field at each time step is used to represent, respectively, the mean bias and the uncertainty of radar estimates.

c. Using a Kalman filter (Kalman, 1960), the two rainfall fields are optimally combined such that the overall estimation uncertainty is minimised. In the Kalman filter the radar data and the interpolated rain gauge estimates act, respectively, as "a priori estimate" and "measurement". The degree of "uncertainty" of each type of estimate constitutes a gain value (the so-called Kalman gain, $K_t$) at each radar grid location, and determines the proportion of each type of estimate that is used to compute the merged output. The use of this gain value ensures the minimisation of the overall estimation uncertainty and is expressed as

$$K_t = V_{\varepsilon_t^{\mathrm{RD}}} \left( V_{\varepsilon_t^{\mathrm{RD}}} + V_{\varepsilon_t^{\mathrm{RG}}} \right)^{-1}, \tag{1}$$

and the optimally merged output, BAY (i.e. the a posteriori estimates $y_t''$ in the Kalman filter) can be obtained from

$$y_t'' = y_t' + K_t \left( y_t^{\mathrm{RG}} - y_t' \right), \tag{2}$$

where $y_t'$ is the "unbiased" radar rainfall estimate (i.e. $y_t' = y_t^{\mathrm{RD}} - \mu_{\varepsilon_t^{\mathrm{RD}}}$) used as the a priori estimate in the Kalman filter.

It can be seen that the Kalman gain is a function of the co-variances of radar and rain gauge estimation errors. When $V_{\varepsilon_t^{\mathrm{RD}}} \gg V_{\varepsilon_t^{\mathrm{RG}}}$ (or $K_t \approx 1$, i.e. radar estimates have significantly higher uncertainty than the rain gauge ones), the radar estimates are less trustworthy and the output estimates will be very similar to the interpolated rain gauge field. In contrast, when $V_{\varepsilon_t^{\mathrm{RG}}} \gg V_{\varepsilon_t^{\mathrm{RD}}}$ (or $K_t \approx 0$), the output will be closer to the radar estimates.

It is in steps b and c where the problems associated with the Bayesian merging technique, and geostatistical techniques in general, arise. The (second-order) co-variance function that these techniques employ to characterise radar–rain gauge errors cannot well capture local singularity structures. Instead, in second-order models singularities may be mistakenly regarded as errors in the radar data, thus leading to higher estimated radar uncertainty, $V_{\varepsilon_t^{\mathrm{RD}}}$. As a result, the radar data will be trusted less, leading to smoother merged outputs, which are closer to the interpolated rain gauge field.

## 2.2 Local singularity analysis

Various types of hazardous geo-processes, including precipitation, often result in anomalous amounts of energy release or mass accumulation confined to narrow intervals in time and/or space. The property of anomalous amounts of energy release or mass accumulation is termed *singularity* and it is often associated with structures depicting fractality or multi-fractality (Agterberg, 2007; Cheng, 1999; Lovejoy and Mandelbrot, 1985; Schertzer and Lovejoy, 1987). Several mathematical models and methodologies have been developed to respectively characterise and treat singularities. In this work, the local singularity analysis proposed by Cheng et al. (1994) has been adopted to identify and extract singularities form rainfall fields. Cheng's method, which has been widely used for estimation of geo-chemical concentrations (Agterberg, 2007; Cheng and Zhao, 2011; Cheng et al., 1994), employs the definition of coarse Hölder exponent to characterise singularities. According to this model, singularities are defined by the fact that the areal average measure (in this work, areal rainfall) centred on point $x$ (taken as the centre of a radar pixel) varies as a power function of the area (Evertsz and
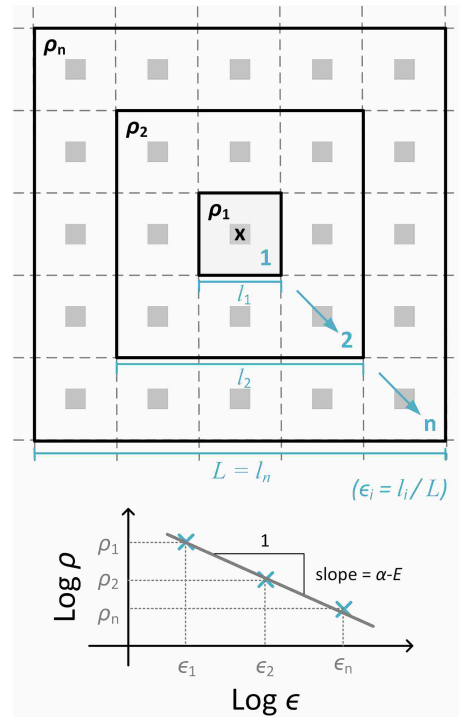


**Figure 2.** Schematic of the local singularity analysis (adapted from Wang et al., 2014).

Mandelbrot, 1992). This power-law relationship can be formulated as an equation (Cheng et al., 1994):

$$\rho(x, \epsilon) = c(x) \epsilon^{\alpha(x) - E}, \tag{3}$$

where $\rho(x, \epsilon)$ represents the density of measure (e.g. concentration of geo-data; in the context of this paper, rainfall intensity) over a square area with side-length $l$ and associated scale $\epsilon$ ($\epsilon = l/L$, where $L$ is the side-length of the largest square area under consideration) centred at a specific location $x$; $c(x)$ is a constant value (in the context of this paper, a constant intensity value) at $x$; $\alpha(x)$ is the singularity index (or the coarse Hölder exponent); and $E = 2$ is the Euclidean dimension of the plane.

A schematic of the estimation of the constant value $c(x)$ and singularity index $\alpha(x)$ from gridded data is provided in Fig. 2. For a given pixel with centre $x$ (centre of top plot in Fig. 2), the mean rainfall intensities at different spatial scales (centred in $x$) can be calculated (i.e. rainfall intensities $\rho_1$, $\rho_2$, ..., $\rho_n$, respectively at scales $\epsilon_1$, $\epsilon_2$, ..., $\epsilon_n$). Then, the logarithms of these mean values and the associated spatial scales are compared (bottom plot in Fig. 2). The constant value $c(x)$ and singularity index $\alpha(x)$ of the data set can be derived by applying a simple linear regression analysis, where the slope and the $y$-intercept of the regression line correspond, respectively, to the terms $(\alpha(x) - E)$ and $\log c(x)$. A detailed explanation of the computation of $c(x)$ and $\alpha(x)$ can be found in previous studies (Agterberg, 2012b; Chen et al., 2007; Cheng et al., 1994). It is worth mentioning that the estimation of

$c(x)$ and $\alpha(x)$ can be trusted only if a good linear relation is observed (i.e. if the scaling behaviour is well followed).

Going back to the definition of singularity, Eq. (3) constitutes a useful tool to decompose an areal rainfall intensity at a given location $x$ into two components (Wang et al., 2012): (1) the background (or non-singular, NS) magnitude $c(x)$, which is invariant as measuring scale $\epsilon$ changes and is more approximately normal than the original field; and (2) a local "scaling" multiplier, the magnitude of which changes as the measuring scale $\epsilon$ changes, according to the local singularity index $\alpha(x)$. When $\alpha(x) < 2$, the rainfall magnitude strikingly increases as the measuring scale $\epsilon$ decreases (namely local enrichment); this corresponds to a "peak" singularity. In contrast, when $\alpha(x) > 2$, the rainfall magnitude decreases as $\epsilon$ decreases (i.e. local depletion), and it is therefore a "trough" singularity. When $\alpha(x) = 2$, there is no singularity: the rainfall intensity $\rho(x, \epsilon)$ within a $\epsilon \times \epsilon$ area remains the same as scale changes (i.e. $\rho(x, \epsilon) = c(x)$).

In practice however, there is a drawback to this local singularity analysis. Because it carries out a "local" analysis, the singularity exponents are usually obtained from a small number of data samples. This increases the uncertainty of the estimation of $\alpha(x)$. The consequence of this drawback is that the singularity is incorrectly estimated or incompletely extracted; therefore, $c(x)$ is an unreliable or incomplete non-singular value. To circumvent this, two numerical strategies were employed in this study. The first one involves constraining the value of the estimated singularity exponents within a certain range. This can avoid obtaining unreasonably large or small singularity exponents. A number of ranges, symmetric to the non-singular condition (i.e. $\alpha(x) = 2$), were selected for testing. They are (from the widest to narrowest intervals): SIN1 = [0, 4], SIN2 = [0.5, 3.5], SIN3 = [1, 3], SIN4 = [1.5, 2.5] and SIN5 = [1.75, 2.25]. These "truncated" singularity ranges were empirically chosen according to the authors' experience and the fact that the distribution of $\alpha(x)$ is seldom largely skewed to a specific side of $\alpha(x) = 2$. It can be generally expected that, the wider the range is, the more singularity information (both local enrichment and depletion) from the radar images is taken into account in the merging process. The impact of different singularity ranges in rainfall estimation and hydraulic simulation is further discussed in Sect. 3.3.

The second numerical strategy is to decompose the rainfall field using an iterative procedure (Agterberg, 2012b; Chen et al., 2007):

$$c^{(k-1)}(x) = c^{(k)}(x)\epsilon^{\alpha^{(k)}(x) - E}, \tag{4}$$

where the iterative index $k = , 0, 1, 2, \ldots, n$. As $k = 0$, $c^{(-1)}(x) = \rho(x, \epsilon)$ (i.e. the original value) and $c^{(0)}(x)$ is the "calculated non-singular" value from the first iteration, which is equal to $c(x)$ from the non-iterative calculation above (Eq. 3). This $c^{(0)}(x)$ is then used as the left-hand-side value of Eq. (4) to calculate the "non-singular" value at the

next iteration, and so on. Substituting Eq. (4) into Eq. (3), one can obtain an iterative local singularity analysis equation:

$$\rho(x, \epsilon) = c^*(x)\epsilon^{\alpha^*(x) - E}, \tag{5}$$

where

$$\begin{cases} c^*(x) = c^{(n)}(x) \\ \alpha^*(x) = \alpha^{(0)}(x) + \sum_{k=1}^{n} \left( \alpha^{(k)}(x) - E \right). \end{cases} \tag{6}$$

The criterion to terminate the iteration procedure is when $\alpha^{(k)}(x) \approx E$ (which is equivalent to $c^{(k-1)}(x) \approx c^{(k)}(x)$). That means the singularity components have been clearly removed from the data.

Moreover, in this work a spatial-scale range of 1–9 km, which results in a total of five rainfall intensity samples (at scales 1, 3, 5, 7 and 9 km), was used in the singularity analysis. This range was selected for two main reasons. Firstly, our analyses revealed that a good linear behaviour was generally observed within this scale range, while small-scale structures were still preserved in the resulting rainfall product. As such, the selected spatial scale range was deemed to represent a good balance between estimation uncertainty (which depends upon the number of samples employed in the calculations) and local feature preservation. Secondly, a scaling break at approximately 8–16 km has been reported in studies in which 1 km radar rainfall data were analysed (Gires et al., 2012; Tchiguirinskaia et al., 2012). This means that the rainfall data at spatial-scale regimes ranging from 1 to 8–16 km comply with the same or similar statistical or physical behaviour. This scaling range has also been used in other applications to represent relatively local characteristics of rainfall fields (Bowler et al., 2006).

Lastly, a 10-iteration singularity analysis was applied in order to ensure that most of the singularity exponents could be extracted. The downside of conducting many iterations is the longer computational time, which may be an issue for real-time applications. Nonetheless, in practice, approximately 4–6 iterations are sufficient for effectively removing most of the singularity.

## 2.3 Incorporation of the local singularity analysis into the Bayesian merging method

The underlying idea of the proposed method is to use the local singularity analysis to decompose each radar image into a non-singular image and a singularity map before applying the Bayesian merging (step (i) in Fig. 1b). The non-singular radar image (NS-RD), which has a distribution closer to normality (thus being more suitable for Gaussian-based treatments), is merged with the point rain gauge data following the Bayesian procedure. This yields a non-singular Bayesian merged field (NS-BAY). Afterwards, the singularity map is applied back and proportionally to the NS-BAY merged field (step (ii) in Fig. 1b), thus yielding a singularity-sensitive merged field (SIN). This is done by multiplying each pixel
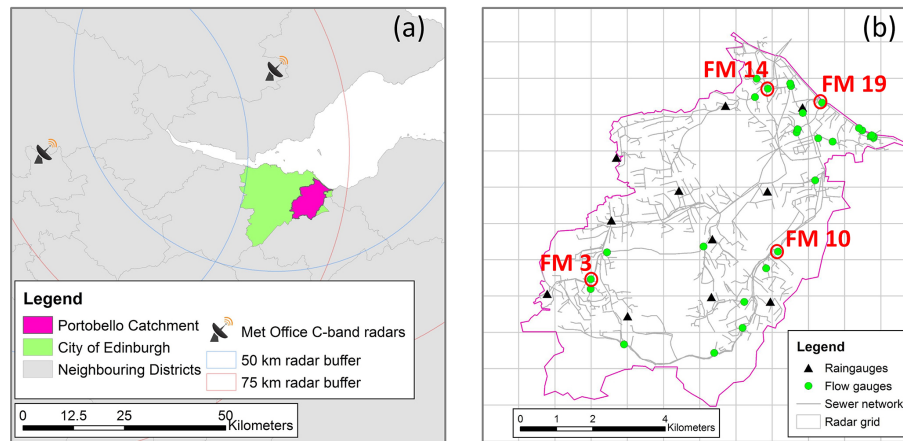
**Figure 3.** Portobello catchment **(a)** general location; **(b)** sensor location, sewer network and radar grid over the catchment. Flow gauges FM 3, 10, 14 and 19 (the circled round markers in **(b)**, respectively located at the up-, mid- and downstream parts of the catchment) are particularly selected for visual inspection in Sect. 3.3.2.

value of the NS-BAY field by the following ratio:

$$r(x) = \epsilon^{\alpha^*(x) - E} \tag{7}$$

which corresponds to the ratio difference between the original radar field (RD) and the non-singular radar field (NS-RD). In this way, a singularity-sensitive merged field (SIN), which better retains the local singularity structures embedded in the original RD field, is obtained.

It is worth noting that the proposed singularity-sensitive merging method does not always increase the reliability of RD estimates. Such increase only happens when the RD estimates exhibit high singularity and thus cannot be well handled using Gaussian approximations.

A particular phenomenon which may cause problems in the application of the proposed methodology and is therefore worth highlighting is the eventual presence of singularity structures in the interpolated rain gauge field (i.e. BK field) and in the resulting supposedly non-singular Bayesian (NS-BAY) merged field. While BK fields are generally highly smooth, singularity structures may appear in the special case in which a rain gauge is located within a convective cell or a local depletion. Singularity structures in the BK field may be preserved in the NS-BAY field. When this is the case, the application of the singularity map back and proportionally to the NS-BAY field may result in double-counting of singularities. This can ultimately result in a merged (SIN) product with more singularities than those originally observed in the radar image. In order avoid this, a "moving window" smoothing has been applied to the BK field before it is merged with the NS-RD field. That is, each pixel value of the BK field is replaced by the mean of the original value and neighbouring pixel values within a 9 km diameter (which is equal to the coarsest scale considered in the local singularity analysis). In this way singularity structures potentially present in the BK field are smoothed-off.

## 3 Case study

The proposed SIN merging method is tested using as case study four storm events observed in the Portobello catchment (Edinburgh, UK) during 2011 and for which radar estimates, dense rain gauge and flow records, as well as a recently calibrated urban drainage model were available. Portobello is a coastal town located 5 km to the east of the city centre of Edinburgh, along the coast of the Firth of Forth, in Scotland (Fig. 3a). The catchment is predominantly urban, of residential character. It stretches over an area of approximately 53 km$^2$, of which 27 km$^2$ are drained by the sewer system. Of the drained area, 46 % corresponds to impervious surfaces. This includes a small western part of Edinburgh city centre and the surrounding southwestern region. The storm water drainage system is predominantly combined with some separate sewers and drains from the southwest to the northeast (towards the sea).

### 3.1 Available models and data sets

#### 3.1.1 Urban storm-water drainage model

A semi-distributed model of the storm-water drainage system of the Portobello catchment, including its sewer system (Fig. 3b), was set up by the water utility of the area in the commercial modelling package InfoWorks CS v13.0. In this model the whole catchment surface is split into sub-catchment units through which rainfall is applied (within each sub-catchment rainfall is assumed to be uniform). Each sub-catchment comprises a mix of pervious and impervious surfaces whose runoff drains to a common outlet point, which corresponds to an inlet node of the sewer system (i.e. a gully or a manhole). Each sub-catchment is characterised by a number of parameters, including total area, length, slope, proportion of each land use, amongst others. Based upon

**Table 1.** Selected rainfall events over the Portobello catchment.

| Event | Date | Duration (h) | RG total (mm) | RG peak intensity (mm h$^{-1}$) | RD total (mm) | RD peak intensity (mm h$^{-1}$) |
|-------|------|-------------|---------------|---------------------------------|---------------|---------------------------------|
| Storm 1 | 6–7 May 2011 | 7 | 9.25 | 11.21 | 9.67 | 7.93 |
| Storm 2 | 23 May 2011 | 7 | 7.70 | 5.03 | 11.02 | 4.10 |
| Storm 3 | 21–22 Jun 2011 | 24 | 32.96 | 8.46 | 26.21 | 3.33 |
| Storm 4 | 8 Jun 2011 | 11 | 5.03 | 7.11 | 4.69 | 2.41 |

Note: the accumulation and peak intensity values shown in this table correspond to areal mean values for the entire domain under consideration (as shown in Fig. 3b).

these parameters, the runoff volume at each sub-catchment is estimated using the NewUK rainfall–runoff model (Osborne, 2001). The estimated runoff is then routed to the sub-catchment outlet using the Wallingford (double linear reservoir) model (HR Wallingford, 1983). Sewers are modelled as one-dimensional conduits and flows within them are simulated based on the full de Saint-Venant equations (i.e. fully hydrodynamic model).

The Portobello model contains a total of 1116 sub-catchments, with areas ranging between 0.02 and 24.42 ha and a mean area of 2.3 ha. Sub-catchment slopes range from 0.0 to 0.63 m m$^{-1}$, with a mean slope of 0.031 m m$^{-1}$. The model of the sewer system comprises 2917 nodes and 2907 conduits, in addition to 14 pumps. The total length of modelled sewers is 250 km. The sewer system ranges in height from 186.6 mAOD at Comiston to 3.8 mAOD along the Firth of Forth. 2 % of the modelled pipes have a gradient between 0.1 and 0.25 m m$^{-1}$, 55 % have a gradient between 0.01 and 0.1 m m$^{-1}$, and 43 % have a gradient < 0.01 m m$^{-1}$.

Following UK standards (WaPUG, 2002) and using solely rain gauge data as input, the model of the Portobello catchment was manually calibrated in 2011 based upon three storm events recorded during the medium-term flow survey described below. For the three storm events used for model calibration, the following mean performance statistics were obtained: Nash–Sutcliffe efficiency of 0.5782, root mean square error of 0.0373 m$^3$ s$^{-1}$, coefficient of determination ($R^2$) of 0.7756 and regression coefficient ($\beta$), which provides a measurement of conditional bias, of 0.8826.

### 3.1.2 Local monitoring data (medium term survey)

Local rainfall and flow data were collected in the Portobello catchment through a medium-term flow survey carried out between April and June 2011. The survey comprised 12 tipping bucket rain gauges and 28 flow monitoring stations (each comprising a depth and a velocity sensor, based upon which flow rates were estimated). Both rain gauge and flow records were available at a temporal resolution of 2 min. However, rain gauge records were linearly interpolated to 5 min, in order to ensure agreement with the temporal resolution at which radar estimates were available (see Sect. 3.1.3).

The location of the local flow and rain gauges is shown in Fig. 3b.

### 3.1.3 Radar rainfall data

The Portobello catchment is within the coverage of C-band radars operated by the UK Met Office (Fig. 3b). Radar rainfall estimates for the same period as the local flow survey (i.e. April–June 2011) were obtained through the British Atmospheric Data Centre (BADC). These estimates were available at spatial and temporal resolutions of 1 km and 5 min, respectively, and correspond to a quality controlled multi-radar composite product generated with the UK Met Office Nimrod system (Golding, 1998), which incorporates corrections for the different errors inherent to radar rainfall measurements, including identification and removal of anomalous propagation (e.g. beam blockage and clutter interference), attenuation correction and vertical profile correction (for a full description of the Nimrod system, the reader may refer to Harrison et al. (2000, 2009)).

### 3.1.4 Storm events selected for analysis

During the monitoring period (April–June 2011), four relevant storm events were captured which comply with UK standards for calibration and verification of urban drainage models (i.e. these events have instantaneous rainfall rates > 5 mm h$^{-1}$ and accumulation > 5 mm) (Gooch, 2009). Three of these events (referred to as Storms 1, 2 and 3) were used for calibration of the urban drainage model of the Portobello catchment (following UK standards, as mentioned above). In this study, all four storm events, including one not used in the calibration of the model (Storm 4), were used to test the proposed singularity-sensitive merging method. The dates and main statistics of the four selected events are summarised in Table 1. It is worth mentioning that, given the response time of the catchment, as well as inter-event time definition thresholds (IETD) recommended in the literature (Guo and Adams, 1998), a minimum IETD of 6 h was used as criteria to differentiate rainfall events.

As can be seen in Table 1, Storms 1, 2 and 4 are of relatively short duration, whilst Storm 3 is of much longer du-

ration. Moreover, well-structured storm cell clusters crossing the catchment area can be found in Storms 1, 3 and 4, but not in Storm 2. This is reflected in the lower RG peak intensity observed in Storm 2.

## 3.2 Evaluation methodology

The performance of the proposed singularity-sensitive Bayesian method (SIN hereafter) is assessed by inter-comparison against radar (RD), rain gauge (RG) and block-kriged (BK) interpolated RG estimates, as well against adjusted estimates resulting from the original Bayesian (non singularity-sensitive) technique (BAY) and the commonly used mean field bias (MFB) adjustment method. It is important to note that, in this work, the MFB was implemented in a relatively dynamic way by computing a sample cumulative bias ($B_t$) at each time step as $B_t = \sum \text{RG}_{\text{catchment}} / \sum \text{RD}_{\text{co-located}}$, where $\text{RG}_{\text{catchment}}$ and $\text{RD}_{\text{co-located}}$ represent the rain gauge and the co-located radar grid rainfall estimates over the experimental catchment during the last hour. The MFB adjusted estimates are obtained by multiplying the bias ($B_t$) at each particular time step by the original rainfall field; that is, $\text{MFB} = B_t \cdot \text{RD}$.

Two evaluation strategies were applied:

1. Through analysis of the different rainfall estimates, using as main reference local rain gauge records, while also inter-comparing the behaviour of other estimates.

2. Through analysis of the hydraulic outputs obtained by feeding the different rainfall estimates as input to the hydraulic model of the Portobello catchment and comparison of these with available flow records. Note that the RG estimates were applied to the model using Thiessen polygons.

Both evaluation strategies have inherent limitations which are next described. However, they provide useful and complementary insights into the performance of the proposed merging method.

The first strategy is a natural and widespread way of assessing the performance of rainfall products. However, the fact that all precipitation estimates entail errors and that the true rainfall field is unknown, in addition to the differences in the spatial and temporal resolutions of RG and RD estimates (and the resulting merged rainfall products), renders any direct comparison of rainfall estimates imperfect (Brandes et al., 2001). In the particular case of rain gauge records, which are used as main reference in the present investigation, errors can arise from a variety of sources (Einfalt and Michaelides, 2008). In order of general importance, systematic errors common to all rain gauges include errors due to wind field deformation above the gauge orifice, errors due to wetting loss in the internal walls of the collector, errors due to evaporation from the container, and errors due to in- and out-splashing of water. In addition, tipping bucket rain gauges, such as the ones used in this investigation, are known to underestimate rainfall at higher intensities because of the rainwater amount that is lost during the tipping movement of the bucket (La Barbera et al., 2002). This is a systematic error unique to the tipping bucket rain gauge and is of the same order of importance as wind-induced losses. Besides these systematic errors, tipping bucket rain gauge records are also subject to local random errors, mainly related to their discrete sampling mechanism (Ciach, 2003; Habib et al., 2008). The order of magnitude of the two main error sources associated with tipping bucket rain gauges is 2–10 % for wind-induced losses (depending on wind speed, weight of precipitation and gauge construction parameters) (Sevruk and Nešpor, 1998) and up to 10 % for rainfall intensities of $100\,\text{mm}\,\text{h}^{-1}$ and 20 % for intensities of $200\,\text{mm}\,\text{h}^{-1}$, for water losses during the tipping action (Luyckx and Berlamont, 2001; Molini et al., 2005). Given that the storm events under investigation were not extremely heavy and that a basic quality-control was applied to ensure the quality of rain gauge measurements (e.g. manual comparison between neighbouring rain gauge data), the rain gauge records used in this study can be deemed acceptable as reference for the evaluation of other rainfall products. However, they are not perfect and the reader must keep in mind that the true rainfall field remains unknown.

The second strategy (i.e. hydraulic evaluation) allows some of the limitations of the rainfall evaluation strategy to be overcome, and is particularly useful when dense flow records are available, as is the case in the Portobello catchment. However, it has two main deficiencies: the fact that flow records (obtained based upon depth and velocity measurements) used in the evaluation contain errors, and the fact that the hydraulic modelling results encompass uncertainties from different sources in addition to rainfall input uncertainty (Deletic, 2012; Kavetski et al., 2006). In this regard, it is worth reminding the reader that the hydraulic model used in this study was calibrated using as input RG data. In fact, the model was calibrated based upon the data of Storms 1–3, which are used for testing in the present paper (Storm 4 is the only event not used in the calibration of the model). Since the model was "attuned" for RG inputs, it favours all RG-derived (and RG-emulating) rainfall estimates. Moreover, the relatively coarse spatial resolution of rainfall inputs (in this case $\sim 1\,\text{RG}/4.4\,\text{km}^2$) may have led to further biases in the model (Kavetski et al., 2006). It would be desirable to re-calibrate the hydraulic model using as input the different rainfall products analysed in this study. However, this would entail a significant amount of work which falls outside of the scope of the present study. In spite of these limitations, we believe that the hydraulic evaluation strategy using the currently available hydraulic model still provides useful insights into the performance of the proposed SIN merging method in relation to other rainfall estimates, and complements the findings of the rainfall analysis.

**Hydrol. Earth Syst. Sci., 19, 4001–4021, 2015**

**www.hydrol-earth-syst-sci.net/19/4001/2015/**

### 3.2.1 Methodology for analysis of rainfall estimates

The performance of the SIN rainfall products in relation to other rainfall estimates (including RD, RG, BAY and MFB) is evaluated in terms of accumulations and rainfall rates at the areal level (i.e. at a scale corresponding to the area over which the Portobello catchment stretches) and at individual point gauge locations. In addition, a qualitative assessment of the spatial structure of the different (gridded) rainfall products is carried out based upon visual inspection of images of the rainfall fields at the time of areal average peak intensity.

In view of the high density and coverage of the RG network over the Portobello catchment, the areal average RG estimates in the areal level analysis are assumed to be a good approximation of the "true" areal (average) rainfall over the experimental catchment (i.e. the areal reduction effect is expected to be minor – Bell, 1976) and are therefore used as reference to evaluate the performance of the different areal gridded rainfall estimates. The areal analysis includes comparison of event areal average accumulations and peak intensities, as well as comparison of intensities throughout the duration of each event (through scatterplots using RG areal average intensities as reference).

In the analysis of rainfall estimates at rain gauge point locations a cross-validation strategy was adopted and three performance statistics are used. The cross-validation strategy, also referred to as "leave-one-out", is an iterative method in which, at each iteration, data from one RG site is omitted from the calculations and the value at the "hidden" (i.e. omitted) location is estimated using the remaining data. Performance statistics are then computed from the comparison between the estimated and the known (but not used) values (Velasco-Forero et al., 2009). The following three performance statistics are used in the present study. Firstly, a sample bias ratio ($B$) is used to quantify the cumulative bias between gridded rainfall estimates (i.e. RD, BK, MFB, BAY and SIN) and RG estimates at each RG location over the event period under consideration. $B = 1$ means no cumulative bias between the RG and the given gridded rainfall estimates (i.e. equal rainfall accumulation recorded by RG and the gridded product at the given gauge location); $B > 1$ means that the accumulations of the gridded estimates at the point locations are greater than those recorded by RG, and $B < 1$ means the opposite. In addition to the comparison of rainfall accumulations, a simple linear regression analysis is applied to each pair of "instantaneous" (rain rate) point RG records and the co-located gridded estimates. The results of the regression analysis are presented in terms of $R^2$ (coefficient of determination) and $\beta$ (regression coefficient, i.e. the slope or gradient of the linear regression). These two measures provide an indication of how well RG rates are replicated by the different rainfall estimates at each gauging location, both in terms of pattern and accuracy. The $R^2$ measure ranges from 0 to 1 and describes how much of the "observed" (RG) variability is explained by the "mod-

elled" (RD/BK/adjusted) one. In practical terms, $R^2$ provides a measurement of the similarity between the patterns of the observed (i.e. RG) and "modelled" (i.e. gridded estimates) rainfall time series at a given gauging location. However, systematic bias (under- or over-estimation) of the modelled estimates cannot be detected from this measure (Krause et al., 2005). The regression coefficient, $\beta$, is therefore employed to provide this supplementary information to the $R^2$. $\beta \approx 1$ represents good agreement in the magnitude of the rainfall rates recorded by RG and those of the gridded estimates; $\beta > 1$ means that the rain rates associated with the gridded estimate are higher in the mean (by a factor of $\beta$) than those recorded by RG; and $\beta < 1$ means the opposite (i.e. rain rates of gridded estimates are lower in the mean than RG ones).

### 3.2.2 Methodology for analysis of hydraulic outputs

A qualitative analysis of the hydraulic outputs is carried out based upon visual inspection of recorded vs. simulated flow hydrographs (for the different rainfall inputs) at different points of the catchment. Furthermore, similar to the rainfall analysis, a simple linear regression analysis is applied to each pair of recorded and simulated flow time series (at each flow gauging location). The performance of the associated hydraulic simulations is evaluated using the $R^2$ and $\beta$ statistics obtained from the linear regression analysis. In addition, the "weighted" coefficient of determination ($R_{\mathrm{w}}^2$) is employed to quantify the joint performance of hydrological efficiency. This measure is defined as (Krause et al., 2005)

$$R_{\mathrm{w}}^2 = \begin{cases} |\beta| \cdot R^2 & \text{for } \beta \leq 1 \\ |\beta|^{-1} \cdot R^2 & \text{for } \beta > 1, \end{cases} \tag{8}$$

where higher $R_{\mathrm{w}}^2$ values correspond to better hydraulic performance.

In order to minimise the influence of the errors in the flow measurements, the available flow records were quality-controlled (QC) before carrying out the statistical analysis of hydraulic outputs. The QC was carried out following UK guidelines (WaPUG, 2002). It included analysis of depth vs. flow scatterplots at each monitoring location (the shape and spread of the resulting scatterplots provides insights into the quality and consistency of depth and velocity records at each site), as well as visual inspection of the observed hydrographs at each location. Whenever a flow monitor was deemed unreliable, it was manually removed from the analysis. Likewise, with the purpose of preventing systematic hydraulic modelling errors from affecting results, whenever the model was found to be unable to replicate the recorded flows at a given location, the given flow monitor was also manually removed from the analysis. This left us with a total of 16 flow monitors for analysis. Moreover, all records associated with depth measurements below 0.1 m were left out when estimating performance statistics; this is due to the fact that at low depths, both velocity and depth records become unreliable (the 0.1 m threshold was adopted based upon UK guidelines

**Table 2.** Areal average rainfall accumulations and peak intensities for the different rainfall products.

| | Input | Storm 1 | Storm 2 | Storm 3 | Storm 4 |
|---|---|---|---|---|---|
| AVG rainfall | RG | 9.25 | 7.70 | 32.96 | 5.03 |
| accumulation | RD | 9.80 | 11.02 | 26.21 | 4.69 |
| (mm) | BK | 8.87 | 7.77 | 31.41 | 4.08 |
| | MFB | 8.64 | 7.25 | 31.87 | 4.55 |
| | BAY | 8.76 | 7.79 | 27.24 | 3.81 |
| | SIN1 | 9.37 | 7.95 | 33.37 | 4.82 |
| | SIN2 | 9.37 | 7.95 | 33.38 | 4.83 |
| | SIN3 | 9.38 | 7.95 | 34.43 | 4.85 |
| | SIN4 | 9.41 | 7.96 | 33.73 | 4.94 |
| | SIN5 | 9.61 | 8.04 | 31.57 | 5.23 |
| AVG rainfall peak | RG | 11.21 | 5.03 | 8.46 | 7.11 |
| intensity over 5 min | RD | 7.93 | 4.10 | 3.33 | 2.41 |
| (mm h$^{-1}$) | BK | 10.66 | 4.54 | 7.59 | 5.09 |
| | MFB | 9.06 | 4.68 | 4.05 | 3.34 |
| | BAY | 10.47 | 3.80 | 6.82 | 4.92 |
| | SIN1 | 13.17 | 5.08 | 8.00 | 6.97 |
| | SIN2 | 13.17 | 5.08 | 8.00 | 6.97 |
| | SIN3 | 13.17 | 5.08 | 8.00 | 6.98 |
| | SIN4 | 13.19 | 5.08 | 8.01 | 7.03 |
| | SIN5 | 13.53 | 5.09 | 8.27 | 7.25 |

and recommendations from studies focusing on the performance of flow gauges; Marshall and McIntyre, 2008).

## 3.3 Results and discussion

### 3.3.1 Rainfall estimates

#### Areal rainfall estimates

Table 2 shows the areal average (AVG) accumulations and peak intensities for the different rainfall products for the four storm events under consideration. As can be seen, the difference between RD and RG event areal average accumulations is generally small, yet it is event-varying. For Storms 1 and 2, the areal average RG totals are slightly overestimated by the RD estimates, while for Storms 3 and 4 they are underestimated, with the underestimation being largest in Storm 3. In terms of areal average peak intensities, RD estimates appear to consistently underestimate the areal average peak intensities recorded by RG. The relative difference between RG and RD areal average peak intensities is approximately 20–30 % for Storms 1 and 2, and it is as high as 60 % for Storms 3 and 4. This indicates that the RD estimates could not satisfactorily capture instantaneous rainfall rates, particularly high rainfall rate values, and corroborates the need for dynamic adjustment of RD estimates using local RG measurements.

As would be expected, the BK estimates exhibit areal average accumulations and peak intensities similar to those of the RG. Small differences are observed (in general BK values are slightly lower than RG ones) which can be generally attributed to the area-point rainfall differences (Anagnostou and Krajewski, 1999). These differences become more evi-
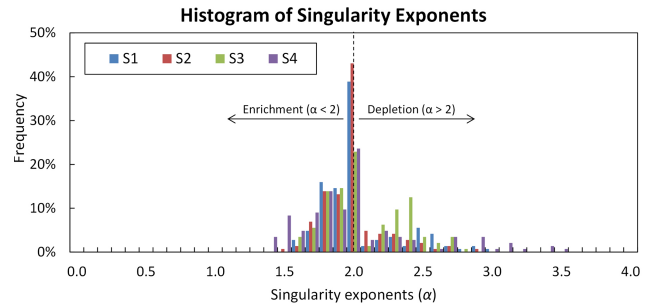


**Figure 4.** Histogram of singularity exponents ($\alpha$) at the time of areal peak intensity for the four selected storm events.

dent when analysing results at individual gauge locations (in next section).

When looking at the adjusted rainfall products (i.e. MFB, BAY and SIN), it can be seen that all of them can improve the original RD estimates, but the degree of improvement is different for each method. As expected, the MFB successfully reduces the difference in event areal average accumulations (i.e. bias), leading to areal average accumulations close to those recorded by RG. In terms of peak intensities, the MFB method leads to some improvement, but the resulting peak intensities are still significantly lower than the RG ones. Although the MFB was applied dynamically with an hourly frequency of bias correction, these results suggest that more dynamic and spatially varying (higher order) methods than the MFB are required in order to successfully adjust radar rainfall estimates for urban hydrological applications.

The BAY estimates show the least improvement in terms of event bias, with a general tendency to underestimate RG areal accumulations, which is even more marked than for BK estimates. This is particularly the case in Storms 3 and 4, in which strong singularity structures, as represented by the high frequency of $\alpha$ values different from 2 (Fig. 4), were observed. This tendency to underestimate can be attributed to a combined effect of the BAY method "over-trusting" the BK estimates (which show a slight underestimation tendency at the areal level), in addition to smoothing off the singularity structures (often associated to strong intensities) originally present in the RD image. With regard to peak intensities, the BAY estimates display a larger improvement than the MFB ones, which demonstrates the benefits of more dynamic and spatially varying adjustment methods. However, the areal average peak intensities of the BAY estimates still underestimate the "true" (RG) areal peak intensities. Lastly, the SIN estimates, particularly SIN ranges 1–4, exhibit very good performance: both areal average accumulations and peak intensities of SIN estimates are close to those of RG, and no systematic over- or underestimation is observed in SIN estimates. The better performance of SIN estimates suggests that the singularity analysis can in fact improve the original BAY merging method. With regard to the impact of the singularity range, it can be seen that, as the range becomes narrower
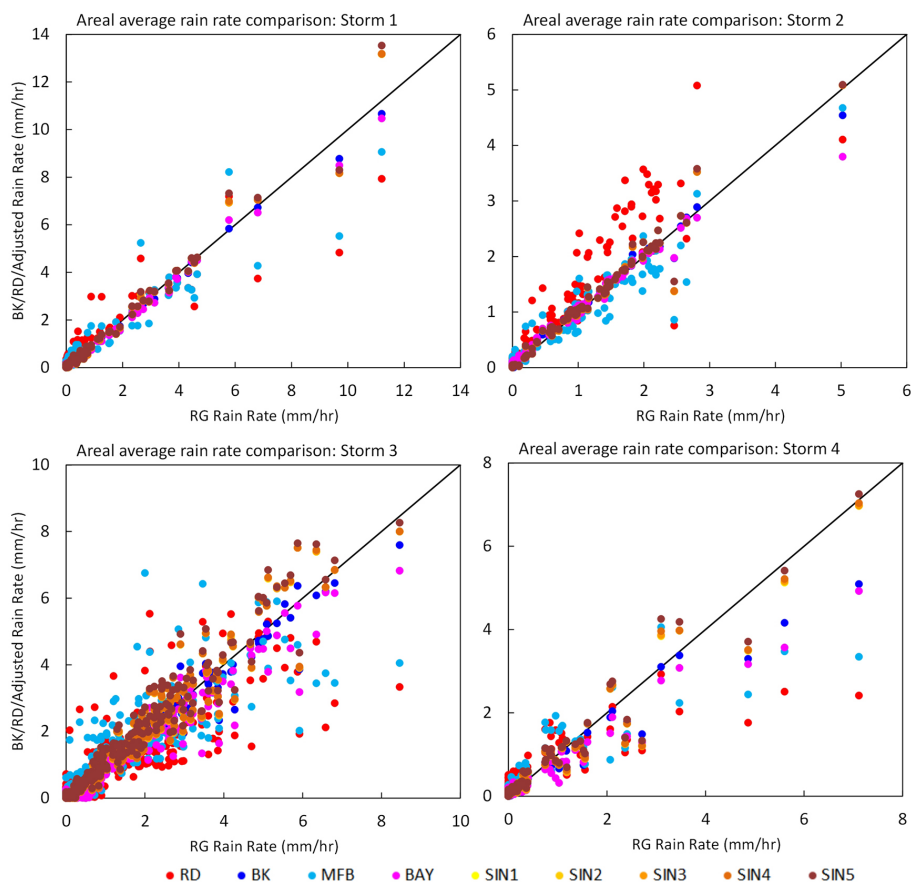
Hydrol. Earth Syst. Sci., 19, 4001–4021, 2015

www.hydrol-earth-syst-sci.net/19/4001/2015/

**Figure 5.** Scatterplots of instantaneous areal average RG vs. RD/BK/MFB/BAY/SIN rainfall rates over the Portobello catchment for the four selected events, where SIN1–SIN5 represent the SIN estimates with different "truncated" singularity ranges (from widest to narrowest).

(from SIN1 to SIN5), the areal accumulations and peak intensities tend to be higher. A particular steep increment is observed in the areal accumulations and peak intensities of SIN5, in relation to the other singularity ranges, which results in a slight overestimation of accumulations and peak intensities, as compared to the RG estimates. This is especially apparent in Storms 3 and 4. The increment in rainfall accumulation and peak intensities as the singularity range becomes narrower can be explained by the fact that at narrower ranges, only part of the singularity structures are removed before merging, and a big proportion of them remains in the radar image. This is evident from Fig. 4, where it can be seen that for Storms 3 and 4 a significant number of singularity exponents spread beyond the narrowest singularity range (i.e. SIN5: $\alpha \in [1.75, 2.25]$). The problem arises because some singularity features remain in the radar image before the BAY merging is applied, and these may be partially preserved throughout the merging process. Afterwards, when the extracted singularity component is applied back and proportionally to the merged rainfall field, it may interact (in a nonlinear fashion) with the singularity structures preserved throughout the BAY merging, thus leading to an overestima-

tion of extremes (whether these are enrichments or depletions). These results suggest that a better approach is to be sure to remove most of the singularity structures before carrying out the merging. Therefore, very narrow ranges such as SIN5 should be avoided. In fact, it can be seen that the intermediate range of SIN3 (i.e. $\alpha \in [1, 3]$) covers most singularity exponents (see Fig. 4). Indeed, using wider singularity ranges leads to very similar results as those obtained when using the SIN3 range (notice the similarity between SIN1, SIN2 and SIN3 estimates in Table 2). This suggests that the singularity range of [1, 3] is appropriate.

Figure 5 shows a further comparison of instantaneous areal average RG intensities vs. areal average BK, RD, BAY and SIN intensities throughout each of the storm events under consideration. As expected and in line with the analysis above, the areal BK estimates are generally in good agreement with areal RG estimates. Some underestimations can be observed at high RG rainfall intensities; nonetheless, most of them are fairly minor and are still within a reasonable range, which can be attributed to areal-point differences. With regard to the RD estimates, it can be seen that they tend to overestimate small rainfall intensities and underestimate the

peak intensities, with underestimation being more evident in the events with relatively high intensities (i.e. Storms 1, 3 and 4). Unlike BK estimates, the difference between areal RD and RG peak intensities is too large to be entirely explained by areal-point differences. The fact that the RD estimates display relatively good performance in (long-duration) accumulations but not in (short-duration) instantaneous intensities corroborates the claim that RD estimates fail to capture the short-term dynamics of small-scale rainfall. As mentioned in Sect. 1, the main reason for this lies in the use of a static $Z$–$R$ conversion function, which represents a statistical compromise for the range of rainfall rates that frequently occur (whereas the occurrence of very small and large intensities is relatively rare). Concerning the MFB method, from Fig. 5 it can be seen that it fails to satisfactorily improve RD instantaneous rainfall rates; this is particularly evident at high intensities, at which, similarly to RD estimates, MFB estimates perform poorly. An accurate representation of peak intensities is of outmost importance in the modelling and forecasting of urban pluvial flooding. This confirms that, being a first-order technique, the MFB adjustment method may be insufficient for urban-scale applications. In contrast, it can be seen that over- and underestimation errors in RD estimates can be improved by second- or higher-order adjustment techniques, such as BAY and SIN. In fact, in terms of instantaneous rainfall rates the BAY estimates display a significantly better performance than the MFB ones. However, the BAY estimates still fail to properly reproduce the highest intensities. These shortcomings of the BAY method seem to be overcome by the incorporation of the singularity analysis: indeed, the SIN estimates exhibit the best overall performance, particularly at peak intensities. In agreement with the results displayed in Table 2, in Fig. 5 it can be seen that as the singularity range becomes narrower, rainfall estimates with slightly higher intensities are generated. Moreover, it can be noticed that wider singularity ranges lead to more conservative results and appear to be a good choice.

**Rainfall estimates at gauging locations**

The aforementioned features of the different rainfall estimates are further highlighted through analysis at each rain gauge location; the associated statistics, including sample bias ($B$), regression coefficient ($\beta$) and coefficient of determination ($R^2$), are summarised in Fig. 6.

As expected, the RD estimates (before adjustment) display the largest differences from point RG estimates: in general, they possess the largest cumulative bias ($B$) and the lowest $R^2$, and their statistics show great variability. Moreover, the distribution of the $\beta$ values indicates that RD estimates tend to largely underestimate RG instantaneous rainfall rates. This is the case for all storm events, except for Storm 2, for which rainfall intensities were low on average.

Similarly to the results of the areal (average) analysis, the individual-site BK estimates display the closest behaviour to

the RG ones. This is of course expected given that the BK estimates are obtained by simple interpolation of point RG data. It can be seen that the BK estimates are nearly unbiased ($B$ is very close to 1) and possess the highest $R^2$ medians (i.e. closest to 1), as well as the narrowest $R^2$ boxes and whiskers. However, when looking at the distribution of $\beta$ values of the BK estimates, it can be seen that most of the time the whole boxes and whiskers are below 1. This reflects a systematic underestimation of rainfall rates at point gauging locations, which is discussed below.

With regard to the adjusted rainfall estimates, the MFB method is found to bring original radar estimates slightly closer to RG ones, but the improvement seems insufficient. As expected, the main improvement of MFB estimates is found in the bias ($B$), which is significantly reduced (thus becoming closer to 1). In terms of instantaneous rainfall rates, the improvement provided by MFB is very limited. This is reflected in the low $R^2$ values and in the poor $\beta$ scores, which remain remarkably close to those of the original RD estimates. Similar to the results of the areal analysis, these results suggest that the MFB adjustment method is insufficient for urban-scale applications, in which small-scale rainfall dynamics are critical.

When looking at the statistics of the BAY estimates, it can be noticed that these behave similarly to the BK ones: their bias is also small, the $R^2$ is generally high and the $\beta$ values are systematically below 1. The similarity in the behaviour of BK and BAY estimates suggests that for the selected events, the BAY method tends to trust the (smooth interpolated) BK estimates more than the RD estimates. As explained in Sects. 1 and 2, this is the main shortcoming of the BAY method. The systematic underestimation of rainfall rates observed in BK and BAY estimates (reflected by $\beta$ values systematically below 1) can be partially explained by the areal reduction effect. However, considering that the individual-site comparison was conducted using instantaneous rainfall estimates for very short time intervals (i.e. 5 min), one would expect the tendency of the areal-point differences to be of higher randomness, rather than of a systematic nature. This suggests that the systematic underestimation may be a joint consequence of the areal reduction effect and of the underlying second-order approximation (which smooths off some local extreme magnitudes).

With regard to the SIN estimates, it can be seen that their bias is small (close to 1) and that the distribution of their $R^2$ values is somewhere between that of the BAY and RD estimates. This indicates that, as compared to the original BAY estimates, the SIN method incorporates more spatial features from the RD estimates throughout the merging process, while retaining the accuracy of the RG estimates. In terms of $\beta$, it can be seen that although the median values are usually below 1, they are generally much closer to 1 than other rainfall estimates. Moreover, their distribution is more variable than that associated with BK and BAY estimates (which display a systematic underestimation). In line with the re-
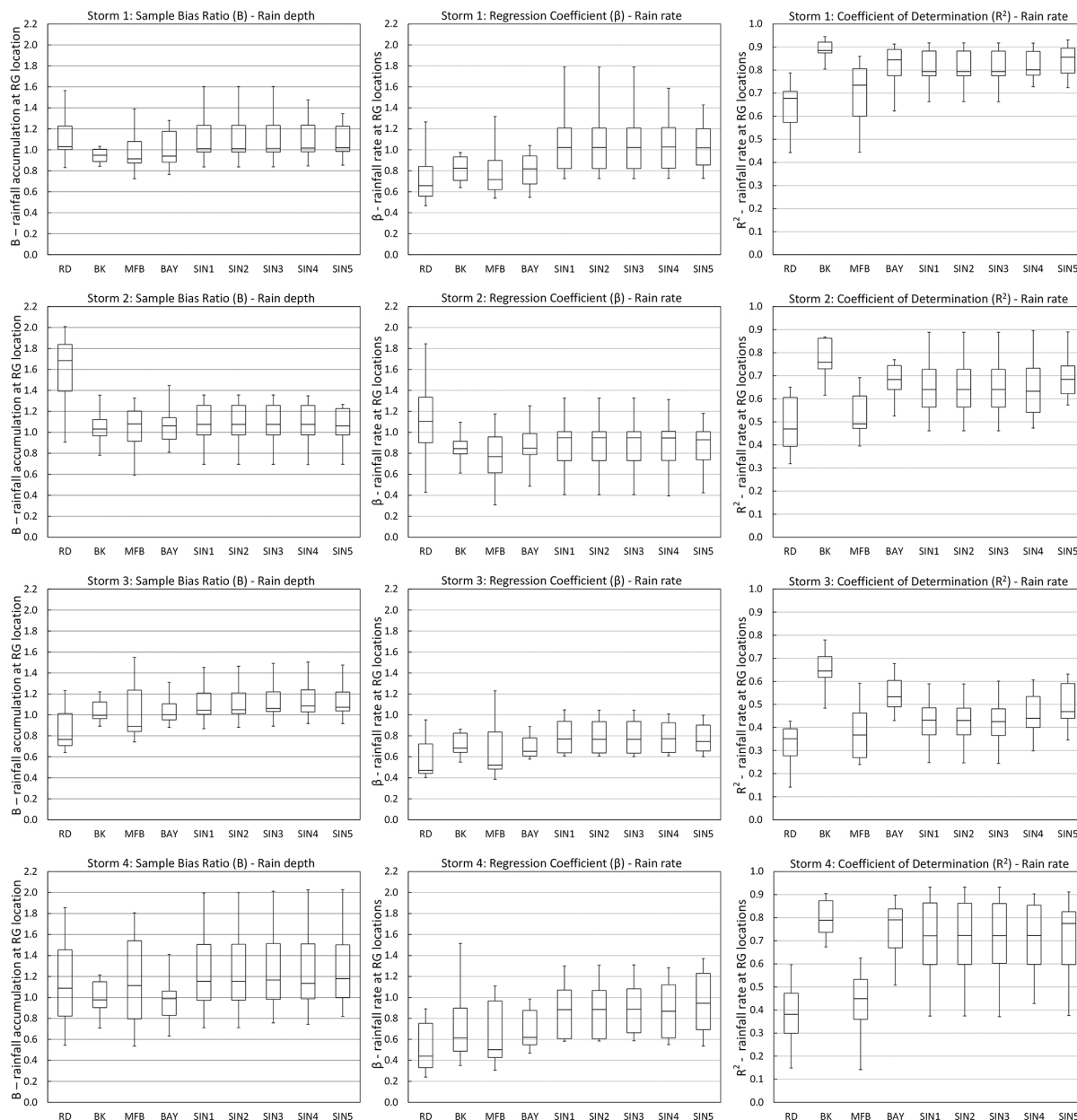
**Figure 6.** Boxplots displaying the distribution of sample bias ratio ($B$) (left column panels), regression coefficient ($\beta$) (middle column panels) and coefficient of determination ($R^2$) (right column panels) estimated between the different gridded rainfall estimates and RG records at individual rain gauge locations following a cross-validation approach.

sults of the areal analysis, this serves to further highlight two important features of the proposed SIN method: it generally respects the commonly observed areal reduction effect, and it integrates more small-scale randomness from RD data in the data-merging process. Regarding the singularity ranges, a similar behaviour can be observed for all of them, although a slight tendency can be observed for the bias ($B$) and $\beta$ values to increase as the singularity range becomes narrower. This is in agreement with the results of the areal analysis and suggests that working with intermediate ranges, as opposed

to very narrow ones which can truncate the singular structures, is advisable.

**Spatial structure of rainfall fields**

Snapshot images of the different gridded rainfall products at the time of peak areal intensity for the four storm events under consideration are shown in Fig. 7. Due to space constraints, images of only one of the SIN ranges (the intermediate one, SIN3: $\alpha \in [1, 3]$) are shown in this figure. As men-
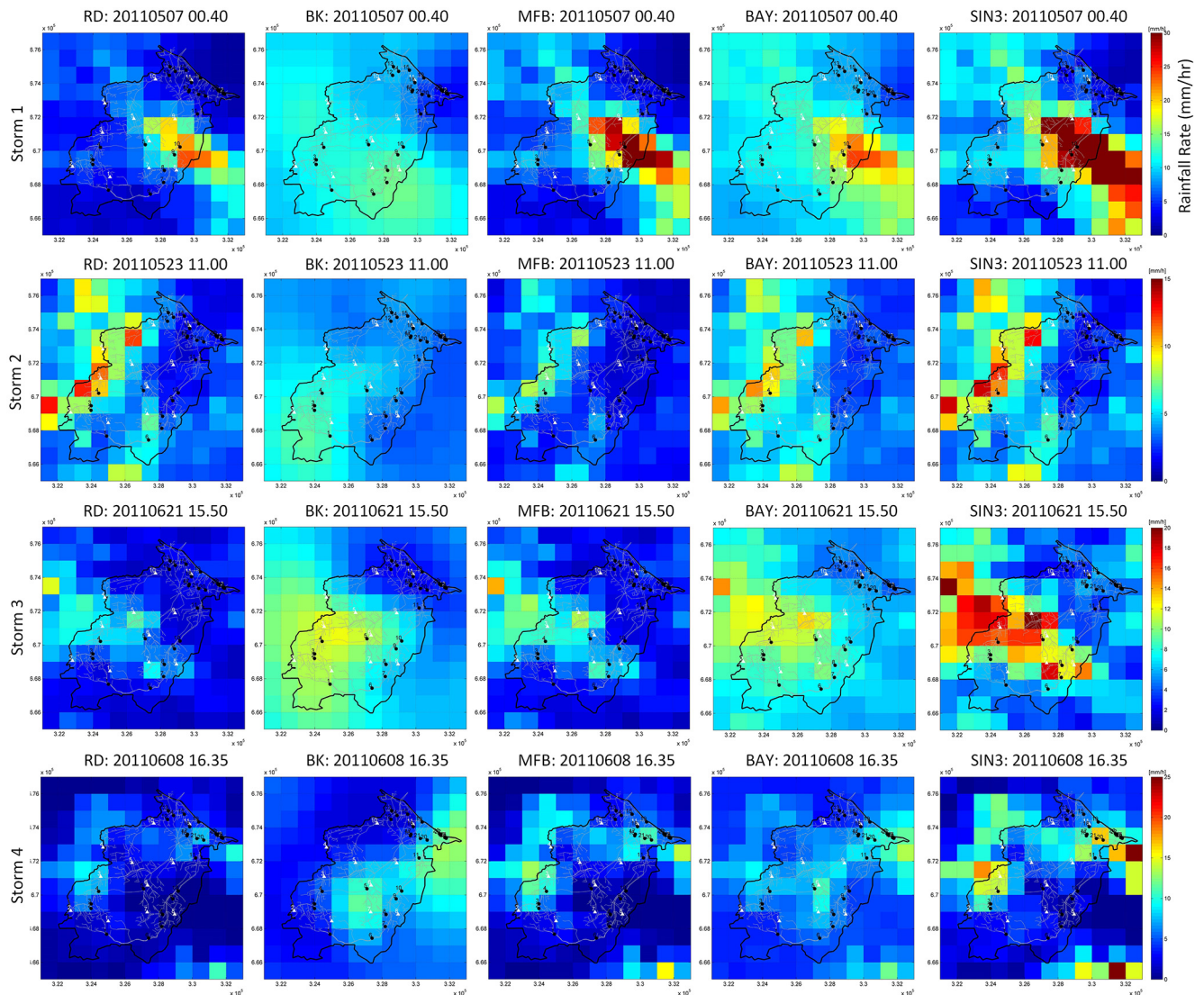
**Figure 7.** Snapshot images of the different spatial rainfall products at the time of peak areal intensity for Storms 1 (top panels) to 4 (bottom panels) over the Portobello catchment. From left to right panels: RD, BK, MFB, BAY and SIN3 (with singularity range [1, 3]) estimates. The black polygon indicates the boundary of the Portobello catchment, and the black and white markers respectively represent the location of flow and rain gauges.

tioned in the previous sections, the SIN3 range covers most of the singularity indices and consistently led to good results for the storms under consideration.

It can be seen that the spatial structure of the BK rainfall field (fully based upon rain gauge data) is highly symmetric and smooth, and is rather unrealistic. With regard to the adjusted rainfall products (MFB, BAY and SIN), it can be noticed that the proportion of radar (RD) and BK interpolated rain gauge features that are preserved varies according to the method. The MFB fields fully inherit the spatial structure of the RD fields; the only change is that the actual intensity values are scaled up or down by an areal ratio derived from the sample bias between mean rain gauge and

radar rainfall estimates. In agreement with the quantitative results presented above, it can be seen that the structure of the BAY peak rainfall fields is often similar to that of the BK ones and is smoother than the original RD image. Singularity structures are often present in rainfall fields during peak intensity periods (such as the ones shown in Fig. 7). As explained in Sect. 2, the presence of these structures causes the RD fields to be considered highly uncertain and therefore these are less trusted in the BAY merging process. This results in BAY peak intensity merged fields closer to the BK ones, instead of to RD ones. Some spatial features from RD can be still observed in the BAY fields – for example, in the lower-right area of the BAY image of Storm 1 and in the
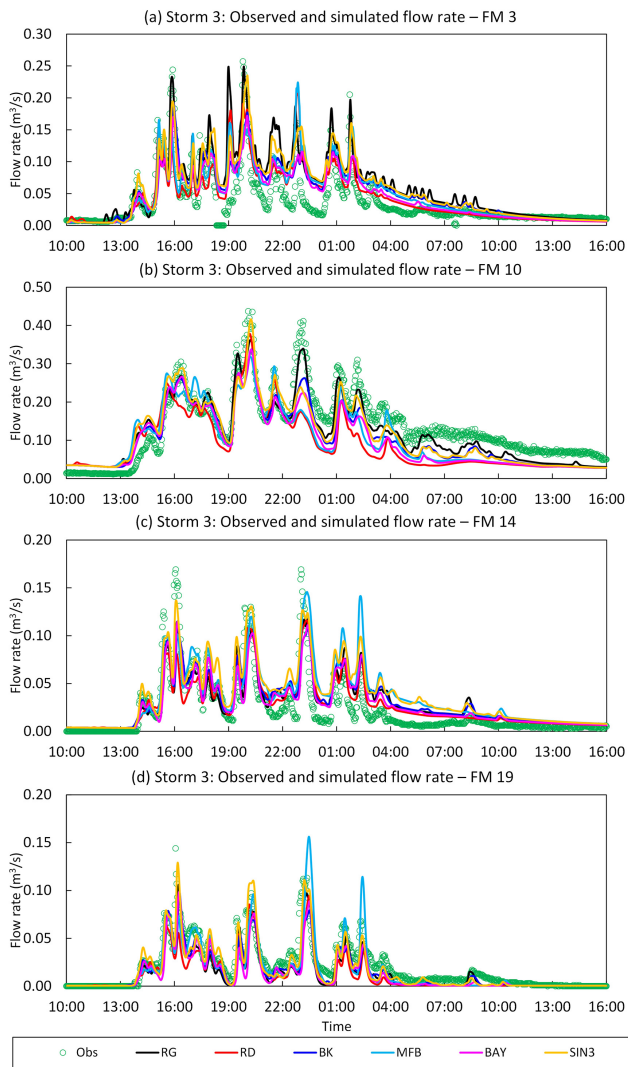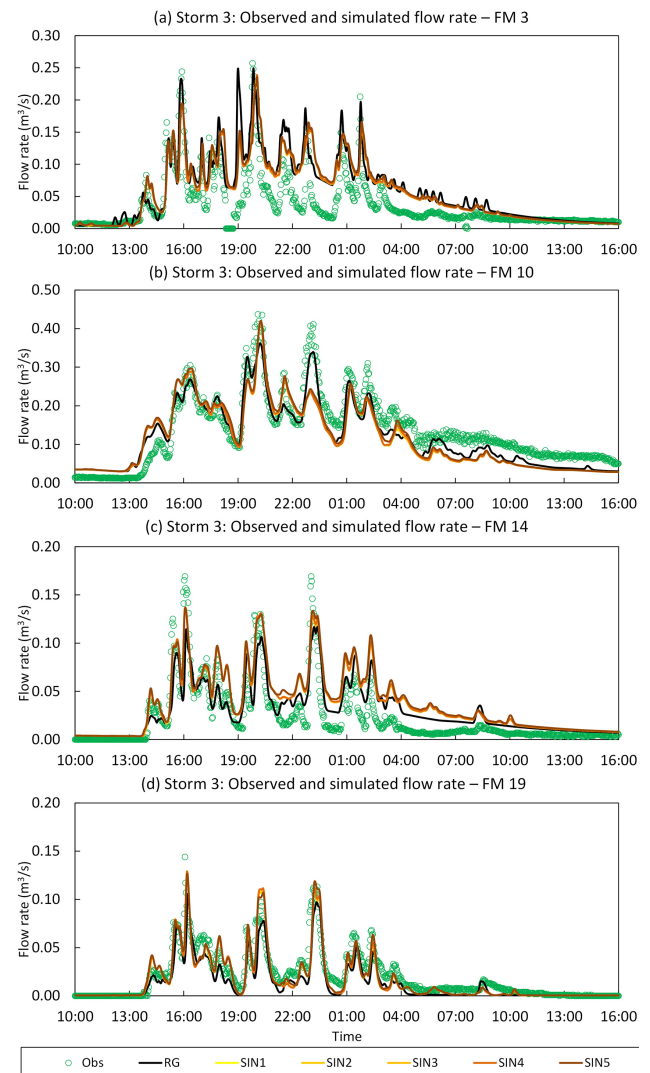
**Figure 8.** Observed flows vs. simulated flows with RG, RD, BK, BAY and SIN3 rainfall inputs at selected flow gauging sites of Portobello catchment during Storm 3. Selected gauging sites: FM3: upstream end of the catchment (top panels); FM10: mid-stream area (bottom panels). The location of the selected monitoring sites is shown in Fig. 3.

**Figure 9.** Observed flows vs. simulated flows with SIN1–SIN5 rainfall inputs at selected flow gauging sites of Portobello catchment during Storm 3. Selected gauging sites: FM14: upstream end of a small branch of the sewer system (top panels); FM19: downstream end of the catchment (bottom panels). The location of the selected monitoring sites is shown in Fig. 3.

middle-left area of the BAY image of Storm 3 (see Fig. 7, top). However, these features appear to be much smoother and spreading over a larger area, as compared to their structure in the original RD image. As compared to the BAY peak rainfall fields, the SIN fields display less smooth and more realistic structures which preserve more features of the original RD fields. In general, the inspection of the snapshot images of the different rainfall products confirms the findings of the areal and point gauge analyses regarding the ability of the SIN method to better preserve the singularity structures present in rainfall fields (and captured by RD) throughout the merging process, as compared to the original BAY merging method.

### 3.3.2 Hydraulic modelling results

Figures 8 and 9 shows example observed vs. simulated flow hydrographs for the different rainfall inputs at four gauging locations (see locations in Fig. 3) during Storm 3. Note that Storm 3 is an event with high rainfall accumulations, rainfall rates and strong singularity structures. Figure 10 summarises the performance statistics resulting from the simple linear regression analysis (i.e. $\beta$, $R^2$ and $R_w^2$) conducted at each flow gauging station for each storm event.

From the hydrographs in Figs. 8 and 9 it can be seen that, in terms of pattern and timing, all simulated flows (resulting from the different rainfall inputs) are generally in good
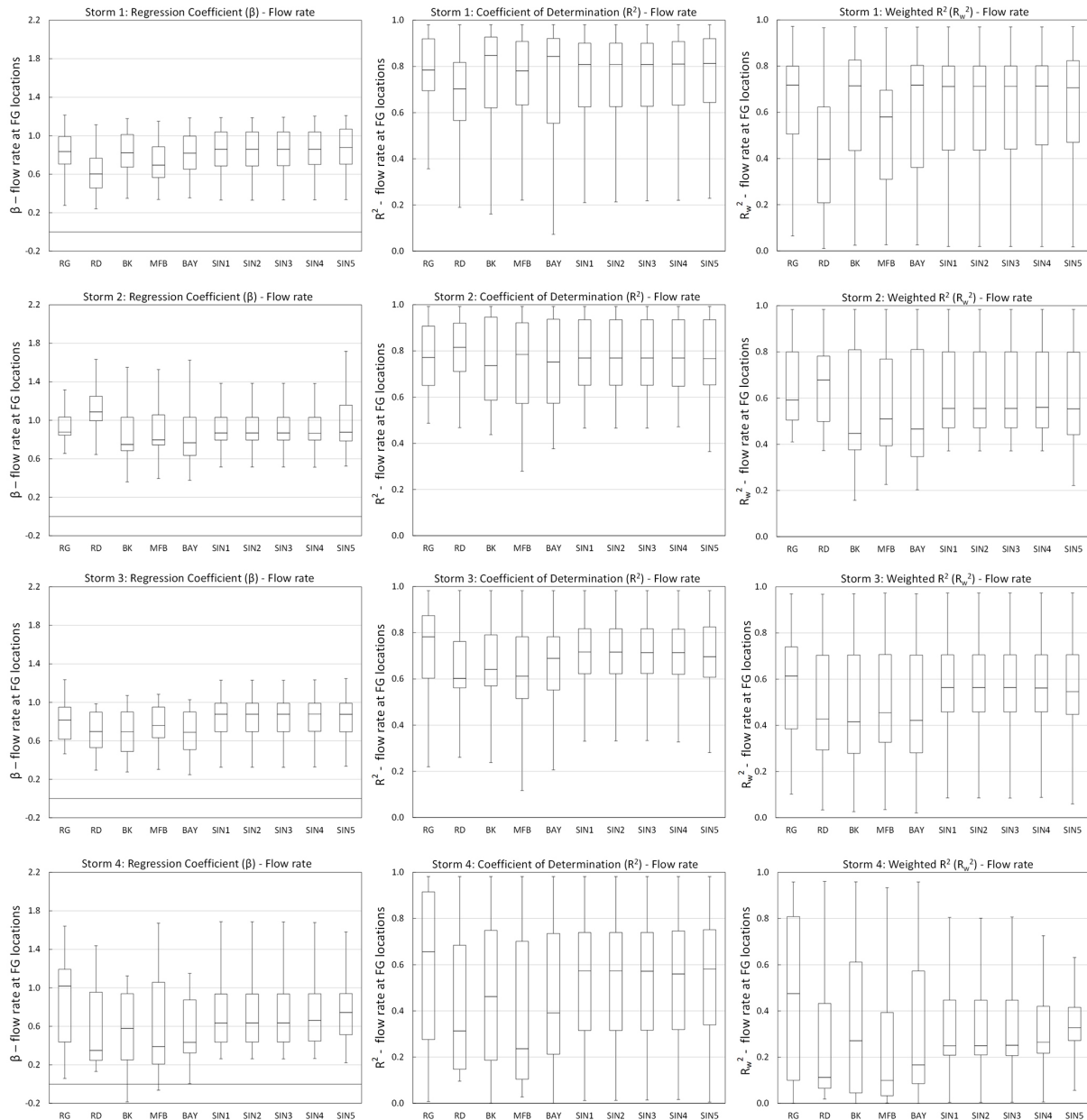
**Figure 10.** Boxplots displaying the distribution of regression coefficient ($\beta$) (left column panels), coefficient of determination ($R^2$) (middle column panels) and weighted coefficient of determination ($R_w^2$) (right column panels) statistics derived from the linear regression analysis conducted for each pair of recorded and simulated flow time series at each gauging location.

agreement with observations. This indicates that all rainfall products, including the RD before adjustment, can well capture the general dynamics of rainfall fields. The main difference between the simulated flows lies in their ability to reproduce flow peaks, which in turn is a function of the ability of the different rainfall estimates to reproduce peak rainfall rates in terms of magnitude, timing and spatial distribution. In line with the results of the rainfall analysis, the flow hydrographs associated with BK estimates are close to the ones associated with RG records; however, the former display a

smoother behaviour and generally lead to flow peaks that are lower than the recorded ones and the ones resulting from RG inputs. The RD outputs can well match some of the observed flow peaks, but significantly underestimate others (e.g. flow peak at around 23 h in FM10); this can be attributed to the underestimation of peak intensities observed in RD estimates. The MFB associated flows show little improvement over the RD ones and in many cases they even lead to a worse performance (e.g. see overestimation of peak flows at around 02:00 UTC in FM14 and FM19). This is further confirmed

by the statistics in Fig. 10 and corroborates the claim that the application of a "blanket" MFB correction over the area of interest is insufficient for urban applications. The BAY outputs, on the other hand, show a consistent improvement over the RD outputs. Nonetheless, in agreement with the results of the rainfall analysis, the BAY estimates behave similarly to the BK ones and lead to smooth flow peaks which often underestimate observations. The SIN outputs also show a consistent improvement over RD estimates, but, unlike BAY outputs, the SIN outputs do not smooth off flow peaks and instead show a better ability to reproduce these, sometimes leading to a better match of observed flow peaks than RG associated outputs (which were used as input for the calibration of the model). With regard to the singularity ranges, it can be seen that the differences observed in the SIN1–SIN5 rainfall estimates are mostly filtered out when converting rainfall to runoff. As a result, the flow outputs of all SIN estimates are very similar and their hydrographs can barely be differentiated.

The preliminary conclusions drawn from the visual inspection of the selected hydrographs are corroborated by the statistics in Fig. 10. As would be expected, the RG outputs generally show the best performance in all statistics (except for Storm 2). It is nonetheless noteworthy that the $\beta$ values for RG estimates (for which the model was calibrated) are generally below 1. This reveals a slight bias of the model to underestimate flows and partially explains the fact that $\beta$ values associated with all rainfall inputs are mostly below 1. With regard to the BK (i.e. interpolated RG) associated outputs, the tendency to underestimate, which is observed in the rainfall analysis, becomes even more evident in the hydraulic outputs: BK's $\beta$ values are significantly lower than 1 and lower than the $\beta$ values associated with RG estimates. Different from the results of the rainfall analysis, in which those products closest to RG estimates (including BK) displayed the best performance in terms of $R^2$, in the hydraulic analysis BK outputs generally lead to a deterioration in $R^2$ values (see statistics of Storms 2, 3 and 4). This suggests that the smoothing caused in the BK interpolation affects the small-scale dynamics of rainfall fields, leading to poor representation of associated flow dynamics. Contrary to RG-associated outputs, RD flow estimates generally display the worst performance. In agreement with the results of the rainfall analysis, RD outputs show a tendency to largely underestimate flows (as indicated by $\beta$ values well below one and much lower than those obtained for RG outputs). Moreover, they display relatively low $R^2$ and associated $R_w^2$ values. Nonetheless, a special case is observed in Storm 2, when RD estimates, which in the rainfall analysis displayed the poorest performance, yielded the best flow simulations, thus emphasising the added value that RD estimates can provide, as well as the complementary information provided by the hydraulic evaluation strategy. In the cases in which RD outputs perform poorly (i.e. in Storms 1, 3 and 4), all adjusted rainfall estimates lead to improvements over RD hydraulic results,

with the degree of improvement varying according to the adjustment technique. In Storm 2, when RD outputs displayed the best performance, the different merging methods showed to retain different degrees of RD features. Overall, it can be seen that the MFB estimates provide little improvement over the original RD estimates. In the cases in which RD led to systematic underestimation of flows, the MFB estimates managed to slightly reduce this underestimation by bringing $\beta$ values closer to 1, as compared to those of the original RD estimates. However, in Storm 2, in which RD outputs performed best, MFB caused a large deterioration of $\beta$ values, whereas the SIN estimates managed to keep $\beta$ scores closer to 1. In terms of $R^2$, the MFB estimates do not provide much improvement and can actually lead to a deterioration of this statistic (e.g. Storms 2 and 4), suggesting that the application of the MFB adjustment can alter the spatial-temporal structure of the original RD fields. The BAY outputs show a greater improvement than MFB, particularly in terms of $R^2$. Nonetheless, in agreement with the rainfall analysis and with the visual inspection of hydrographs, the BAY outputs behave remarkably similarly to BK ones. One of the main features of the BAY outputs is that they lead to systematically lower flows than RG estimates (note $\beta$ values consistently lower than those of RG outputs). This confirms the smoothing of rainfall peaks that occurs when second-order approximations are applied. Lastly, it can be seen that the SIN outputs display the greatest improvement over original RD outputs, both in terms of correcting systematic bias, as well as in terms of well reproducing rainfall and associated flow patterns. As compared to MFB and BAY outputs, the SIN outputs generally display $\beta$ values closer to 1, higher $R^2$ values (sometimes even higher than those of RG outputs) and consequently higher $R_w^2$ values. The better performance of SIN hydraulic outputs over BAY ones, particularly in terms of $\beta$, provides an a posteriori confirmation that it was right, in the SIN method, to view the singularities in the radar field as actual features of the real rainfall. Similarly as was observed in the hydrographs in Figs. 8 and 9, the performance of the different SIN ranges is very similar. In line with the results of the rainfall analysis, SIN5 shows a slight tendency towards higher flows, which sometimes resulted in better hydraulic statistics. Nonetheless, the differences are small and, based upon the findings of the rainfall analysis, the adoption of a wider SIN range and removal of most singularities appears to be a more conservative option. However, this aspect must be further investigated using a wider range of storm events and pilot catchments.

Regarding the difference in hydraulic performance between the events used for model calibration (i.e. Storms 1–3) and the independent event (i.e. Storm 4), it can be seen that the statistics of the hydraulic outputs during Storm 4 are generally lower than for the other three storm events (Fig. 10). This can be expected, given that the model was "tuned" to give a good fit for the calibration events. However, as discussed above, the general features and relative performance

of the hydraulic outputs associated with the different rainfall inputs was generally consistent in all storm events under consideration. The particular differences that were observed were due to the nature of a given storm and not to the fact that a given event was used for model calibration or not.

## 4  Conclusions and future work

In this paper, a new gauge-based radar rainfall adjustment method was proposed, which aims at better merging rainfall estimates obtained from rain gauges and radars, at the small spatial and temporal scales characteristic of urban catchments. The proposed method incorporates a local singularity analysis into the Bayesian merging technique (Cheng et al., 1994; Todini, 2001). Through this incorporation, the merging process preserves the fine-scale singularity (non-Gaussian) structures present in rainfall fields and captured by radar, which are often associated with local extremes and are generally smoothed off by currently available radar–rain gauge merging techniques, mainly based upon Gaussian approximations.

Using as case study four storm events observed in the Portobello catchment ($53 \, \text{km}^2$) (Edinburgh, UK) in 2011, the performance of the proposed singularity-sensitive Bayesian data-merging (SIN) method, in terms of adjusted rainfall estimates and the subsequent runoff estimates, was evaluated and compared against that of the original Bayesian data-merging (BAY) technique and the widely used mean field bias correction (MFB) method. This analysis clearly brought out the benefits of introducing the singularity-sensitive method. The results suggest that the proposed SIN method can effectively identify, extract and preserve the singular structures present in radar images while retaining the accuracy of rain gauge (RG) estimates. This is reflected in the better ability of the SIN method to reproduce instantaneous rainfall rates, rainfall accumulations and associated runoff flows. This method clearly outperforms the commonly used MFB adjustment, which simply fails to reproduce the dynamics of rainfall in urban areas, and the original BAY method, which shows an overall good performance but smooths off peak rainfall magnitudes, thus leading to underestimation of runoff extremes.

In this study the sensitivity of the SIN results to the "degree" of singularity that is removed from the radar image and preserved throughout the merging process was also tested. While the impact of it was found to be generally small, the results suggest that partially removing singularities could have a negative impact on the results. Therefore, removing most singularity exponents from the original radar image is advisable.

While the proposed singularity method has shown great potential to improve the merging of radar and rain gauge data for urban hydrological applications, further testing including more storm events and pilot catchments is still required in order to ensure that the results are not case specific and to draw more robust conclusions about the applicability of the proposed method. Other aspects on which further work is recommended are the following:

a. The current version of the singularity-sensitive method shows a slight tendency to overestimate rainfall rates and accumulations. This is likely to be due to one of two aspects, or a combination of them:

- In the eventual case in which a rain gauge is located within the core of a convective cell, the resulting interpolated (block-kriged; BK) field may end up having singularity structures and, as explained in Sect. 2.3, this may ultimately lead to "double-counting" of singularities. For the time being, a moving-window smoothing has been applied on the BK field before it is merged with the non-singular radar field, so as to remove singularity structures potentially present in the BK field. While this has proven to be an acceptable solution, we believe it can be further refined. Other methods for dealing with this particular problem have been tested, including extraction of singularities from the point rain gauge records using the singularity exponents derived from the co-located radar pixels, and extraction of singularities from the BK field using local singularity analysis. However, these have proven unstable and highly uncertain. Further work to better deal with this issue is required.

- The asymmetric distribution of singularity exponents and the numerical stability of singularity extraction from a small set of data samples. This drawback could be improved by forcing the mean of non-singular components to remain equal to the original radar estimates (Agterberg, 2012a). Alternatively, other techniques for singularity identification and extraction could be used. For example, the wavelet transformation (Kumar and Foufoula-Georgiou, 1993; Mallat and Hwang, 1992; Robertson et al., 2003; Struzik, 1999), Principal Component Analysis (Gonzalez-Audicana et al., 2004; Zheng et al., 2007) and Empirical Mode Decomposition (Nunes et al., 2003, 2005) techniques are widely recommended in the literature.

b. Given that the proposed singularity-sensitive merging method is particularly intended to improve rainfall estimates for (small-scale) urban areas, it would be interesting to test it using higher spatial-temporal resolution data (e.g. from X-band radars).

Lastly, a suggestion often made to us and therefore worth briefly discussing is to use a transformation in order to bring the distribution of the radar field closer to normality before the merging (be it with the Bayesian or other geo-statistical

method) is conducted. However, doing this would somehow miss the point of the proposed method. The key point here is that for a non-Gaussian structure, moments beyond the second order are important, as each brings new information worth preserving. To create a more "normal" field is not the purpose of the singularity extraction; instead, it is the consequence after removing singularities from the rainfall fields, which can be physically associated with abnormal energy concentration, such as "convective" cells, and which in the proposed method are set aside to ensure their preservation throughout the merging process.

**The Supplement related to this article is available online at doi:10.5194/hess-19-4001-2015-supplement.**

# References

Agterberg, F. P.: Mixtures of multiplicative cascade models in geochemistry, Nonlin. Processes Geophys., 14, 201–209, doi:10.5194/npg-14-201-2007, 2007.

Agterberg, F. P.: Multifractals and geostatistics, J. Geochem. Explor., 122, 113–122, doi:10.1016/j.gexplo.2012.04.001, 2012a.

Agterberg, F. P.: Sampling and analysis of chemical element concentration distribution in rock units and orebodies, Nonlin. Processes Geophys., 19, 23–44, doi:10.5194/npg-19-23-2012, 2012b.

Anagnostou, E. N. and Krajewski, W. F.: Real-time radar rainfall estimation, Part I: Algorithm formulation, J. Atmos. Ocean. Tech., 16, 189–197, 1999.

Bell, F. C.: The Areal Reduction Factor in Rainfall Frequency Estimation, Centre for Ecology & Hydrology (CEH), Wallingford, 1976.

Bowler, N. E., Pierce, C. E., and Seed, A. W.: STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Q. J. Roy Meteorol. Soc., 132, 2127–2155, doi:10.1256/qj.04.100, 2006.

Brandes, E. A., Ryzhkov, A. V and Zrnić, D. S.: An evaluation of radar rainfall estimates from specific differential phase, J. Atmos. Ocean. Tech., 18, 363–375, doi:10.1175/1520-0426(2001)018<0363:AEORRE>2.0.CO;2, 2001.

Chen, Z., Cheng, Q., Chen, J., and Xie, S.: A novel iterative approach for mapping local singularities from geochemical data, Nonlin. Processes Geophys., 14, 317–324, doi:10.5194/npg-14-317-2007, 2007.

Cheng, Q.: Multifractality and spatial statistics, Comput. Geosci., 25, 949–961, doi:10.1016/S0098-3004(99)00060-6, 1999.

Cheng, Q. and Zhao, P.: Singularity theories and methods for characterizing mineralization processes and mapping geo-anomalies for mineral deposit prediction, Geosci. Front., 2, 67–79, doi:10.1016/j.gsf.2010.12.003, 2011.

Cheng, Q., Agterberg, F. P., and Ballantyne, S. B.: The separation of geochemical anomalies from background by fractal methods, J. Geochem. Explor., 51, 109–130, doi:10.1016/0375-6742(94)90013-2, 1994.

Ciach, G. J.: Local random errors in tipping-bucket rain gauge measurements, J. Atmos. Ocean. Tech., 20, 752–759, doi:10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2, 2003.

Collier, C. G.: Accuracy of rainfall estimates by radar, part I: Calibration by telemetering raingauges, J. Hydrol., 83, 207–223, doi:10.1016/0022-1694(86)90152-6, 1986.

Collier, C. G.: Applications of Weather Radar Systems, 2nd Edn., Wiley, Chichester, England, 1996.

Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T. D., Rauch, W., Bertrand-Krajewski, J. L., and Tait, S.: Assessing uncertainties in urban drainage models, Phys. Chem. Earth Pt. A/B/C, 42–44, 3–10, doi:10.1016/j.pce.2011.04.007, 2012.

Einfalt, T. and Michaelides, S.: Quality control of precipitation data, in: Precipitation: Advances in Measurement, Estimation and Prediction SE-5, edited by: Michaelides, S., Springer, Berlin, Heidelberg, 101–126, 2008.

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G., and Vieux, B.: Towards a roadmap for use of radar rainfall data in urban drainage, J. Hydrol., 299, 186–202, doi:10.1016/j.jhydrol.2004.08.004, 2004.

Einfalt, T., Jessen, M., and Mehlig, B.: Comparison of radar and raingauge measurements during heavy rainfall, Water Sci. Technol., 51, 195–201, 2005.

Evertsz, C. J. G. and Mandelbrot, B. B.: Multifractal measures, in: Chaos and Fractals, edited by: Peitgen, H.-O., Jurgens, H., and Saupe, D., Springer, New York, 922–953, 1992.

Fulton, R. A., Breidenbach, J. P., Dong-Jun, S., and Miller, D. A.: The WSR-88D rainfall algorithm, Weather Forecast., 13, 377–395, doi:10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2, 1998.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, Q. J. Roy Meteorol. Soc., 132, 1669–1692, doi:10.1256/qj.05.190, 2006.

Gires, A., Onof, C., Maksimović, C., Schertzer, D., Tchiguirinskaia, I., and Simões, N.: Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multi-

fractal downscaling: a case study, J. Hydrol., 442–443, 117–128, doi:10.1016/j.jhydrol.2012.04.005, 2012.

Golding, B. W.: Nimrod: a system for generating automated very short range forecasts, Meteorol. Appl., 5, 1–16, doi:10.1017/S1350482798000577, 1998.

Gonzalez-Audicana, M., Saleta, J. L., Catalan, R. G., Garcia, R.: Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition, IEEE T. Geosci. Remote, 42, 1291–1299, doi:10.1109/TGRS.2004.825593, 2004.

Gooch, M. N.: Use of rainfall data from flow surveys, WaPUG User Note No. 6, Chartered Institute of Water and Environmental Management (CIWEM), London, 2009.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.

Guo, Y. and Adams, B. J.: Hydrologic analysis of urban catchments with event-based probabilistic models: 1. Runoff volume, Water Resour. Res., 34, 3421–3431, doi:10.1029/98WR02449, 1998.

Habib, E., Meselhe, E. and Aduvala, A.: Effect of local errors of tipping-bucket rain gauges on rainfall–runoff simulations, J. Hydrol. Eng., 13, 488–496, doi:10.1061/(ASCE)1084-0699(2008)13:6(488), 2008.

Harrison, D. L., Driscoll, S. J., and Kitchen, M.: Improving precipitation estimates from weather radar using quality control and correction techniques, Meteorol. Appl., 7, 135–144, doi:10.1017/S1350482700001468, 2000.

Harrison, D. L., Scovell, R. W., and Kitchen, M.: High-resolution precipitation estimates for hydrological uses, P. I. Civil Eng-Wat. M., 162, 125–135, doi:10.1680/wama.2009.162.2.125, 2009.

HR Wallingford: Wallingford Procedure for Design and Analysis of Urban Storm Drainage, Wallingford, UK, 1983.

Kalman, R. E.: A new approach to linear filtering and prediction problems, J. Basic Eng.-T. ASME, 82, 35–45, doi:10.1115/1.3662552, 1960.

Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, Water Resour. Res., 42, W03407, doi:10.1029/2005WR004368, 2006.

Krajewski, W. F.: Cokriging radar-rainfall and rain gage data, J. Geophys. Res., 92, 9571–9580, doi:10.1029/JD092iD08p09571, 1987.

Krajewski, W. F. and Smith, J. A.: Radar hydrology: rainfall estimation, Adv. Water Resour., 25, 1387–1394, doi:10.1016/S0309-1708(02)00062-3, 2002.

Krämer, S., Fuchs, L., and Verworn, H.: Aspects of radar rainfall forecasts and their effectiveness for real time control-the example of the sewer system of the City of Vienna, Water Pract. Technol., 2, 42–49, 2007.

Krause, P., Boyle, D. P., and Bäse, F.: Comparison of different efficiency criteria for hydrological model assessment, Adv. Geosci., 5, 89–97, doi:10.5194/adgeo-5-89-2005, 2005.

Kumar, P. and Foufoula-Georgiou, E.: A multicomponent decomposition of spatial rainfall fields: 1. Segregation of large- and small-scale features using wavelet transforms, Water Resour. Res., 29, 2515–2532, doi:10.1029/93WR00548, 1993.

La Barbera, P., Lanza, L. G., and Stagi, L.: Tipping bucket mechanical errors and their influence on rainfall statistics and extremes, Water Sci. Technol., 45, 1–9, 2002.

Liguori, S., Rico-Ramirez, M. A., Schellart, A. N. A., and Saul, A. J.: Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments, Atmos. Res., 103, 80–95, doi:10.1016/j.atmosres.2011.05.004, 2011.

Lovejoy, S. and Mandelbrot, B. B.: Fractal properties of rain, and a fractal model, Tellus A, 37, 209–232, doi:10.1111/j.1600-0870.1985.tb00423.x1985.

Luyckx, G. and Berlamont, J.: Simplified method to correct rainfall measurements from tipping bucket rain gauges, in: Urban Drainage Modeling: Proceedings of the Speciality Symposium of the World Water and Environmental Resource Congress, edited by: Brashear, R. W. and Maksimović, Č., American Society of Civil Engineers, Washington, D.C., 767–776, 2001.

Mallat, S. and Hwang, W.-L.: Singularity detection and processing with wavelets, IEEE T. Inform. Theory, 38, 617–643, doi:10.1109/18.119727, 1992.

Marshall, J. and Palmer, W.: The distribution of raindrops with size, J. Meteorol., 5, 165–166, doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2, 1948.

Marshall, M. and McIntyre, N.: Field verification of bed-mounted ADV meters, Proc. ICE – Water Manage., 161, 199–206, doi:10.1680/wama.2008.161.4.199, 2008.

Mazzetti, C.: Data interpolation and multi-sensors Bayesian combinations, RainMusic User's Manual and References, PROGEA srl, Bologna, Italy, 2012.

Mazzetti, C. and Todini, E.: Combining raingauges and radar precipitation measurements using a Bayesian approach, in: geoENV IV – Geostatistics for Environmental Applications, edited by: Sanchez-Vila, X., Carrera, J., and Gómez-Hernández, J. J., Kluwer Academic Publishers, Springer, the Netherlands, 401–412, 2004.

Molini, A., Lanza, L. G., and La Barbera, P.: The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications, Hydrol. Process., 19, 1073–1088, doi:10.1002/hyp.5646, 2005.

Nunes, J., Bouaoune, Y., Delechelle, E., Niang, O., and Bunel, P.: Image analysis by bidimensional empirical mode decomposition, Image Vis. Comput., 21, 1019–1026, doi:10.1016/S0262-8856(03)00094-5, 2003.

Nunes, J., Guyot, S., and Deléchelle, E.: Texture analysis based on local analysis of the bidimensional empirical mode decomposition, Mach. Vis. Appl., 16, 177–188, doi:10.1007/s00138-004-0170-5, 2005.

Ochoa-Rodríguez, S., Wang, L.-P., Grist, A., Allitt, R., Onof, C., and Maksimović, Č.: Improving the applicability of radar rainfall estimates for urban pluvial flood modelling and forecasting, in: Urban Drainage Group Autumn Conference and Exhibition 2013: Future Thinking and Challenges, 13–15 November 2013, Nottingham, UK, 19, 2013.

Osborne, M. P.: A New Runoff Volume Model, Wastewater Planning Users Group WaPUG, Chartered Institute of Water and Environmental Management (CIWEM), UK, 2001.

Robertson, A. N., Farrar, C. R., and Sohn, H.: Singularity detection for structural health monitoring using Hölder exponents, Mech. Syst. Signal Pr., 17, 1163–1184, doi:10.1006/mssp.2002.1569, 2003.

Schellart, A. N. A., Shepherd, W. J., and Saul, A. J.: Influence of rainfall estimation error and spatial variability on sewer flow prediction at a small urban scale, Adv. Water Resour., 45, 65–75, doi:10.1016/j.advwatres.2011.10.012, 2012.

Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res., 92, 9693–9714, doi:10.1029/JD092iD08p09693, 1987.

Schertzer, D., Tchiguirinskaia, I., and Lovejoy, S.: Multifractality: at least three moments!, Interactive comment on "Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrolog" by F. Lombardo et al., Hydrol. Earth Syst. Sci. Discuss., 10, C3103–C3109, 2013.

Seo, D. and Smith, J.: Rainfall estimation using raingages and radar – a Bayesian approach: 1. Derivation of estimators, Stoch. Hydrol. Hydraul., 5, 17–29, doi:10.1007/BF01544175, 1991.

Sevruk, B. and Nešpor, V.: Empirical and theoretical assessment of the wind induced error of rain measurement, Water Sci. Technol., 37, 171–178, doi:10.1016/S0273-1223(98)00330-8, 1998.

Sinclair, S. and Pegram, G.: Combining radar and rain gauge rainfall estimates using conditional merging, Atmos. Sci. Lett., 6, 19–22, doi:10.1002/asl.85, 2005.

Smith, J. A., Baeck, M. L., Meierdiercks, K. L., Miller, A. J., and Krajewski, W. F.: Radar rainfall estimation for flash flood forecasting in small urban watersheds, Adv. Water Resour., 30, 2087–2097, doi:10.1016/j.advwatres.2006.09.007, 2007.

Smith, J. A., Hui, E., Steiner, M., Baeck, M. L., Krajewski, W. F., and Ntelekos, A. A.: Variability of rainfall rate and raindrop size distributions in heavy rain, Water Resour. Res., 45, W04430, doi:10.1029/2008WR006840, 2009.

Struzik, Z. R.: Local effective Hölder exponent estimation on the wavelet transform maxima tree, in: Fractals: Theory and Applications in Engineering, edited by: Dekking, M., Véhel, J. L., Lutton, E., and Tricot, C., Springer, London, UK, 93–112, 1999.

Tchiguirinskaia, I., Schertzer, D., Hoang, C. T., and Lovejoy, S.: Multifractal study of three storms with different dynamics over the Paris region, in: Weather Radar and Hydrology, edited by: Moore, R. J., Cole, S. J., and Illingworth, A. J., IAHS, International Association of Hydrological Sciences (IAHS), Exeter, UK, 421–426, 2011.

Thorndahl, S., Nielsen, J. E., and Rasmussen, M. R.: Bias adjustment and advection interpolation of long-term high resolution radar rainfall series, J. Hydrol., 508, 214–226, doi:10.1016/j.jhydrol.2013.10.056, 2014.

Todini, E.: A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, Hydrol. Earth Syst. Sci., 5, 187–199, doi:10.5194/hess-5-187-2001, 2001.

Ulbrich, C. W.: Natural variations in the analytical form of the raindrop size distribution, J. Clim. Appl. Meteorol., 22, 1764–1775, doi:10.1175/1520-0450(1983)022<1764:NVITAF>2.0.CO;2, 1983.

Velasco-Forero, C. A., Sempere-Torres, D., Cassiraga, E. F., and Gómez-Hernández, J.: A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data, Adv. Water Resour., 32, 986–1002, doi:10.1016/j.advwatres.2008.10.004, 2009.

Vieux, B. E. and Bedient, P. B.: Assessing urban hydrologic prediction accuracy through event reconstruction, J. Hydrol., 299, 217–236, doi:10.1016/j.jhydrol.2004.08.005, 2004.

Villarini, G., Smith, J. A., Lynn Baeck, M., Sturdevant-Rees, P., and Krajewski, W. F.: Radar analyses of extreme rainfall and flooding in urban drainage basins, J. Hydrol., 381, 266–286, doi:10.1016/j.jhydrol.2009.11.048, 2010.

Wackernagel, H.: Multivariate Geostatistics, An Introduction with Applications, Springer, Berlin, 2003.

Wang, L.-P. and Onof, C.: High-resolution rainfall field reconstruction based upon Kriging and local singularity analysis, in: Hydrofractals '13, HF-10, 17–19 October 2013, Kos Island, Greece, 2013.

Wang, L.-P., Onof, C., Ochoa-Rodríguez, S., and Simões, N.: Analysis of kriged rainfields using multifractals, in: 9th International Workshop on Precipitation in Urban Areas: Urban Challenges in Rainfall Analysis, 6–9 December 2012, St. Moritz, Switzerland, 138–142, 2012.

Wang, L.-P., Ochoa-Rodríguez, S., Simões, N. E., Onof, C., and Maksimović, Č.: Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology, Water Sci. Technol., 68, 737–747, doi:10.2166/wst.2013.300, 2013.

Wang, L.-P., Ochoa-Rodríguez, S., Willems, P., and Onof, C.: Improving the applicability of gauge-based radar rainfall adjustment methods to urban pluvial flood modelling and forecasting using local singularity analysis, in: International Symposium on Weather Radar and Hydrology (WRaH), 7–9 April 2014, Washington, D.C., 10 pp., 2014.

WaPUG: Code of Practice for the Hydraulic Modelling of Sewer Systems, Wastewater Planning Users Group WaPUG, Chartered Institute of Water and Environmental Management (CIWEM), UK, 2002.

Zheng, Y., Hou, X., Bian, T., and Qin, Z.: Effective image fusion rules Of multi-scale image decomposition. in: 5th International Symposium on Image and Signal Processing and Analysis, 27–29 September 2007, Istanbul, Turkey, 362–366, doi:10.1109/ISPA.2007.4383720, 2007.
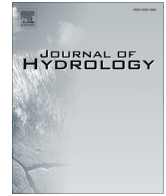
# Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment

Li-Pen Wang [a,*], Susana Ochoa-Rodríguez [b], Johan Van Assel [c], Rui Daniel Pina [b], Mieke Pessemier [c], Stefan Kroll [c], Patrick Willems [a], Christian Onof [b]

[a] Hydraulics Laboratory, KU Leuven, 3001 Leuven, Belgium
[b] Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK
[c] Aquafin NV, Dijkstraat 8, 2630 Aartselaar, Belgium

## ARTICLE INFO

## SUMMARY

Rainfall estimates of the highest possible accuracy and resolution are required for urban hydrological applications, given the small size and fast response which characterise urban catchments. While radar rainfall estimates have the advantage of well capturing the spatial structure of rainfall fields and its variation in time, the commonly available radar rainfall products (typically at ~1 km/5–10 min resolution) may still fail to satisfy the accuracy and resolution – in particular temporal resolution – requirements of urban hydrology. A methodology is proposed in this paper, to produce higher temporal resolution, more accurate radar rainfall estimates, suitable for urban hydrological applications. The proposed methodology entails two main steps: (1) Temporal interpolation of radar images from the originally-available temporal resolutions (i.e. 5–10 min) to finer resolutions at which local rain gauge data are commonly available (i.e. 1–2 min). This is done using a novel interpolation technique, based upon the multi-scale variational optical flow technique, and which can well capture the small-scale rainfall structures relevant at urban scales. (2) Local and dynamic gauge-based adjustment of the higher temporal resolution radar rainfall estimates is performed afterwards, by means of the Bayesian data merging method. The proposed methodology is tested using as case study a total of 8 storm events observed in the Cranbrook (UK) and Herent (BE) urban catchments, for which radar rainfall estimates, local rain gauge and depth/flow records, as well as recently calibrated urban drainage models were available. The results suggest that the proposed methodology can provide significantly improved radar rainfall estimates and thereby generate more accurate runoff simulations at urban scales, over and above the benefits derived from the mere application of Bayesian merging at the original temporal resolution at which radar estimates are available. The benefits of the proposed temporal interpolation + merging methodology are particularly evident in storm events with strong and fast-changing (convective-like) rain cells.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Rainfall estimates of the highest possible accuracy and resolution are required for urban hydrological applications, given the small size and fast response which characterise urban catchments (Berne et al., 2004; Collier, 2009; Fabry et al., 1994; Liguori et al., 2012; Ochoa-Rodríguez et al., in press). Due to their ability to well capture the spatial characteristics of rainfall fields and their evolution in time, radar rainfall estimates are playing an increasingly important role in urban hydrological applications (Krajewski and Smith, 2002; Krämer et al., 2007; Schellart et al., 2012; Villarini et al., 2010; Wang et al., 2011a). However, the operational radar rainfall products provided by national weather services (typically at ~1 km/5–10 min resolution) may still fail to meet the demanding requirements of urban hydrology, both in terms of accuracy and resolution.

As regards accuracy, since radar quantitative precipitation estimates (QPEs) are an indirect measurement of rainfall, they are subject to multiple sources of error. Firstly, radar reflectivity

* Corresponding author. Tel.: +32 16322236.
E-mail addresses: Lipen.Wang@bwk.kuleuven.be (L.-P. Wang), s.ochoa-rodriguez@imperial.ac.uk (S. Ochoa-Rodríguez), johan.vanassel@aquafin.be (J. Van Assel), r.pina13@imperial.ac.uk (R.D. Pina), mieke.pessemier@aquafin.be (M. Pessemier), stefan.kroll@aquafin.be (S. Kroll), Patrick.Willems@bwk.kuleuven.be (P. Willems), c.onof@imperial.ac.uk (C. Onof).

measurements, from which QPEs are subsequently derived, may be affected by factors such as radar beam blockage, attenuation, ground clutter and anomalous propagation of the signal (Collier, 1996; Einfalt et al., 2004; Harrison et al., 2000). A number of corrections are usually applied in order to reduce errors arising from these sources; however, it is virtually impossible to have error-free reflectivity measurements. Additional errors arise in the conversion of reflectivity measurements ($Z$) to rainfall rates ($R$), which is usually done using the $Z–R$ relationship, $Z = aR^b$ (Marshall and Palmer, 1948). The variables $a$ and $b$ can be theoretically linked to rain drop size distribution and are generally deduced by physical approximation or empirical calibration based upon long-term comparisons (Collier, 1986; Krajewski and Smith, 2002). However, the highly dynamic nature of rain drop size distribution, even within a single storm event (Smith et al., 2009; Ulbrich, 1983), renders the use of a static $Z–R$ relationship – as used for single-polarisation radars – imperfect, especially when extreme rainfall rates are observed (Einfalt et al., 2005; Goudenhoofdt and Delobbe, 2013). In the case of dual-polarisation radars (dual-pol radars hereafter), dual-polarisation parameters provide useful information which enables significant reduction of some of the sources of error mentioned above (especially detection of non-weather echoes and improved attenuation correction), as well as a dynamic adjustment of the $Z–R$ relationship according to drop-size distribution; this results in much more accurate rainfall rate estimates (Bringi and Chandrasekar, 2001). Dual-pol radars are being deployed in a number of countries around the world (e.g. Sugier and Tabary, 2006; Chandrasekar et al., 2009; Bringi et al., 2011; Wang et al., 2011b; Kim et al., 2012; Vasiloff, 2012; Berkowitz et al., 2013), and it is likely that in the near future most existing single-pol radars will be upgraded to dual-pol. However, despite the advantages of this new technology and the better QPEs that can be achieved with it, dual polarisation will not change the limitations inherent to radar, such as the fact that rainfall is measured indirectly, often well above the ground and often far away from the radar, which results in beam broadening and range degradation.

The uncertainties in QPEs propagate through hydrological and hydraulic models and their effect is particularly evident at the small scales of urban catchments (Collier, 2009; Schellart et al., 2012; Vieux and Bedient, 2004). Gauge-based adjustment of radar QPEs has proven effective to reduce these errors and improve the accuracy of the estimates, thus improving their applicability for hydrological applications (Harrison et al., 2009). However, most gauge-based adjustment methods have been tested and applied at large spatial and temporal scales (Anagnostou and Krajewski, 1999; Cole and Moore, 2008; Fulton et al., 1998; Germann et al., 2009; Gerstner and Heinemann, 2008; Goudenhoofdt and Delobbe, 2009; Harrison et al., 2009; Seo and Smith, 1991; Todini, 2001). Relatively few tests have been conducted at urban/small scales and all of them have concluded that at these scales more dynamic and localised adjustments are required (Borup et al., submitted; Sinclair and Pegram, 2005; Wang et al., 2013b). In fact, at urban scales, commonly used coarse-scale methods such as Mean Field Bias (MFB) correction have proven to be insufficient, while other more dynamic and higher (statistical-) order methods (e.g. geostatistical methods) have exhibited a better ability to reproduce fine-scale rainfall structures and dynamics (Wang et al., 2015).

With regards to the resolution of radar QPEs, recent studies suggest that the currently commonly available resolutions (i.e. $\sim$1 km/5–10 min) may be insufficient for urban-scale applications. In fact, the effect of insufficient spatial–temporal information of rainfall inputs on urban hydrological simulations may be as significant as that caused by insufficient accuracy. This is especially the case when the drainage area of interest is small (Gires et al., 2014, 2012; Schellart et al., 2012; Wang et al., 2012). Therefore, the impact of rainfall data resolution should not be ignored in urban hydrology. Although the spatial and temporal resolution of rainfall inputs are strongly related, a number of studies have suggested that the latter generally constitutes a more critical factor than the former (Ochoa-Rodríguez et al., in press; Singh, 1997; Thorndahl et al., 2014) and that temporal resolutions of $\sim$1–2 min (i.e. below those currently available) are required for urban hydrological applications, while spatial resolutions of $\sim$1 km (i.e. close to those currently available) appear to be sufficient. The predominant effect of temporal resolution, as well as the above mentioned resolution requirements for urban hydrological applications are illustrated in Fig. 1 (adapted from Ochoa-Rodríguez et al., in press). Moreover, it is often the case that local rain gauge (RG) records are available at temporal resolutions finer than those of radar QPEs. However, in order to perform local gauge-based adjustments, the RG records are usually aggregated to the temporal resolution of radar QPEs, thus losing valuable information. In fact, for small-scale applications, recent studies suggest that performing gauge-based adjustment of radar QPEs at shorter time intervals leads to better results than doing so at longer intervals (Borup et al., submitted; Thorndahl et al., 2014). Traditional strategies for obtaining higher temporal resolution radar QPEs include changes in radar scanning and sampling strategies (Delobbe et al., 2008; Gill et al., 2006; Sadjadi, 2000; Tabary, 2007; Zhang et al., 2005) and stochastic downscaling (Deidda, 2000; Gires et al., 2012; Gupta and Waymire, 1993; Koutsoyiannis and Onof, 2001; Marsan et al., 1996; Pegram and Clothier, 2001; Segond et al., 2006; Tessier et al., 1993; Wang et al., 2010). The former is not always possible, due to hardware and operational limitations. With regards to the stochastic temporal downscaling, albeit applications exist that meet the high temporal resolution requirements mentioned above ($\sim$1–2 min; e.g. Gires et al., 2014), they result in large ensembles, which are difficult to use operationally, given the runtimes associated to urban hydrodynamic models (Leandro et al., 2014). More recently, an advection based temporal interpolation method, combined with MFB correction, has been proposed by (Nielsen et al., 2014; Thorndahl et al., 2014). Although this method has shown to improve the performance of hourly and daily radar QPEs (assessed through comparison against ground rain gauge measurements), its performance at sub hourly scales was inconsistent. The unsatisfactory performance at smaller scales may be explained by the way in which storm movement is estimated in advection-based techniques, as well as by the MFB-based adjustment techniques that were employed, which may be insufficient to well capture and preserve the small rainfall structures relevant at urban scales (Gao et al., 1999; Germann and Zawadzki, 2002; Van Horne, 2003; Laroche and Zawadzki, 1995; Rinehart and Garvey, 1978; Wang et al., 2013b; Weickert and Schnörr, 2001; Wilson et al., 2004).

In this paper a methodology is proposed for producing accurate radar rainfall estimates with high temporal resolution, suitable for urban hydrological applications. Similar to the method proposed by Nielsen et al. (2014), the procedure proposed herein entails two main steps: (1) temporal interpolation of radar images, followed by (2) gauge-based adjustment of radar QPEs at short time intervals (1–2 min). Nonetheless, different from Nielsen's method, the techniques employed in this study to carry out the aforementioned steps are particularly well suited to capture and reproduce small-scale rainfall structures, thus making the proposed method more appropriate for urban hydrological applications. For the first step, a novel temporal interpolation technique, based upon the multi-scale variational optical flow technique, is proposed to generate high temporal resolution (i.e. 1–2 min) radar
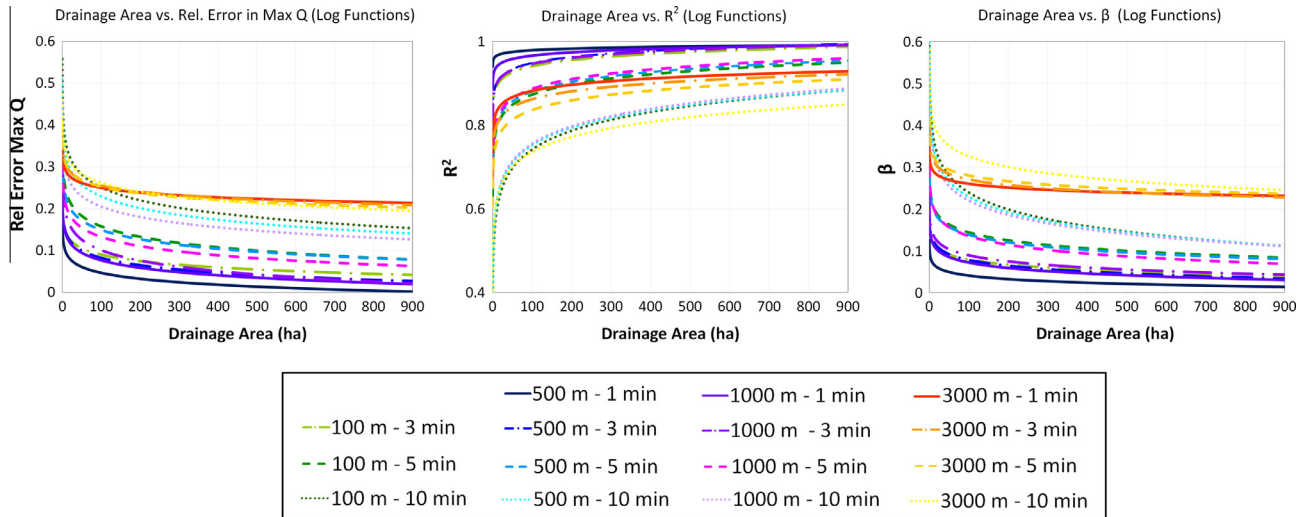
**Fig. 1.** Errors in urban runoff estimates in relation to drainage area, as a result of the coarsening of the spatial and temporal resolution of rainfall inputs. Logarithmic functions were fitted to the hydrodynamic simulation results obtained for 7 urban catchments in North-West Europe, forced with data from 9 storm events recorded at high resolution (100 m/1 min) by a polarimetric X-band radar. Coarser resolution rainfall estimates were generated by averaging in space, whereas coarser resolution temporal estimates were generated by sampling radar images at the desired temporal resolution, thus replicating radar scanning strategies. Performance statistics were estimated using as reference the runoff estimates associated to the highest available resolution rainfall estimates (adapted from Ochoa-Rodríguez et al., in press).

rainfall estimates. The optical flow technique was initially developed to characterise detailed image motion and has also been employed for short-term rainfall forecasting. In both realms it has shown to outperform many other techniques. Considering the nature of this technique, it is expected to be suitable for carrying out temporal interpolation of radar images, especially when small-scale rainfall structures are critical. For the second step, the Bayesian data merging technique is used (Todini, 2001); this has shown to outperform many other merging techniques, both at large and small (urban) scales (Mazzetti and Todini, 2004; Wang et al., 2013b), is applied locally and dynamically.

The proposed methodology is tested using as case study a total of eight storm events observed between 2012 and 2014 in the Cranbrook (London, UK) and Herent (Leuven, Belgium) urban catchments. For these pilot locations radar estimates, local rain gauge measurements, flow records, as well as recently calibrated urban drainage models were available.

The paper is organised as follows. In Section 2 a detailed explanation is provided of the temporal interpolation + gauge-based adjustment methodology. In Section 3 we present the case study, including a description of the two study catchments and available dataset, the strategy and performance criteria used to evaluate the proposed methodology, and the results of the testing. Lastly, in Section 4 the main conclusions are presented and future work is discussed.

## 2. Methodology

### 2.1. Overview

A schematic of the proposed methodology for producing accurate radar rainfall estimates with high temporal resolution is shown in Fig. 2. Let $RD_t$ and $RG_t$, respectively, be radar and rain gauge rainfall records at a specific time step $t$, but each with different temporal resolutions: $\Delta T$ for radar images and $\Delta t$ for point rain gauge records, where $\Delta T = n \cdot \Delta t$ and $n \geqslant 1$, as is often the case in urban catchments (i.e. the temporal resolution of local rain gauge records is finer than that at which radar rainfall estimates are available). The first step of the proposed method is to derive a field of movement vectors **W** between two successive radar images (i.e. $RD_t$ and $RD_{t+\Delta T}$) based upon the *optical flow technique*. Assuming

that this movement field remains constant within each time interval ($\Delta T$), intermediate radar images ($RD'$) with temporal resolution $\Delta t$ are derived through *temporal interpolation with occlusion reasoning*. Afterwards, the original RD and the intermediate (interpolated) $RD'$ images are *dynamically merged* with the coincidental RG records at time steps $\Delta t$, using the *Bayesian data merging* method. This yields high temporal-resolution radar rainfall estimates with better accuracy ($RD''$ with temporal resolution $\Delta t$). Note that for the first two steps (i.e. movement field estimation and temporal interpolation), radar images over a large domain (>100 km × 100 km) are used, while for the last step (i.e. dynamic data merging) only a sub-domain of the radar images is used (which coincides with the urban area of interest and with the coverage of the local rain gauge network).

The key techniques that are employed in the proposed methodology are: (1) optical flow estimation, (2) temporal interpolation with occlusion reasoning, and (3) Bayesian data merging. Each of these techniques is described in the following sub-sections.

### 2.2. Optical flow estimation

Optical flow estimation is a technique used to characterise image motion, which has been widely applied in the field of computer vision, e.g. object detection and video compression (Ehrhardt et al., 2007; Herbst et al., 2009; Larsen et al., 1997; Sadek et al., 2012). Due to the nature of the technique, it is in particular suitable for handling consecutive images with short time intervals (e.g. video frames). Consequently, it has also been applied to meteorological data processing and proven to be successful in short-term rainfall forecasting (Bowler et al., 2004; Cheung and Yeung, 2012; Krajewski and Smith, 2002).

The optical flow constraint (OFC, also known as the grey value constancy assumption) constitutes the core of the optical flow estimation (Bab-hadiashar and Suter, 1998; Baker et al., 2011; Brox et al., 2004; Bruhn et al., 2005; Horn and Schunck, 1981; Papenberg et al., 2006; Sun et al., 2010):

$$I(x, y, t) = I(x + u, y + v, t + \Delta T), \tag{1}$$

where $I(x, y, t)$ is the grey value (i.e. radar rainfall rate, in our case) of a pixel located at $\mathbf{x} = (x, y)$ of a given image $\mathbf{I}_t$ at time step $t$, and $\mathbf{w} = (u, v)$ is the movement vector to be estimated between images
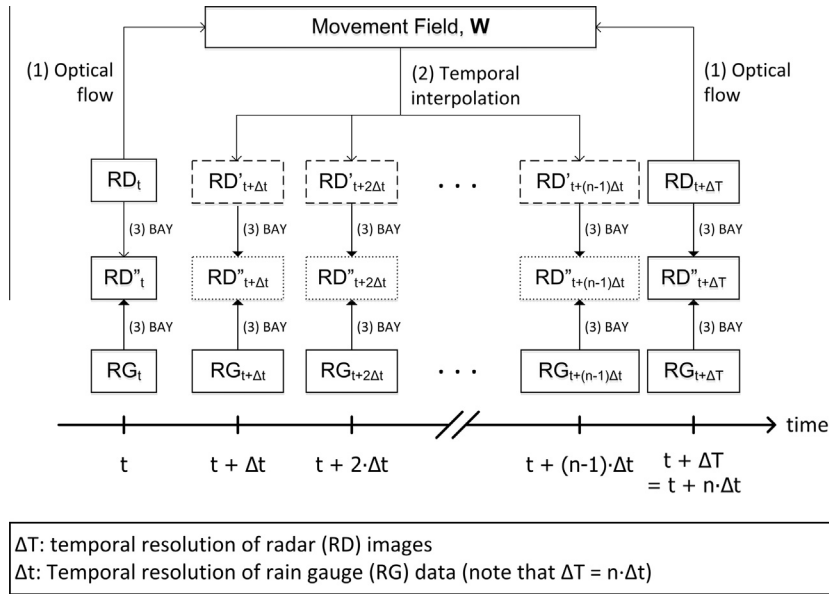
**Fig. 2.** Schematic of the proposed methodology for generating accurate radar rainfall images with high temporal resolution. RD and RG (with solid outlines) respectively represent radar and rain gauge observations; RD′ (with dashed outlines) corresponds to temporally-interpolated radar rainfall estimates; and RD″ with dotted outlines) corresponds to interpolated plus gauge-adjusted (through Bayesian (BAY) merging) radar rainfall estimates.

$I_t$ and $I_{t+\Delta T}$. This constraint indicates that image pixels only translate within $\Delta T$, but their grey values remain invariant. However, this is usually not the case, particularly for natural scenes. For instance, in the case of high-resolution radar rainfall images, the decay or growth of rainfall cell intensities can be observed even during a short duration (Foresti and Seed, 2014; Radhakrishna et al., 2012). In order to obtain better results, some smoothing treatment (e.g. median filter) is usually applied to the images before optical flow estimation is conducted (Bowler et al., 2004; Cheung and Yeung, 2012). This smoothing process is indeed helpful for solving OFC and has proven to be useful in large-scale applications (e.g. country-wide rainfall nowcasting and short-term flood forecasting; Bowler et al., 2006). However, for urban-scale applications, the smoothing process is usually not desirable because it may lead to significant underestimation of extreme rainfall magnitudes and associated runoff estimates (Wang et al., 2015).

In this work an alternative approach to handling the variations in pixel values between consecutive images is used. In this approach, an additional constraint, called gradient constancy assumption, is incorporated into the optical flow estimation (Brox et al., 2004; Papenberg et al., 2006; Sun et al., 2010). It is expressed as:

$$\nabla I(x, y, t) = \nabla I(x + u, y + v, t + \Delta T), \tag{2}$$

where $\nabla = (\partial_x, \partial_y)$ denotes the spatial gradient. This expression allows slight changes in pixel values, which somewhat relaxes the grey value constancy constraint. However, this constraint, as well as the OFC, focuses exclusively on the movement of individual pixels; their interactions with neighbouring pixels are not taken into account. This may cause the so-called aperture problem, which means that there may be spatial discontinuities (or holes) in the resulting movement fields. This is unrealistic and undesirable for applications. To prevent this, it is necessary to introduce a third constraint which ensures the smoothness of the movement field. This can be achieved by minimising the Laplacian of the movement fields (i.e. $\Delta \mathbf{w}$ or $\nabla^2 \mathbf{w}$), or equivalently by minimising the two components of the expansion of $\nabla^2 \mathbf{w}$, i.e., $\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ and

$\nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}$ (Bowler et al., 2004; Brox et al., 2004; Horn and Schunck, 1981).

To completely incorporate the three constraints described above into the optical flow estimation, a variational approach is used in this work, which was proposed by Brox et al. (2004) and is termed:

$$E(u, v) = E_{Data} + \alpha \cdot E_{smooth}, \tag{3}$$

where

$$E_{Data}(u, v) = \int_{\Omega} \Psi(|I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x})|^2) d\mathbf{x} + \gamma \cdot \int_{\Omega} \Psi(|\nabla I(\mathbf{x} + \mathbf{w}) - \nabla I(\mathbf{x})|^2) d\mathbf{x}, \tag{4}$$

and

$$E_{smooth}(u, v) = \int_{\Omega} \Psi(\nabla^2 u + \nabla^2 v) d\mathbf{x}. \tag{5}$$

where $\Psi$ is a penalty function and $\Omega$ is the radar domain. These two constraints define an energy functional that 'penalises' the deviation of image pixel values (i.e. $E_{Data}$) and the variation of movement fields (i.e. $E_{smooth}$). The aim is therefore to determine a movement field that minimises the overall (penalty) energy $E(u, v)$, which is a linear combination of each constraint with a given weight. In our research, weight values of $\alpha = 0.5$ and $\gamma = 50.0$ were assigned, based upon literature review (Sun et al., 2010) and a number of parametric analyses that confirmed the robustness of this selection. In addition, following the suggestion given in Brox et al. (2004) and Sun et al. (2010), the Charbonnier penalty function $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ (where $\epsilon = 0.001$) is employed to better retain the convexity of each energy term during the minimisation process.

This overall (penalty) energy functional can be minimised by solving the associated Euler–Lagrange equations. Because it is a classic approach to optimisation problems, the numerical formulation of this problem will not be repeated in this paper; for details readers are referred to Brox et al. (2004). Nonetheless, it is worth mentioning that another advantage of Brox's approach is the use of the multi-scale numerical strategy. This is done by firstly aggregating (in space) the original images into a number of images with coarser spatial resolutions. The estimation then starts by solving

the coarsest version of the problem (i.e. energy functional on the images with the coarsest spatial resolution). The associated solution is then used as the initial guess for solving a finer version of the problem. This iteration then continues until the original problem (i.e. at the original resolution) is solved. Numerically, this strategy can effectively prevent the estimation from being trapped in a local minimum and has proven to be able to largely improve optical flow estimation (Brox et al., 2004; Sun et al., 2010). Physically, this strategy explicitly takes into account the characteristic differences of images (in particular natural scenes) at various spatial and temporal scales, which has proven to be critical for rainfall field modelling (Deidda et al., 1999; Foresti and Seed, 2014; Germann and Zawadzki, 2004; Goudenhoofdt and Delobbe, 2013; Lovejoy and Schertzer, 1990; Wang et al., 2013a).

### 2.3. Temporal interpolation with occlusion reasoning

The temporal interpolation method used in this study is a bi-directional (i.e. forward–backward) advection-based one, which is similar to the one used in Nielsen et al. (2014). However, Nielsen et al. (2014) did not take into account the occlusion effect that is usually found in the temporal interpolation of consecutive images. Neglecting this effect may cause problems, such as the difficulty in interpolating pixel values along the image boundary and the generation of an over-smooth (in space) intermediate image, which may be problematic in the case of radar images over urban areas.

In order to include the occlusion effect, an additional technique, called occlusion reasoning, is employed in this paper before the intermediate images are interpolated (Herbst et al., 2009; Sadek et al., 2012). The algorithm used to conduct the temporal interpolation with occlusion reasoning is summarised as follows:

(0) Estimation of forward and backward movement fields ($\mathbf{w}^f = (u^f, v^f)$ and $\mathbf{w}^b = (u^b, v^b)$) using optical flow technique described in Section 2.2.
(1) Let $I(\mathbf{x}_{t+i\Delta t})$ be the grey value (or intensity) of a given pixel at location $\mathbf{x}_{t+i\Delta t} = (x, y)$ of an intermediate image $\mathbf{I}_{t+i\Delta t}$ (where $i = 1, 2, \ldots, n - 1$).
(2) By forward projection (i.e. using only the forward movement field $\mathbf{w}^f = (u^f, v^f)$), the corresponding pixels (i.e., $\mathbf{x}_t^f$ and $\mathbf{x}_{t+\Delta T}^f$) respectively in images $\mathbf{I}_t$ and $\mathbf{I}_{t+\Delta T}$ can be inversely identified. They can be termed:

$$\mathbf{x}_t^f = (x, y) - \delta(u^f, v^f), \tag{6}$$

and

$$\mathbf{x}_{t+\Delta T}^f = (x, y) + (1 - \delta)(u^f, v^f), \tag{7}$$

where $\delta = i/n$.

(2-1) If the (absolute) difference between $I(\mathbf{x}_t^f)$ and $I(\mathbf{x}_{t+\Delta T}^f)$ is smaller than a given threshold $\tau$ (that means, within time interval $\Delta T$, $I(\mathbf{x}_t^f)$ is not in a significant decay or growth stage, so presumably occlusion does not occur), forward interpolation is conducted; i.e., $I(\mathbf{x}_{t+i\Delta t}) = (1 - \delta) \cdot I(\mathbf{x}_t^f) + \delta \cdot I(\mathbf{x}_{t+T}^f)$.
(2-2) If the difference is larger than the threshold, no forward interpolation is conducted (the targeting pixel value will temporarily be a hole).
(3) (Backward projection) Based upon the backward movement field, the same algorithm as the forward projection is applied to fill the holes.
(4) If there are still holes left, they will be filled by the linear combination of $I(\mathbf{x}_t^f)$ and $I(\mathbf{x}_{t+\Delta T}^b)$; i.e., $I(\mathbf{x}_{t+i\Delta t}) = (1 - \delta) \cdot I(\mathbf{x}_t^f) + \delta \cdot I(\mathbf{x}_{t+T}^b)$.

### 2.4. Bayesian data merging

The Bayesian data merging method (BAY) is a dynamic adjustment method (applied independently at each time step) intended for real-time applications (Todini, 2001). In previous studies this method has shown to outperform other commonly-used adjustment methods (Mazzetti and Todini, 2004; Wang et al., 2013b). This is why it was adopted to conduct the gauge-based adjustment of radar estimates in the present study.

The underlying idea of the BAY method is to analyse and quantify the uncertainty of rainfall estimates (in terms of the co-variance of estimation errors) from multiple data sources – in this case, radar and rain gauge sensors – and then combine these estimates in such a way that the overall (estimation) uncertainty is minimised.

The implementation of the BAY merging method includes the techniques of block-kriging interpolation and Kalman filter (Kalman, 1960). The former is used to generate a field of rain gauge estimates (which represent the 'measurements' in the Kalman filter) and the latter is used to combine radar (defining the 'a priori' estimate) and block-kriged rain gauge rainfall estimates, to produce the merged rainfall output (i.e. the 'a posteriori' estimate). For a detailed description of the BAY technique the reader may refer to Todini (2001), Wang et al. (2013b) and Wang et al. (2015).

## 3. Case study

The proposed temporal interpolation method, in combination with the Bayesian merging, is tested using as case study a total of eight (8) storm events observed between 2012 and 2014 in the Cranbrook (London, UK) and Herent (Leuven, Belgium) urban catchments. For these events radar estimates, rain gauge and flow records are available. Moreover, recently-calibrated urban drainage models are available for each of the two pilot catchments.

### 3.1. Pilot catchments and datasets

#### 3.1.1. Cranbrook catchment

The Cranbrook catchment (Fig. 3) is located within the London Borough of Redbridge (north-east part of Greater London). It is predominantly urban, of residential and commercial character, and stretches over an area of 8.5 km². Approximately 52% of the catchment corresponds to impervious surfaces. The main water course (the Cran Brook) is about 5.75 km long, of which 5.69 km are culverted and have become part of the storm water drainage system, which is mainly separate. The storm water drainage system of this catchment discharges into the Roding River and, in turn, the Roding River discharges into the river Thames.

*Hydraulic model*: A high-resolution semi-distributed model of the storm-water drainage system of the Cranbrook catchment, including its sewer system, was setup in the integrated catchment modelling software InfoWorks ICM 5.5. In this model the whole catchment surface is split into sub-catchment units to which rainfall is applied (within each sub-catchment rainfall is assumed to be uniform). Each sub-catchment comprises a mix of pervious and impervious surfaces whose runoff drains to a common outlet point, which corresponds to an inlet node of the sewer system (i.e. a gully or a manhole). The sub-catchments are characterised by a number of parameters, including total area, length, slope, proportion of each land use type, amongst others. Based upon these parameters, the runoff volume at each sub-catchment is estimated using a fixed runoff coefficient, which varies according to the type of surface. The estimated runoff is then routed to the sub-catchment outlet using the Wallingford (double linear reservoir) model (Wallingford, 1983). Sewers are modelled as one-dimensional
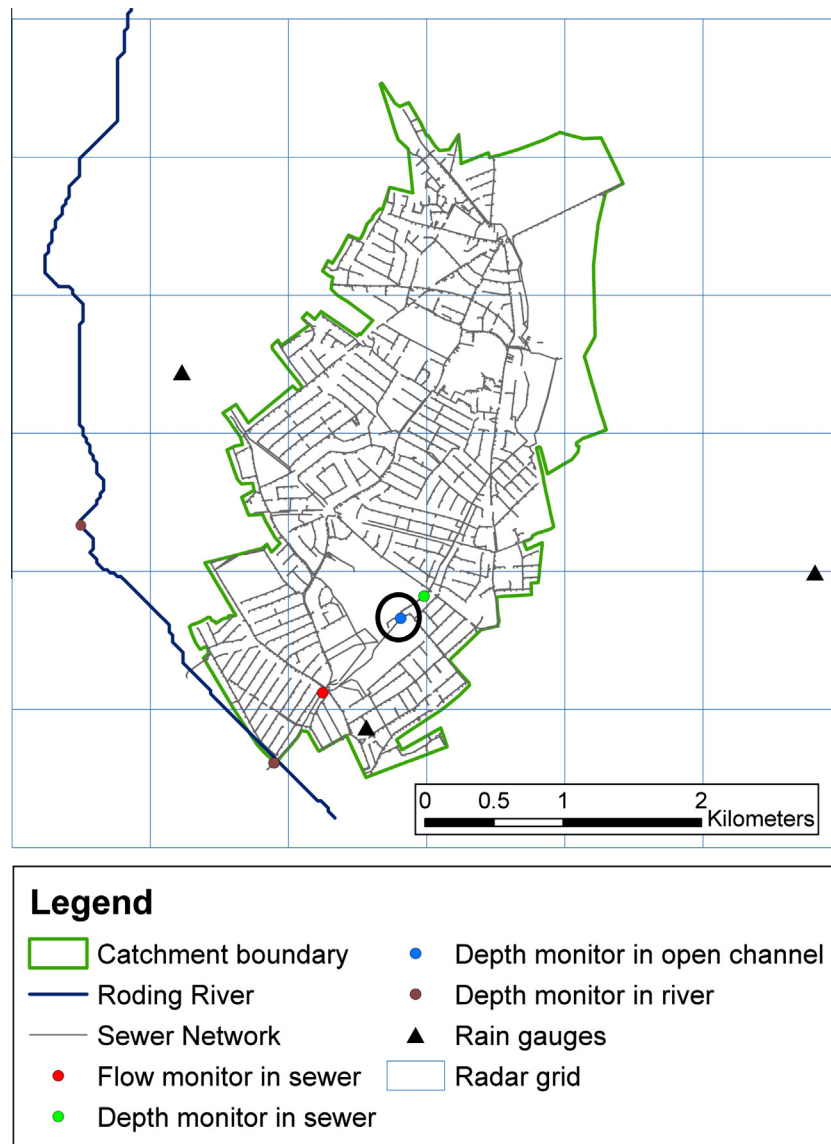
**Fig. 3.** Cranbrook catchment boundary, sewer layout, sensor location and radar grid over the catchment.

conduits and flows within them are simulated based on the full de Saint–Venant equations.

The Cranbrook model contains a total of 4409 sub-catchments, with areas ranging between 0.01 and 40.02 ha and a mean area of 0.19 ha. Sub-catchment slopes range from 0.02 to 0.41 m/m, with a mean slope of 0.05 m/m. The model of the sewer system comprises 6953 nodes and 6994 conduits. The total length of modelled sewers is 147 km. The sewer system ranges in height from 63.15 m above ordnance datum (mAOD) at the upstream end of the catchment to 6.05 mAOD at the downstream end. The modelled pipes have gradients ranging from 0 to 4.44 m/m, with mean and median gradients of 0.20 and 0.09, respectively. The flow in the sewer system is mainly driven by gravity.

Following UK standards (WaPUG, 2002), the model of the Cranbrook catchment was calibrated in early 2014 using data from the monitoring system described below.

*Local monitoring data available for this catchment*: A real time accessible monitoring system has been maintained in the Cranbrook catchment since April 2010. It includes three tipping bucket rain gauges (with 1-min resolution), two pressure sensors for monitoring water levels at the Roding River (downstream boundary condition of the catchment; with 2-min resolution),

one sensor for water depth measurement in sewers (with 2-min resolution), one sensor for flow (i.e. simultaneous water depth and velocity) measurement in sewers (with 2-min resolution), and one sensor for water depth measurement in open channels (with 1-min resolution) (see Fig. 3).

*Radar rainfall data*: The Cranbrook catchment and surroundings are within the coverage of C-band radars operated by the UK Met Office (UKMO). Radar rainfall estimates for the storm events selected for this study – for an 800 km × 800 km domain, centred at the Cranbrook catchment – were obtained through the British Atmospheric Data Centre (BADC). These estimates were available at spatial and temporal resolutions of 1 km and 5 min respectively, and correspond to a quality controlled multi-radar composite product generated with the UK Met Office Nimrod system (Golding, 1998), which includes corrections for the different errors inherent to radar rainfall measurements, as well as an hourly-based nation-wide mean field bias correction (Harrison et al., 2009, 2000).

*Storm events selected for this study*: Four storm events observed between May and September 2014 were selected to test the temporal interpolation + radar adjustment method proposed in this study. These events are different from those used for the

**Table 1**
Selected rainfall events over the Cranbrook catchment.

| ID | Date | Duration (h) | RG# | RG total (mm) | RG peak (mm/h) 1/5/10 min | RD total (mm) | RD peak (mm/h) 5/10 min |
|---|---|---|---|---|---|---|---|
| CBK-S01 | 01–02 May 2014 | 23 | 3 | 25.8 | 36.00/28.80/26.80 | 19.94 | 18.22/17.70 |
| CBK-S02 | 25–26 August 2014 | 20 | 3 | 28.73 | 16.00/8.00/6.80 | 28.84 | 7.01/5.56 |
| CBK-S03 | 19 September 2014 | 3.5 | 3 | 11.47 | 40.00/34.40/32.00 | 12.47 | 25.73/25.73 |
| CBK-S04 | 19 September 2014 | 3 | 3 | 7.07 | 32.00/25.60/18.40 | 10.06 | 20.10/20.02 |

calibration of the model. The dates and main characteristics of these events are summarised in Table 1.

### 3.1.2. Herent area

The catchment of Herent (called after the municipality of the same name) is located immediately northeast of the town of Leuven, the capital town of the Belgian province of Flemish Brabant. It includes several villages, with a predominantly urban residential character, which are separated by more rural areas. Important railway lines and a motorway are cutting through the catchment and contribute partly to the urban drainage system. The drainage area analysed in this study stretches over an area of approximately 25 km², of which about 14% corresponds to impervious surfaces. The sewer system in this area is mostly combined and includes several combined sewer overflows (CSOs) along the main flow path. The drainage system does not follow the natural topographical runoff direction of the catchment (the main brook is situated more along the northern edge), as a result of which

there are also several pumping stations on the main drainage flowpath.

*Hydraulic model*: A semi-distributed model of the drainage system of Herent was setup in the software package InfoWorks CS 15.0 by the water company Aquafin. Same as in the Cranbrook model, runoff volumes are estimated using a fixed runoff coefficient, runoff is routed to the sub-catchment outlets using the Wallingford (double linear reservoir) model (HR Wallingford, 1983), and sewer flow is simulated based upon the full de Saint–Venant equations.

Herent's model contains a total of 2495 sub-catchments, with areas ranging between 0.001 and 16.30 ha and a mean area of 0.40 ha. Sub-catchment slopes range from 0.00 to 0.34 m/m, with a mean slope of 0.01 m/m. The model of the sewer system comprises 3030 nodes and 2987 conduits. The total length of modelled sewers is 222 km. The sewer system ranges in height from 90.22 mAOD at the upstream end of the catchment to 13.12 mAOD at the downstream end. The modelled pipes have gradients ranging
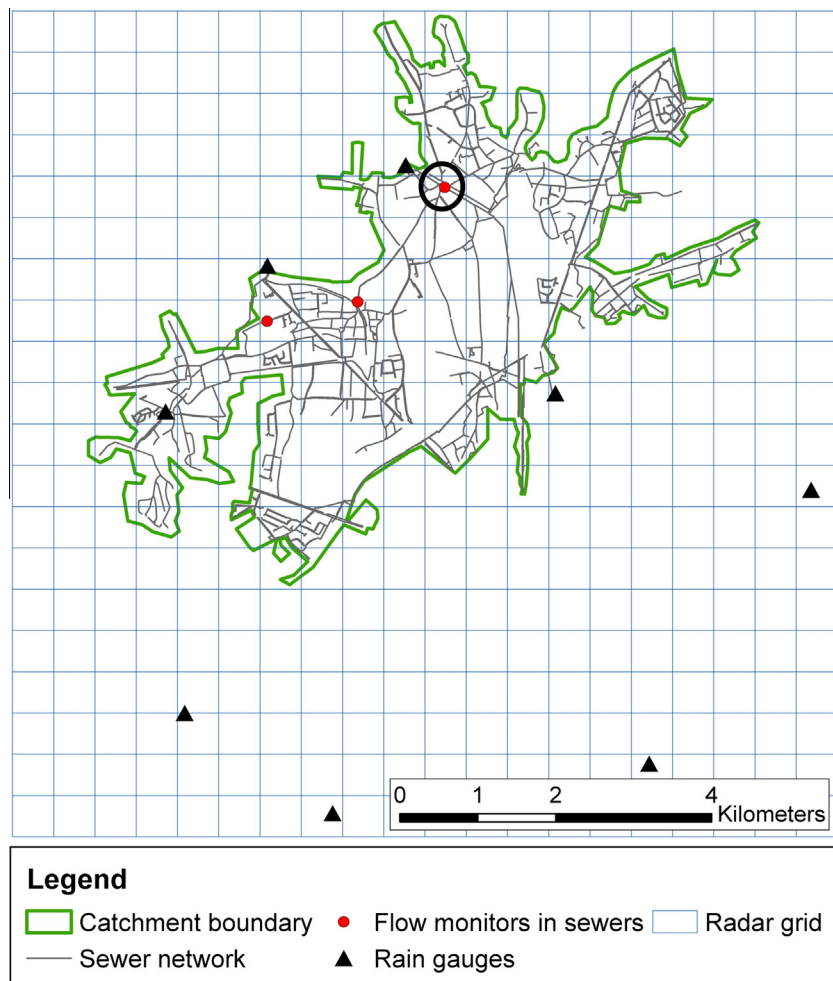


**Fig. 4.** Herent drainage area boundary, sewer layout, sensor location and radar grid over the catchment.

from 0 to 0.68 m/m, with mean and median gradients of 0.01 and 0.003, respectively. While the catchment is relatively flat, the flow in the sewer system is mainly driven by gravity.

The original model (built in the late '90s) was validated using the Flemish standard procedures described in Aquafin's Hydronaut procedure (Aquafin nv, 2014). In the years following, further validation was performed on a regular basis using data from the monitoring system described below, as well as from occasional short term flow surveys.

*Local monitoring data available for this catchment*: A monitoring system has been maintained in Herent and surroundings since 2006. It includes eight tipping bucket rain gauges (initially with 2 min resolution and since 2013 with 1 min resolution) and three flow (i.e. depth and velocity) gauges in sewers (with 2 min resolution) (see Fig. 4).

*Radar rainfall data*: C-band radar rainfall estimates for the whole of Belgium, including Herent, were derived from the radar reflectivity product provided by the Royal Meteorological Institute of Belgium (RMI) (Delobbe and Holleman, 2006). The reflectivity measurements, originally in dBZ units and at temporal and spatial resolutions of 529 m and 5 min respectively, were converted into rainfall rates (mm/h) using the so-called Marshall–Palmer $Z$–$R$ relationship: $Z = 200R^{1.6}$ (Marshall and Palmer, 1948). Although a new radar product is currently under development in RMI (Goudenhoofdt and Delobbe, 2013, 2009), for this study the estimates for Belgium do not include any gauge-based adjustment.

*Storm events selected for this study*: Four storm events observed in Herent between September 2012 and July 2013 were selected to test the temporal interpolation + radar adjustment method proposed in this study. The dates and main characteristics of these events are summarised in Table 2.

### 3.2. Experimental design and evaluation methodology

With the purpose of covering a range of operational conditions, the proposed temporal interpolation and the Bayesian data merging method were applied to radar (RD) rainfall estimates at two commonly-available temporal resolutions: 5 min and 10 min. Although for the pilot locations under consideration (i.e. Cranbrook and Herent) radar estimates were available at 5-min resolution, radar estimates at 10-min resolution are still commonly provided by national weather services (Chen et al., 2007; Gill et al., 2006; Vieux et al., 2003) and were for that reason included in the test. In order to realistically simulate operational radar scanning strategies, in this study the 10-min radar images were obtained through sampling (instead of aggregation). Following the methodology described in Section 2, the original RD estimates (at 5- and 10-min resolutions, respectively) were temporally interpolated to the finer temporal resolution at which rain gauge (RG) estimates were available. Afterwards, the Bayesian data merging was applied.

The performance of the resulting temporally-interpolated + Bayesian-adjusted estimates was assessed by inter-comparison against original RD and RG estimates, as well as against temporally-interpolated-only RD estimates (i.e. without merging) and Bayesian-adjusted-only estimates (i.e. at the original temporal resolution of RD estimates, without temporal interpolation). A summary of the rainfall estimates considered in the analysis, including the notation used to refer to them, is provided in Table 3. Two strategies were employed in the evaluation of the proposed method:

(1) Direct analysis and inter-comparison of the different rainfall estimates, both at areal level as well as at point RG locations.
(2) Analysis of the hydraulic outputs obtained by feeding the different rainfall estimates as input to the hydraulic model of the pilot locations and comparison of these against available water depth records. Note that the RG estimates were applied to the models using Thiessen polygons.

The performance measures employed in the evaluation are described below. Both evaluation strategies, however, have limitations. Firstly, the fact that all rainfall estimates (including RG ones) contain errors and the differences in the spatial resolution of RG and gridded rainfall estimates render any direct comparison of

**Table 2**
Selected rainfall events over Herent area.

| ID | Date | Duration (h) | RG# | RG total (mm) | RG peak (mm/h) 1–2[*]/5/10 min | RD total (mm) | RD peak (mm/h) 5/10 min |
|---|---|---|---|---|---|---|---|
| HER-S01 | 23 September 2012 | 1.8 | 8 | 9.99 | 25.94[*]/20.40/19.16 | 20.00 | 41.89/24.91 |
| HER-S02 | 08 May 2013 | 4 | 7 | 9.22 | 32.48/28.97/20.80 | 7.97 | 18.16/18.16 |
| HER-S03 | 27 July 2013 | 5 | 8 | 21.39 | 86.11/78.94/62.70 | 10.06 | 16.20/16.20 |
| HER-S04 | 27 July 2013 | 3 | 8 | 21.73 | 88.85/78.11/67.25 | 10.75 | 22.23/21.67 |

[*] For event HER-S01 rain gauge records were available at 2 min resolution. For all other events, rain gauge records were available at 1 min resolution.

**Table 3**
Summary of rainfall estimates under consideration in this study.

| Notation | Description | Temporal interpolation | Bayesian adjustment |
|---|---|---|---|
| RG | Local rain gauge rainfall records at highest available temporal resolution (i.e. 2-min for HER-S01 event and 1-min for all other events) | No | No |
| RD 5 | Original radar rainfall estimates sampled at 5-min time intervals | No | No |
| RD 5-1 (or RD 5-2) | Temporally-interpolated radar rainfall estimates, obtained from RD 5 | Yes | No |
| RD 10 | Original radar rainfall estimates sampled at 10-min time intervals | No | No |
| RD 10-1 (or RD 10-2) | Temporally-interpolated radar rainfall estimates, obtained from RD 10 | Yes | No |
| BAY 5 | Bayesian-adjusted rainfall estimates, obtained from RD 5 and 5-min RG rainfall estimates (aggregated from 1- or 2-min RG records) | No | Yes |
| BAY 5-1 (or BAY 5-2) | Temporally-interpolated + Bayesian-adjusted rainfall estimates, obtained from RD 5-1 (or RD 5-2) and RG rainfall estimates | Yes | Yes |
| BAY 10 | Bayesian-adjusted rainfall estimates, obtained from RD 10 and 10-min RG rainfall estimates (aggregated from 1- or 2-min RG records) | No | Yes |
| BAY 10-1 (or BAY 10-2) | Temporally-interpolated + Bayesian-adjusted rainfall estimates, obtained from RD 10-1 (or RD 10-2) and RG rainfall estimates | Yes | Yes |

rainfall estimates imperfect (Brandes et al., 2001). While the hydraulic evaluation strategy allows some of these drawbacks to be overcome, it has other deficiencies: hydraulic (depth and flow) records contain errors, and hydraulic modelling results encompass uncertainties from multiple sources in addition to rainfall input uncertainty (Deletic et al., 2012). Despite these limitations, both strategies provide useful and complementary insights into the performance of the proposed temporal interpolation + Bayesian data merging method.

### 3.2.1. Performance measures

Four performance measures were used to evaluate the proposed temporal interpolation + Bayesian data merging method. The first performance measure (i.e. event sample bias) was only employed in the rainfall evaluation strategy. The other three performance measures were used in both the rainfall and hydraulic evaluation strategies. In the case of the rainfall evaluation strategy, RG records at the highest-available temporal resolution were used as reference or 'observed' values, while the radar and radar-processed (gridded) rainfall estimates being assessed are referred to as 'modelled' estimates. Moreover, in the rainfall evaluation performance measures were estimated both at areal level (see areas shown in Figs. 3 and 4), as well as at RG point locations. In the case of the hydraulic evaluation strategy, water depth records at given locations within the pilot catchments were used as reference or 'observations', and the hydraulic simulation results associated to the different rainfall inputs constitute the 'modelled' estimates. The locations whose water depth records were used in the hydraulic evaluation are circled in black in Figs. 3 and 4; these locations were selected such that consistent records were available for all the storm events under consideration in this study.

(1) Event sample bias ratio ($B$): estimated as $B = \sum R_{gridded} / \sum RG$, where $R_{gridded}$ corresponds to the gridded (i.e. radar-processed) rainfall product being assessed. $B$ is used to quantify the cumulative event bias between gridded rainfall estimates and RG estimates (i.e. unconditional bias). $B = 1$ means no cumulative bias; $B > 1$ means that the accumulations of the gridded estimates are greater than those recorded by RG, and $B < 1$ the opposite.

(2) Coefficient of determination ($R^2$): $R^2$ corresponds to the coefficient of determination of a simple linear regression analysis applied to each pair of observed and modelled time series. The $R^2$ measure ranges from 0 to 1 and describes how much of the observed variability is explained by the modelled one. In practical terms, $R^2$ provides a measurement of the similarity between the patterns of the observed and modelled time series. This measure, however, cannot detect systematic bias (under- or over-estimation) of the modelled estimates (Krause et al., 2005).

(3) Regression coefficient ($\beta$): $\beta$ corresponds to the slope or gradient of the linear regression analysis conducted between each pair of observed and modelled time series. This measure provides information about conditional bias of the modelled estimates. $\beta \approx 1$ represents good agreement in the magnitude of observed and modelled estimates; $\beta > 1$ means that the modelled estimates are higher in the mean (by a factor of $\beta$) than the observations; and $\beta < 1$ means the opposite.

(4) Root mean square error (RMSE): the RMSE represents the standard deviation of the differences between observed ($O$) and modelled ($M$) values. The RMSE is estimated as $\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(O_t - M_t)^2}{n}}$. A perfect fit between observed and
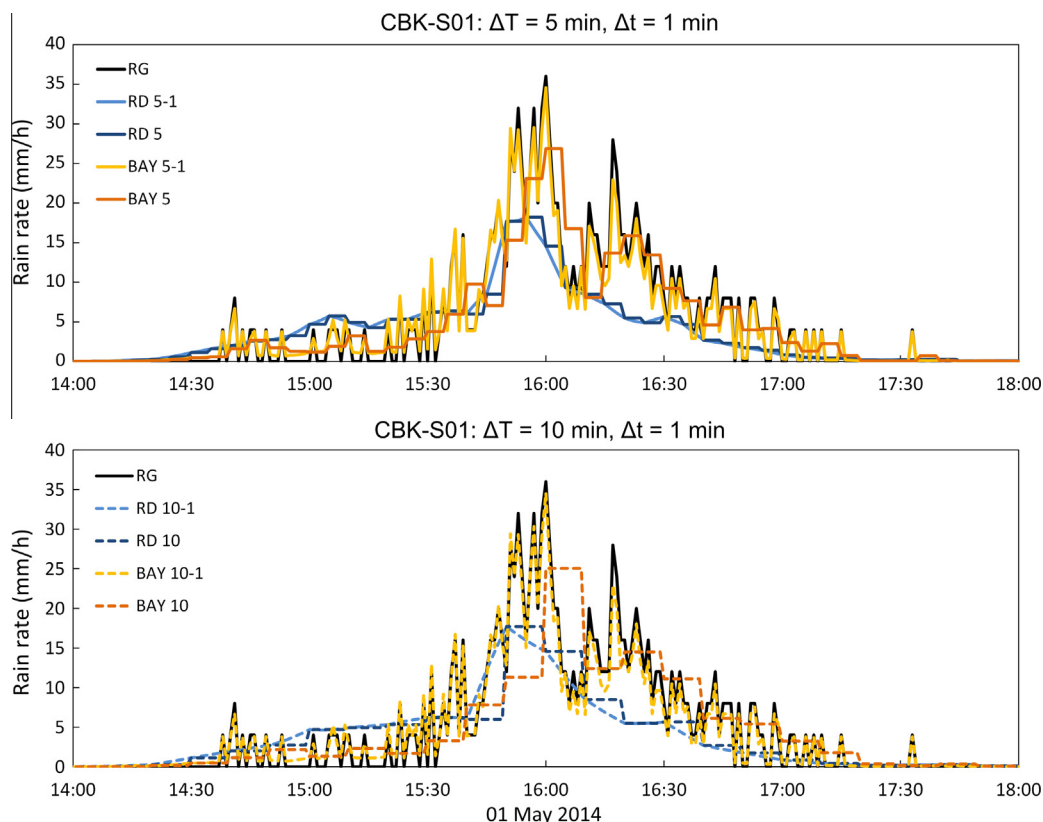


**Fig. 5.** Areal-average rainfall rate profiles of the different estimates under consideration for the CBK-S01 event. Top: RG vs. temporally-interpolated and/or adjusted rainfall estimates derived from 5-min RD estimates; Bottom: RG vs. temporally-interpolated and/or adjusted rainfall estimates derived from 10-min RD estimates.

modelled values corresponds to RMSE = 0. This measure provides information about overall performance of the modelled estimates.

## 3.3. Results and discussion

The results of the rainfall and hydraulic evaluation strategies, including selected images of the rainfall events and associated runoff estimates, as well as tables with summary performance measures, are presented in Figs. 5–11 and Tables 4–7. In what follows the results are analysed per rainfall product, starting with a brief discussion of the original radar estimates, continuing with an analysis of the independent effect of temporal interpolation and Bayesian merging, respectively, and concluding with the analysis of the temporally-interpolated + Bayesian merged estimates. Given that the results of the rainfall evaluation strategy are closely linked to those of the hydraulic evaluation, these are analysed simultaneously.

### 3.3.1. Analysis of raw radar rainfall estimates and associated runoff estimates

As can be seen from Tables 1 and 4 and Fig. 6 (see sample bias ratio), in the Cranbrook catchment the accumulations of RD estimates, both at areal level and at point RG location, are generally similar to those of RG records. This is however not the case for the Herent area (see Tables 2 and 5 and Fig. 7). Except for the HER-S02 event (which was a rather weak event), a large bias between RD and RG accumulations is observed in Herent. This difference in the performance in cumulative rainfall between UKMO and Belgian RMI RD products may be attributed to the different correction routines that are applied, in particular the
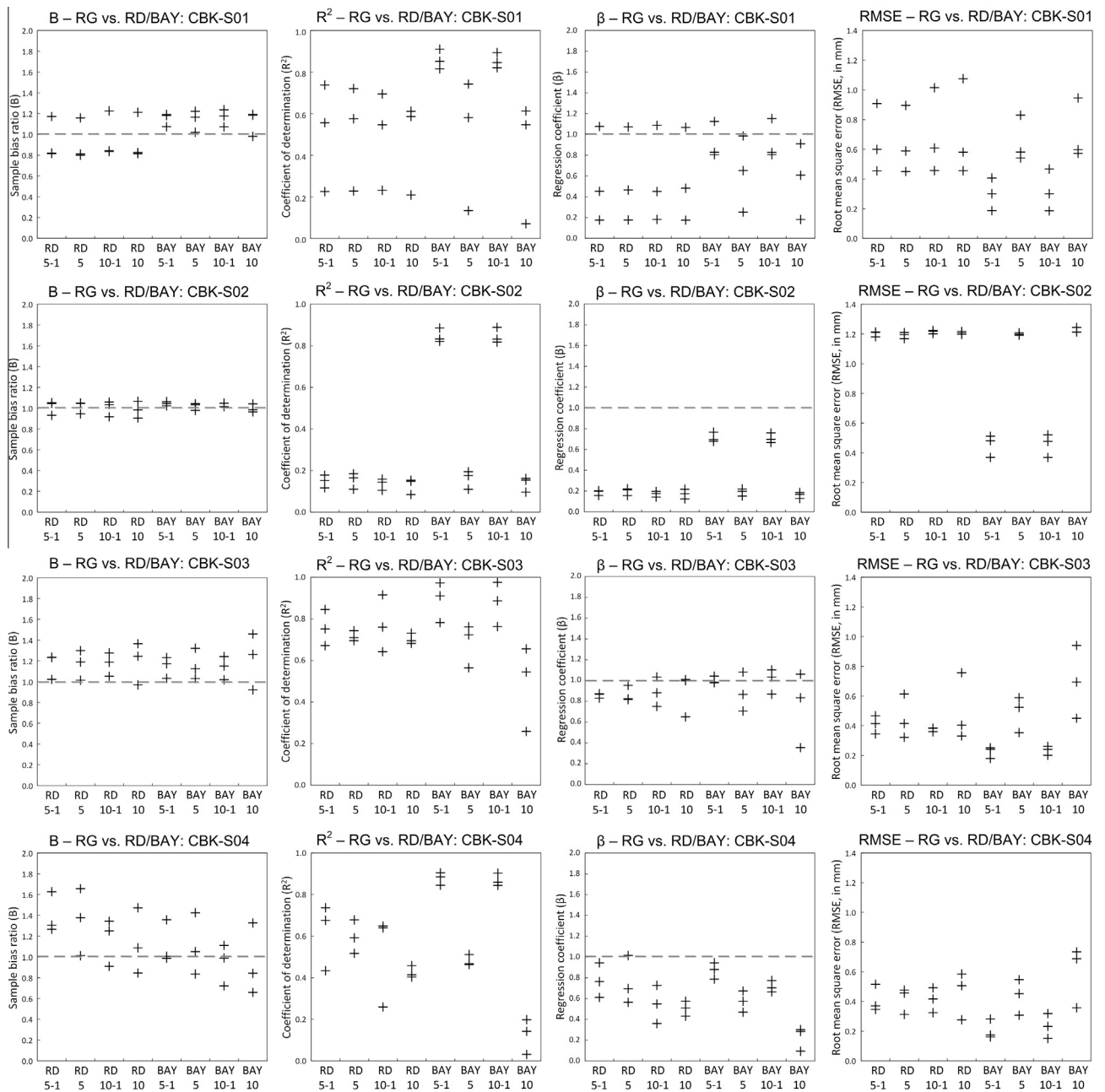


**Fig. 6.** Plots of performance measures at RG point locations for the different gridded rainfall estimates, for the selected storm events over the Cranbrook catchment.
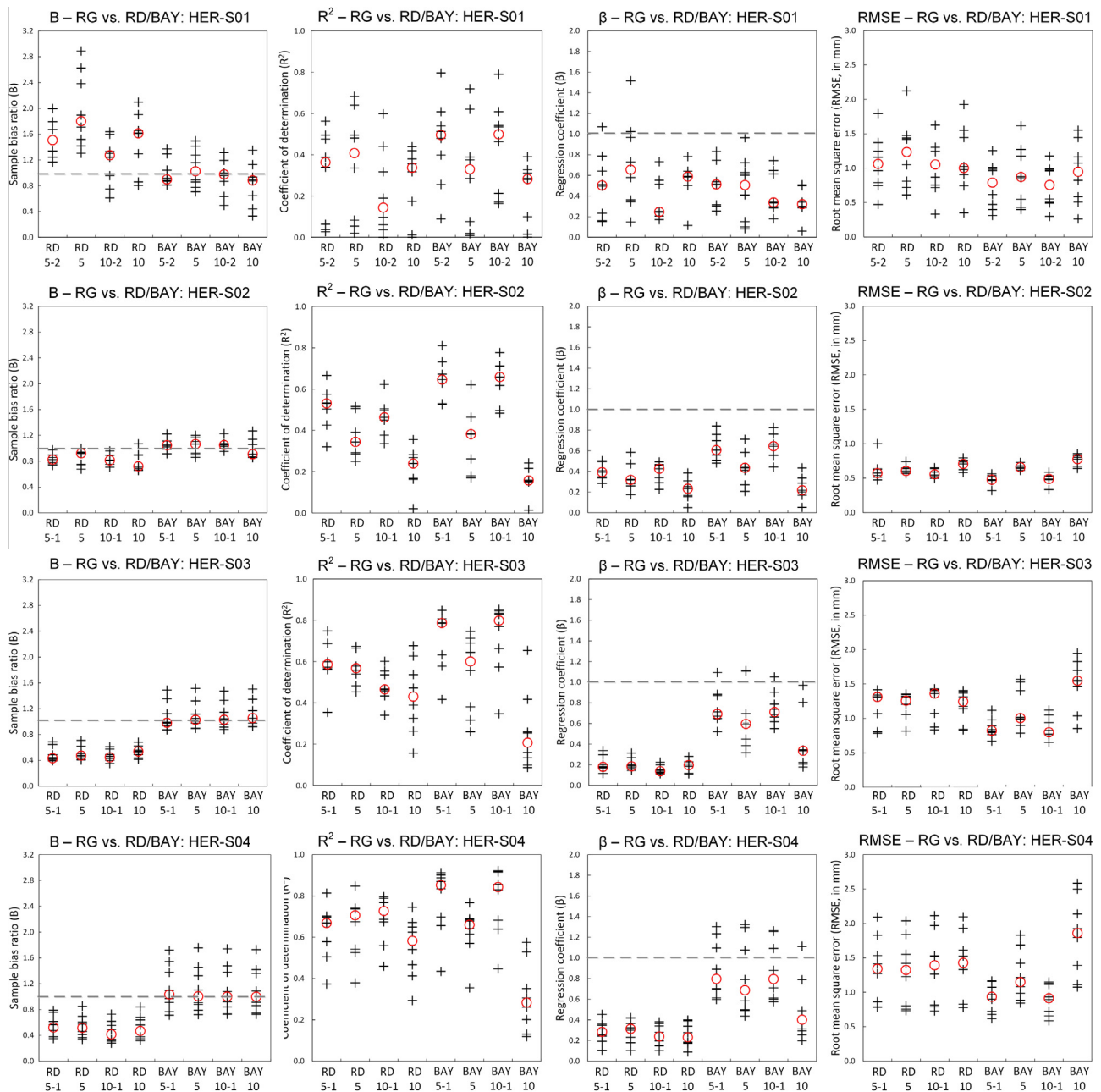
**Fig. 7.** Plots of performance measures at RG point locations for the different gridded rainfall estimates, for the selected storm events over the Herent area. The red circles correspond to the median of the performance measures estimated at all 7 or 8 point RG locations for each rainfall product. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

gauge-based adjustment that is applied to the UKMO RD product on an hourly basis, but not to the RMI product.

In terms of areal-average peak intensities, as compared with the rainfall accumulation performance, a larger bias between RD and RG estimates can generally be seen in both study areas (see Tables 1 and 2 and Fig. 5). Except for the HER-S01 event, RD rain rates appear to consistently underestimate the areal-average peak intensities recorded by RG. For the Cranbrook catchment, the relative difference between RG and RD areal-average (5-min) peak intensities ranges between 12 and 36%, while for the Herent area the relative difference may be up to 80% (see HER-S03 event). Even when the cumulative rainfall bias is small, the underestimation of peak rain rates by RD estimates has a large effect on the subsequent urban runoff estimates (notice how in Figs. 10 and 11 water depth estimates associated to RD 5 and RD 10 inputs

underestimate the recorded water depth peaks). This corroborates the view that the correction processes that have been applied to both UKMO and RMI RD products are insufficient to satisfactorily capture the fine-scale dynamics of rainfall fields, which are critical at urban scales. This confirms the need for local and dynamic adjustment of original RD estimates. Another detail worth noticing is the better performance of RD 5 associated hydraulic outputs, as compared to those associated to RD 10 estimates; this already suggests that there is an added benefit in using higher temporal resolution radar rainfall estimates.

### 3.3.2. Analysis of temporally-interpolated-only radar rainfall estimates and associated runoff estimates

From a visual point of view (see snapshot images in Figs. 8 and 9), the temporally-interpolated RD images show a smooth and
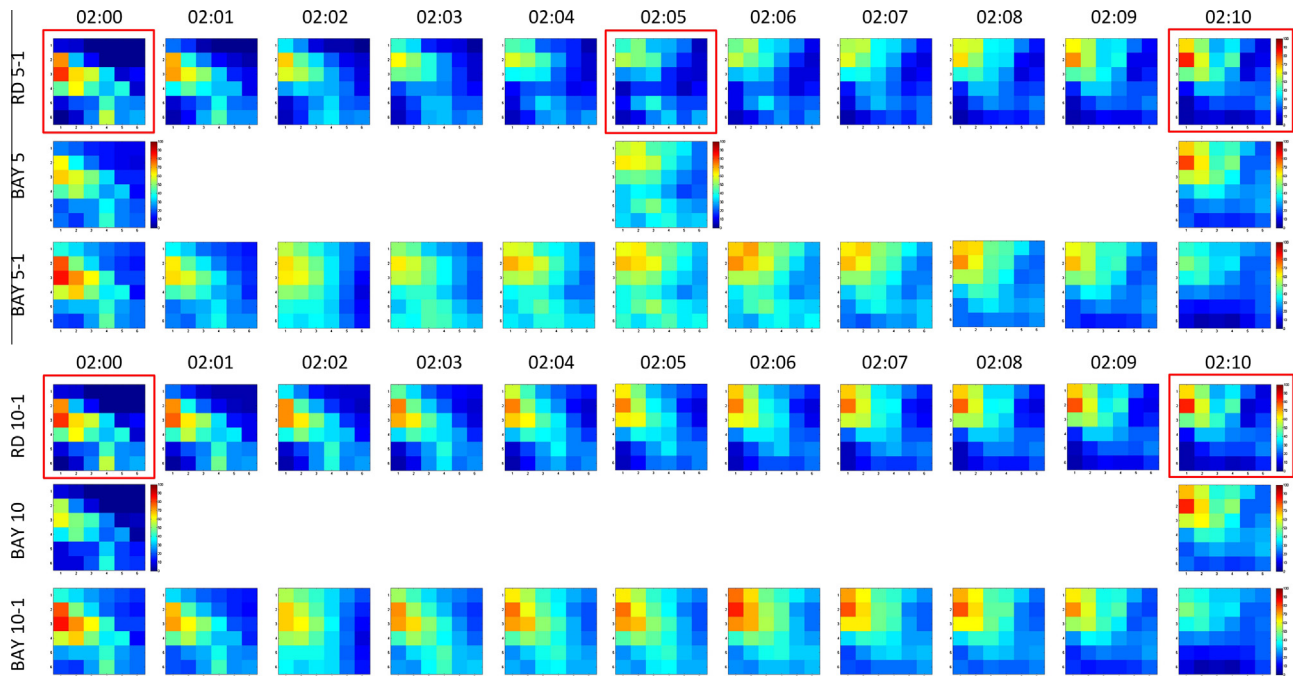
**Fig. 8.** Snapshot images of the original RD measurements in 5- and 10-min time intervals (images with red outlines), interpolated 1-min RD (i.e. RD 5-1 and RD 10-1, respectively interpolated from 5- and 10-min RD measurements) and the Bayesian-adjusted rainfall estimates (i.e. BAY 5, 5-1, 10 and 10-1) during the period of areal average peak intensity for the CBK-S03 event. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
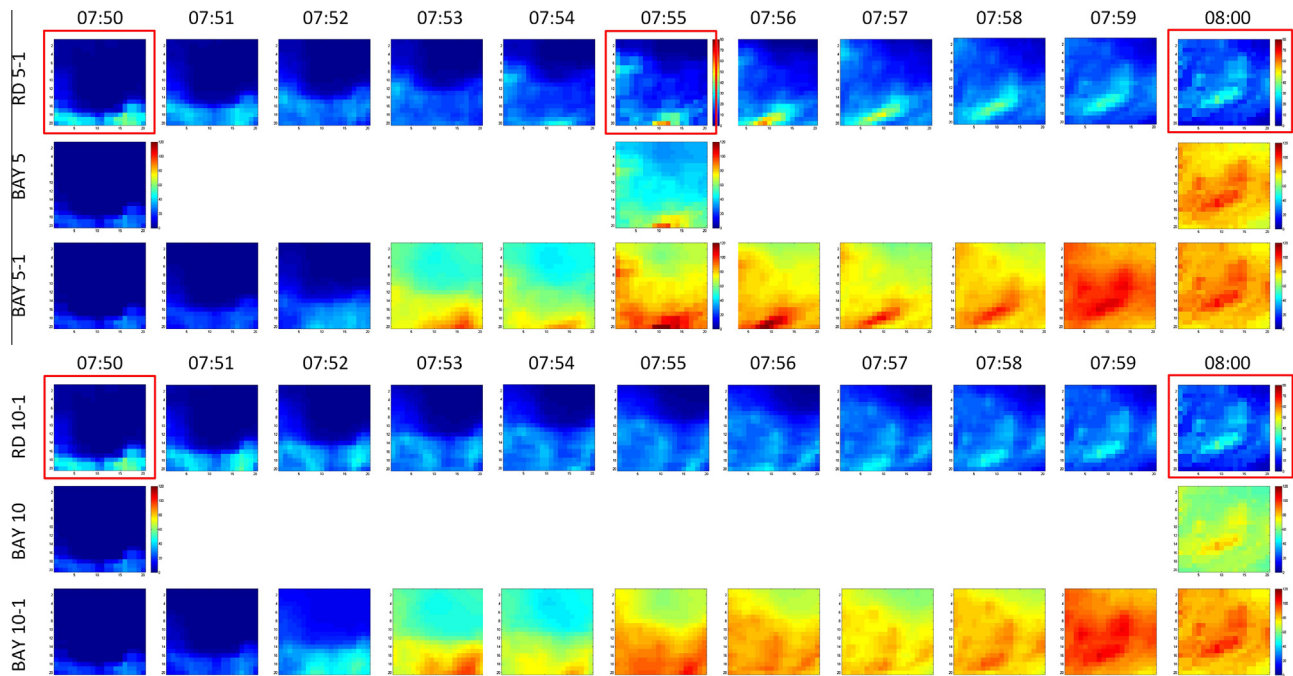


**Fig. 9.** Snapshot images of the original RD measurements in 5- and 10-min time intervals (images with red outlines), interpolated 1-min RD estimates (i.e. RD 5-1 and RD 10-1, respectively interpolated from 5- and 10 min RD measurements) and Bayesian adjusted rainfall estimates (i.e. BAY 5, 5-1, 10 and 10-1) during the period of areal-average peak intensity for the HER-S03 event. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

realistic transition of the rainfall clusters, including small-scale rainfall structures, between the available RD observations. By comparing the interpolated RD image over the Cranbrook catchment at time 02:05 (interpolated from 10 to 1 min) vs. the observed RD at the same time (Fig. 8), high consistency can be observed in the location of the storm cell cluster between the two images, although

some differences are observed in terms of intensities. Based upon visual inspection, it can be concluded that the proposed temporal-interpolation method is suitable for the small scales characteristic of urban areas. The fact that the location and shape of storm clusters is well reproduced suggests that the assumption of a constant field of motion vectors holds well within short time
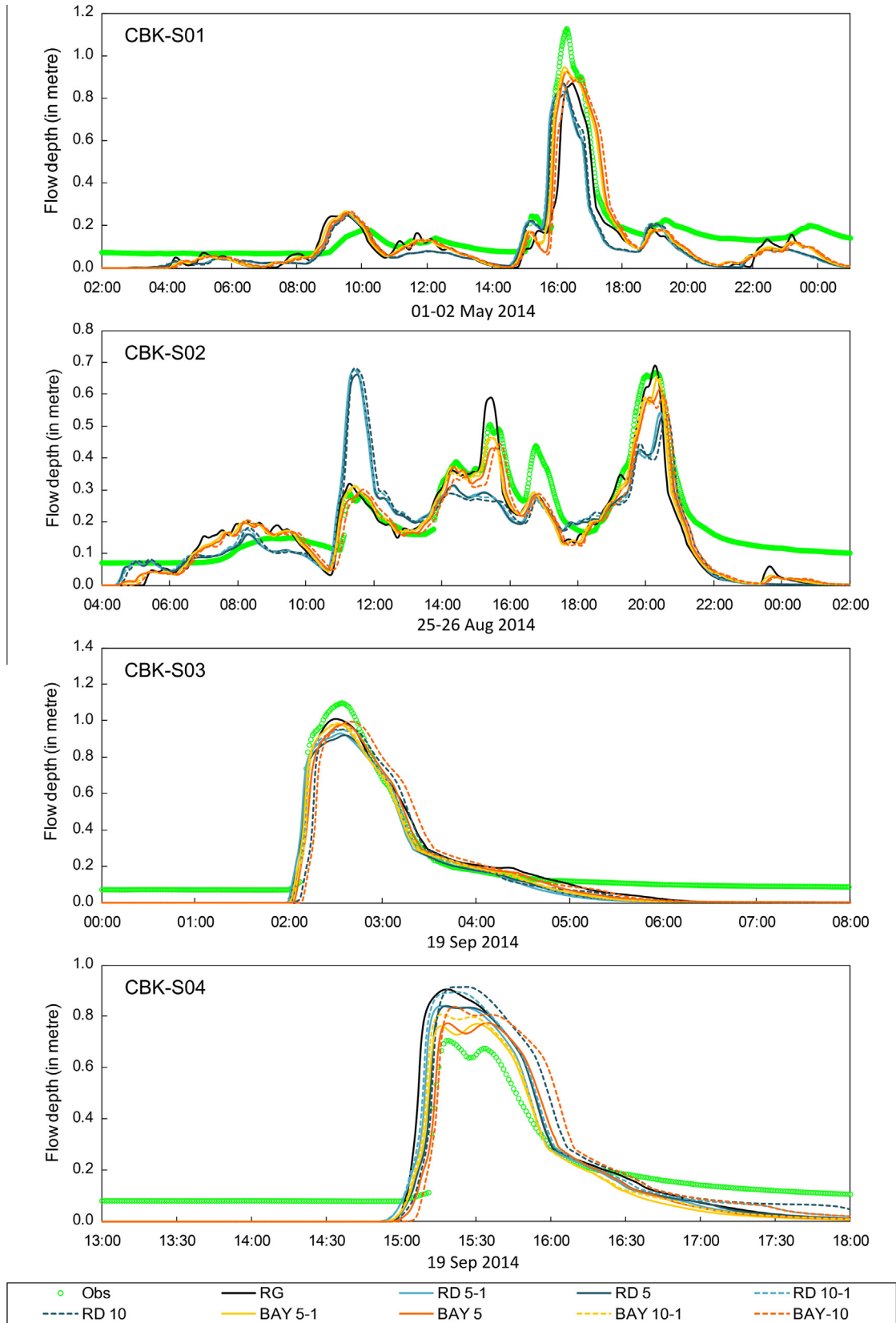
**Fig. 10.** Observed and simulated flow depths for the selected events at the Valentine Open Channel (the circled blue round marker in Fig. 3) in the Cranbrook catchment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

intervals. However, the differences observed in rainfall rates between the interpolated and observed RD images suggests that

assuming a linear change in rainfall rates within the time intervals under consideration may be inappropriate and that additional and
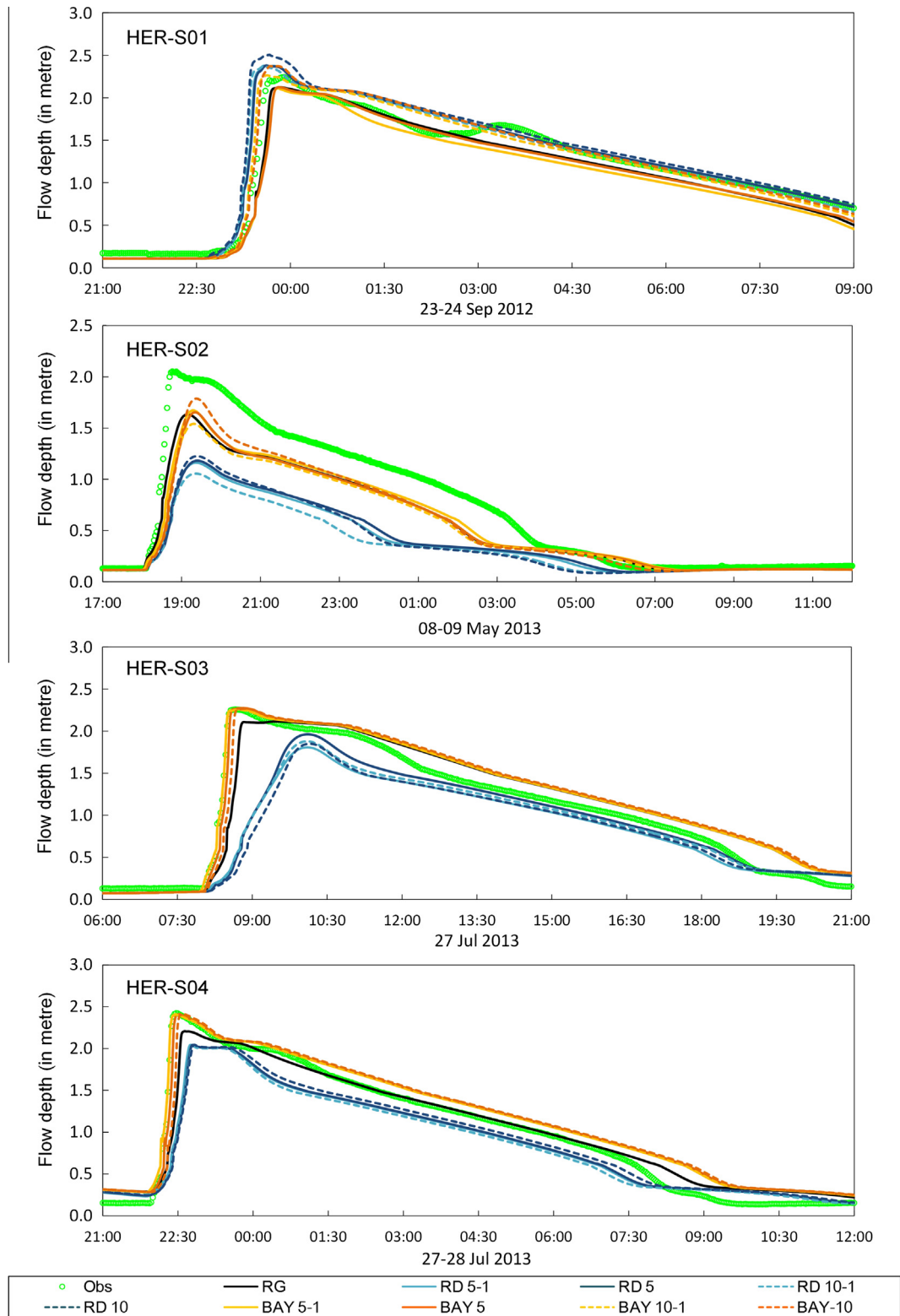
**Fig. 11.** Observed and simulated flow depths for the selected events at the M5 gauging site (the circled red round marker in Fig. 4) in the Herent area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

complementary sources of information (e.g. RG records at higher temporal resolution) are required to compensate for this deficiency.

Quantitatively, at the areal level, it can be seen that the temporal interpolation has a negligible effect in rainfall accumulations: in all cases coefficients $\beta$ and $\beta$ of the interpolated RD estimates stay

very close to those of the original RD estimates (see Tables 4 and 5). In terms of pattern, the intensities of the interpolated RD estimates show higher and generally better dynamics. This is reflected in the generally higher $R^2$ values associated with RD 5-1 and especially RD 10-1 estimates, particularly in the case of storms with strong peaks (see Tables 4 and 5). However, magnitude-wise the

**Table 4**
Summary of performance measures for the areal average gridded rainfall estimates for the selected storm events over the Cranbrook catchment (note that performance measures were estimated using areal-average RG records as reference).

| ID | RD 5-1 | RD 5 | RD 10-1 | RD 10 | BAY 5-1 | BAY 5 | BAY 10-1 | BAY 10 |
|---|---|---|---|---|---|---|---|---|
| *Sample bias ratio – B* | | | | | | | | |
| CBK-S01 | 0.772 | 0.773 | 0.788 | 0.789 | 0.994 | 0.962 | 0.993 | 0.954 |
| CBK-S02 | 1.005 | 1.004 | 1.009 | 1.005 | 1.034 | 1.002 | 1.033 | 0.999 |
| CBK-S03 | 1.081 | 1.088 | 1.136 | 1.129 | 1.135 | 1.144 | 1.142 | 1.171 |
| CBK-S04 | 1.430 | 1.424 | 1.417 | 1.414 | 1.114 | 1.126 | 1.084 | 1.125 |
| *Coefficient of determination – $R^2$* | | | | | | | | |
| CBK-S01 | 0.697 | 0.714 | 0.683 | 0.691 | 0.985 | 0.737 | 0.985 | 0.619 |
| CBK-S02 | 0.315 | 0.310 | 0.304 | 0.277 | 0.962 | 0.302 | 0.962 | 0.284 |
| CBK-S03 | 0.854 | 0.843 | 0.893 | 0.780 | 0.989 | 0.825 | 0.988 | 0.478 |
| CBK-S04 | 0.596 | 0.634 | 0.628 | 0.624 | 0.956 | 0.578 | 0.968 | 0.348 |
| *Regression coefficient – β* | | | | | | | | |
| CBK-S01 | 0.501 | 0.513 | 0.505 | 0.520 | 0.924 | 0.723 | 0.926 | 0.644 |
| CBK-S02 | 0.360 | 0.359 | 0.350 | 0.340 | 0.973 | 0.352 | 0.974 | 0.326 |
| CBK-S03 | 0.830 | 0.832 | 0.895 | 0.824 | 1.093 | 0.971 | 1.092 | 0.712 |
| CBK-S04 | 0.874 | 0.909 | 0.896 | 0.929 | 0.975 | 0.697 | 0.974 | 0.519 |
| *Root mean square error – RMSE (in mm)* | | | | | | | | |
| CBK-S01 | 0.038 | 0.037 | 0.038 | 0.037 | 0.009 | 0.032 | 0.009 | 0.039 |
| CBK-S02 | 0.037 | 0.037 | 0.037 | 0.038 | 0.009 | 0.037 | 0.009 | 0.037 |
| CBK-S03 | 0.054 | 0.056 | 0.046 | 0.066 | 0.022 | 0.063 | 0.022 | 0.112 |
| CBK-S04 | 0.072 | 0.068 | 0.068 | 0.071 | 0.021 | 0.064 | 0.017 | 0.082 |

**Table 5**
Summary of performance measures for the areal average gridded rainfall estimates for the selected storm events over the Herent area (note that performance measures were estimated using areal-average RG records as reference).

| ID | RD 5-1/RD 5-2 | RD 5 | RD 10-1/RD 10-2 | RD 10 | BAY 5-1/BAY 5-2 | BAY 5 | BAY 10-1/BAY 10-2 | BAY 10 |
|---|---|---|---|---|---|---|---|---|
| *Sample bias ratio – B* | | | | | | | | |
| HER-S01 | 1.629 | 2.002 | 1.418 | 1.833 | 1.170 | 1.287 | 1.158 | 1.241 |
| HER-S02 | 0.846 | 0.864 | 0.824 | 0.854 | 1.061 | 1.040 | 1.055 | 1.043 |
| HER-S03 | 0.464 | 0.470 | 0.428 | 0.484 | 1.033 | 1.044 | 1.036 | 1.054 |
| HER-S04 | 0.489 | 0.495 | 0.429 | 0.491 | 1.027 | 1.022 | 1.022 | 1.020 |
| *Coefficient of determination – $R^2$* | | | | | | | | |
| HER-S01 | 0.684 | 0.718 | 0.575 | 0.405 | 0.968 | 0.652 | 0.964 | 0.245 |
| HER-S02 | 0.859 | 0.740 | 0.843 | 0.444 | 0.993 | 0.646 | 0.993 | 0.336 |
| HER-S03 | 0.754 | 0.631 | 0.563 | 0.381 | 0.998 | 0.617 | 0.998 | 0.200 |
| HER-S04 | 0.883 | 0.862 | 0.898 | 0.700 | 0.999 | 0.723 | 0.999 | 0.292 |
| *Regression coefficient – β* | | | | | | | | |
| HER-S01 | 1.093 | 1.300 | 0.803 | 0.787 | 0.970 | 0.734 | 0.958 | 0.420 |
| HER-S02 | 0.592 | 0.556 | 0.579 | 0.444 | 0.984 | 0.750 | 0.985 | 0.492 |
| HER-S03 | 0.216 | 0.196 | 0.152 | 0.156 | 1.032 | 0.777 | 1.033 | 0.411 |
| HER-S04 | 0.309 | 0.311 | 0.277 | 0.281 | 1.024 | 0.849 | 1.023 | 0.516 |
| *Root mean square error – RMSE (in mm)* | | | | | | | | |
| HER-S01 | 0.213 | 0.277 | 0.189 | 0.280 | 0.053 | 0.154 | 0.054 | 0.231 |
| HER-S02 | 0.047 | 0.054 | 0.049 | 0.073 | 0.009 | 0.060 | 0.008 | 0.084 |
| HER-S03 | 0.181 | 0.186 | 0.196 | 0.197 | 0.013 | 0.145 | 0.013 | 0.226 |
| HER-S04 | 0.218 | 0.218 | 0.228 | 0.230 | 0.013 | 0.163 | 0.013 | 0.280 |

intensities of the interpolated RD estimates stay very close to the original RD intensities (see rain rate values in Fig. 5 and β values in Tables 4 and 5). This is due to the assumption of linear change in rainfall intensities mentioned above. Similar trends can be observed at RG point locations (see Figs. 6 and 7). However, when looking at the distribution of statistics at point locations, a larger spread is observed in the results associated to the interpolated RD estimates as compared to those of the original RD observations. This could be partially due to the uncertainty in the interpolation itself, as well as to the uncertainty in the high-resolution RG estimates.

In terms of hydraulic outputs (Tables 6 and 7 and Figs. 10 and 11), the temporally-interpolated RD estimates do not lead to significant improvements, as compared to the original RD rainfall inputs. The difference between the temporally-interpolated-only and the original RD estimates is usually small (i.e. RD 5-1 vs. RD 5 and RD 10-1 vs. RD 10) and mainly lies in the shifting (generally improving) of the timing to peak flow depth. Nonetheless, the improvements observed in the $R^2$ values of the interpolated RD estimates are generally smoothed-out throughout the hydrological and hydraulic modelling. Moreover, it is often found that the RD 5-1 and RD 10-1 estimates result in lower peak flow depths than the original RD estimates, which may be attributed to the inevitable smoothing nature of the interpolation process.

From the analysis of the individual effect of interpolation it can be concluded that while the interpolation helps re-gain the dynamics of rainfall inputs, it is not sufficient to fully correct rainfall estimates. This implies that other sources of error are still present and that a gauge-based adjustment is still necessary.

**Table 6**
Summary statistics of the simulated flow depths at the Valentine open channel (the circled blue round marker in Fig. 3) of the Cranbrook catchment.

| Event ID | RG | RD 5-1 | RD 5 | RD 10-1 | RD 10 | BAY 5-1 | BAY 5 | BAY 10-1 | BAY 10 |
|---|---|---|---|---|---|---|---|---|---|
| *Coefficient of determination – $R^2$* | | | | | | | | | |
| CBK-S01 | 0.827 | 0.796 | 0.838 | 0.789 | 0.874 | 0.886 | 0.858 | 0.881 | 0.775 |
| CBK-S02 | 0.815 | 0.515 | 0.527 | 0.500 | 0.488 | 0.843 | 0.857 | 0.848 | 0.849 |
| CBK-S03 | 0.976 | 0.970 | 0.972 | 0.971 | 0.888 | 0.980 | 0.955 | 0.972 | 0.814 |
| CBK-S04 | 0.835 | 0.907 | 0.963 | 0.891 | 0.963 | 0.939 | 0.953 | 0.924 | 0.903 |
| *Regression coefficient – $\beta$* | | | | | | | | | |
| CBK-S01 | 0.770 | 0.786 | 0.803 | 0.776 | 0.806 | 0.989 | 0.956 | 0.973 | 0.895 |
| CBK-S02 | 1.012 | 0.780 | 0.782 | 0.772 | 0.762 | 0.973 | 0.951 | 0.981 | 0.938 |
| CBK-S03 | 0.942 | 0.935 | 0.929 | 0.949 | 0.909 | 0.947 | 0.929 | 0.950 | 0.882 |
| CBK-S04 | 1.460 | 1.423 | 1.457 | 1.474 | 1.534 | 1.343 | 1.371 | 1.384 | 1.420 |
| *Root mean square error – RMSE (in metre)* | | | | | | | | | |
| CBK-S01 | 0.100 | 0.106 | 0.100 | 0.106 | 0.094 | 0.079 | 0.084 | 0.080 | 0.098 |
| CBK-S02 | 0.076 | 0.110 | 0.108 | 0.111 | 0.112 | 0.070 | 0.068 | 0.070 | 0.070 |
| CBK-S03 | 0.057 | 0.080 | 0.076 | 0.074 | 0.117 | 0.065 | 0.079 | 0.071 | 0.142 |
| CBK-S04 | 0.153 | 0.117 | 0.102 | 0.135 | 0.124 | 0.095 | 0.093 | 0.105 | 0.122 |

**Table 7**
Summary statistics of the comparison of observed and simulated flow depth at the M5 (the circled red round marker in Fig. 4) of the Herent area.

| ID | RG | RD 5-1/RD 5-2 | RD 5 | RD 10-1/RD 10-2 | RD 10 | BAY 5-1/BAY 5-2 | BAY 5 | BAY 10-1/BAY 10-2 | BAY 10 |
|---|---|---|---|---|---|---|---|---|---|
| *Coefficient of determination – $R^2$* | | | | | | | | | |
| HER-S01 | 0.965 | 0.930 | 0.936 | 0.889 | 0.879 | 0.957 | 0.953 | 0.962 | 0.978 |
| HER-S02 | 0.962 | 0.903 | 0.892 | 0.903 | 0.892 | 0.943 | 0.919 | 0.941 | 0.905 |
| HER-S03 | 0.890 | 0.724 | 0.693 | 0.708 | 0.636 | 0.985 | 0.970 | 0.984 | 0.934 |
| HER-S04 | 0.929 | 0.888 | 0.869 | 0.870 | 0.856 | 0.994 | 0.979 | 0.994 | 0.937 |
| *Regression coefficient – $\beta$* | | | | | | | | | |
| HER-S01 | 1.022 | 1.045 | 1.044 | 1.030 | 1.036 | 0.979 | 0.998 | 1.075 | 1.111 |
| HER-S02 | 0.734 | 0.520 | 0.524 | 0.474 | 0.571 | 0.731 | 0.712 | 0.692 | 0.781 |
| HER-S03 | 0.911 | 0.655 | 0.693 | 0.663 | 0.617 | 0.974 | 0.972 | 0.972 | 0.955 |
| HER-S04 | 0.846 | 0.786 | 0.775 | 0.766 | 0.775 | 0.904 | 0.897 | 0.903 | 0.877 |
| *Root mean square error – RMSE (in metre)* | | | | | | | | | |
| HER-S01 | 0.144 | 0.170 | 0.162 | 0.199 | 0.236 | 0.197 | 0.162 | 0.118 | 0.111 |
| HER-S02 | 0.304 | 0.583 | 0.566 | 0.643 | 0.576 | 0.298 | 0.332 | 0.346 | 0.313 |
| HER-S03 | 0.230 | 0.432 | 0.409 | 0.418 | 0.474 | 0.152 | 0.171 | 0.152 | 0.210 |
| HER-S04 | 0.184 | 0.299 | 0.309 | 0.330 | 0.299 | 0.148 | 0.165 | 0.148 | 0.210 |

*3.3.3. Analysis of Bayesian-adjusted-only radar rainfall estimates and associated runoff estimates*

In contrast to the temporal interpolation, the gauge-based adjustment alone generally has a large impact on rainfall estimates and subsequent runoff simulations.

With regards to rainfall estimates, the Bayesian merging alone largely improved the magnitude of peak rainfall rates as well as the event rainfall accumulations, both at areal level and at point RG locations (see magnitudes of peak intensities in Fig. 5 and $\beta$ values in Tables 4 and 5, Figs. 6 and 7). However, the application of the merging at the original temporal resolutions, particularly at 10 min, often caused a deterioration in the timing of peak intensities and in the pattern of rain rate time series in general, as compared to the original RD estimates. This is reflected in a decrease in $R^2$ and $\beta$ values, often accompanied by a deterioration in RMSE values (see Tables 4 and 5 and Figs. 6 and 7). The deterioration of the pattern implies a loss of dynamics which can be attributed to the joint effect of using RD estimates at coarse resolution in combination with RG estimates aggregated to a coarser resolution.

The effects of the Bayesian merging alone propagate through the hydraulic simulations, but are somehow smoothed in the process. On the one hand, the improvement in peak rainfall rates and accumulations clearly translated into improved reproduction of runoff peaks (see Figs. 10 and 11). The loss of dynamics of rainfall time series, on the other hand, did result in shifts (worsening) in the time to runoff peaks, particularly for the BAY 10 rainfall inputs. Nonetheless, the degree of dynamic loss (or alteration in time series pattern) observed in the runoff simulations is not as significant as that initially observed in the BAY 5 and BAY 10 rainfall estimates.

*3.3.4. Analysis of temporally-interpolated and Bayesian-adjusted radar rainfall estimates and associated runoff estimates*

As compared to the temporally-interpolated-only RD images, the temporally-interpolated and Bayesian-adjusted images show a similar shape and displacement of rain cells. Nonetheless, the latter display stronger and less linear changes in rainfall intensities, as a result of incorporating of high resolution RG information (see Figs. 8 and 9). This incorporation allows a more realistic representation of the growth and decay of storm cells, which is particularly evident in the CBK-S03 event (Fig. 8).

An interesting visual feature worth highlighting is the difference in the 07:55 images of BAY 5-1 (observed + merged) and BAY 10-1 (interpolated + merged) estimates in Fig. 9. It can clearly be seen that the pattern of the BAY 10-1 estimates is much smoother and less similar to the associated radar image than that of the BAY 5-1 one. This is because the interpolated radar estimate of the BAY 10-1 is found to be less reliable than the observed radar estimate at time 07:55 employed in the generation of BAY 5-1. As a result, given the theoretical formulation of the Bayesian merging method, the proportion of the RG information that is preserved throughout the adjustment is larger in the case of the BAY 10-1. This finding highlights the advantages of using the Bayesian merging method – as opposed to other adjustment methods, which do not account for the uncertainty of the different data sources being merged – in combination with the temporal interpolation.

Quantitatively, the combined application of temporal interpolation and Bayesian merging led to the best performing rainfall estimates (BAY 5-1 and BAY 10-1), both in terms of magnitude and pattern, at areal level as well as at point RG locations (see Fig. 5, Tables 4 and 5, Figs. 6 and 7). Remarkably, the BAY 10-1 estimates performed better than BAY 5 estimates in almost all events, thus bringing out the advantages of the proposed methodology. The improvements observed in the rainfall estimates are clearly seen in the associated runoff estimates, with BAY 5-1 and BAY 10-1 associated hydraulic outputs generally performing better than others, often even better than the outputs resulting from the original high resolution RG estimates. As compared to the runoff estimates associated to the Bayesian-adjusted-only rainfall inputs, the BAY 5-1 and BAY 10-1 estimates display a better agreement with observations, particularly in terms of the time to runoff peak. This highlights the benefits of applying gauge-based adjustments at higher temporal resolutions.

## 4. Conclusions and future work

In this paper a methodology is proposed to produce higher temporal-resolution (1–2 min), more accurate radar rainfall estimates, aiming at meeting the stringent requirements of urban hydrological applications. The proposed methodology involves two main steps: firstly, radar images are temporally interpolated from the original operational temporal resolutions (5–10 min) to a higher resolution at which local RG records are available (1–2 min). This is done using a novel interpolation technique introduced in the present paper, based upon the multi-scale variational optical flow technique, and which can well capture the small-scale rainfall structures relevant at urban scales. Afterwards, the higher temporal resolution radar rainfall estimates are dynamically merged with local rain gauge records by means of the Bayesian data merging method, which in previous studies had shown to outperform other merging methods, both at large and at small (urban scales).

Using as case study a total of eight storm events observed in the Cranbrook catchment (∼9 km$^2$) (London, UK) and in the Herent area (∼25 km$^2$) (Leuven, Belgium), the performance of the proposed method, in terms of the resulting rainfall estimates and the subsequent runoff estimates, was evaluated and compared against that of the original radar estimates and of the merged-only (i.e. without temporal interpolation) and interpolated-only (i.e. without merging) rainfall estimates. This analysis clearly brought out the benefits of the proposed combination of temporal interpolation and local and dynamic gauge-based adjustment of radar rainfall estimates. The main conclusions and implications of the testing are the following:

- The novel temporal interpolation technique introduced in this study proved able to handle small-scale rainfall features, resulting in realistic intermediate (interpolated) radar images, which served as basis for the subsequent local gauge-based adjustment. While the temporal interpolation alone does not show consistent quantitative results, its benefits are always evident in the case of storm events with fast-changing features.
- The proposed temporal interpolation + merging methodology can significantly improve radar rainfall estimates and associated runoff simulations at urban scales, much more so than the mere application of Bayesian merging at the original temporal resolution at which radar estimates are available. In other words, the results indicate that it is more effective to adjust or merge radar estimates at higher temporal resolutions, as opposed to doing so at coarser ones. In fact, at the coarsest temporal resolution analysed in this study (10 min), the Bayesian merging generally led to smaller improvements, and sometimes even caused deterioration, in both radar QPEs and associated runoff at the small urban scales considered in this study. This finding implies that the nation-wide gauge-based adjusted radar rainfall products which are generated through adjustments often performed at rather coarse temporal resolutions (⩾ 10 min) – e.g. the UKMO-Environment Agency product (Jewell and Gaussiat, 2015) and the Belgian RMI product being developed (Goudenhoofdt and Delobbe, 2009) – may not provide tangible benefits at urban scales.

However, in order to draw more robust conclusions about the applicability of the proposed method, further testing is required based on more storm events and pilot locations. Other aspects on which future work is recommended are the following:

- Testing of the proposed methodology using radar rainfall estimates at higher spatial resolutions. Given the strong relationship between the temporal and spatial resolution of rainfall inputs (Ochoa-Rodríguez et al., in press), it is likely that the benefits of the proposed methodology will become more evident when using radar estimates of higher spatial resolution, which is in agreement with the high temporal resolutions obtained through the methodology demonstrated in this study.
- Application of the temporal interpolation technique proposed herein in combination with the singularity-sensitive Bayesian merging method (Wang et al., 2015), which has proven to better preserve small-scale singularity structures throughout the merging process, thus being more suitable for urban applications. However, the fact that the singularity analysis would be applied on temporally-interpolated radar images, as opposed to actual radar observations, poses particular challenges which need to be carefully examined.
- Development of a stochastic temporal interpolation method, which accounts for the uncertainties associated with the advection model. This stochastic model could serve as basis for a rainfall error model for urban scales.
- Exploration of the use of the improved rainfall estimates obtained with the proposed methodology in a wider range of urban hydrological applications, including calibration of urban drainage models, reconstruction of historical storm events leading to known pluvial flooding, real-time simulation of urban runoff, amongst others.

## References

Anagnostou, E.N., Krajewski, W.F., 1999. Real-time radar rainfall estimation. Part I: algorithm formulation. J. Atmos. Ocean. Technol. 16, 189–197. http://dx.doi.org/10.1175/1520-0426(1999) 016<0189:RTRREP>2.0.CO;2.

Aquafin nv, 2014. Hydronaut Procedure (Internal Modelling Guidelines, in Dutch). Aartselaar, Belgium.

Bab-hadiashar, A., Suter, D., 1998. Robust optic flow computation. Int. J. Comput. Vis. 29, 59–77.

Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M., Szeliski, R., 2011. A database and evaluation methodology for optical flow. Int. J. Comput. Vis. 92, 1–31. http://dx.doi.org/10.1007/s11263-010-0390-2.

Berkowitz, D.S., Schultz, J.A., Vasiloff, S., Elmore, K.L., Payne, C.D., Boettcher, J.B., 2013. Status of dual pol QPE in the WSR-88D Network. In: The 93th AMS, 27th Conference on Hydrology. Austin, USA.

Berne, A., Delrieu, G., Creutin, J.-D., Obled, C., 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. J. Hydrol. 299, 166–179. http://dx.doi.org/10.1016/j.jhydrol.2004.08.002.

Borup, M., Grum, M., Linde, J.J., Mikkelsen, P.S., 2015. Dynamic gauge adjustment of high-resolution X-band radar data for convective rain storms: model-based evaluation against measured combined sewer overflow. J. Hydrol., submitted.

Bowler, N.E.H., Pierce, C.E., Seed, A., 2004. Development of a precipitation nowcasting algorithm based upon optical flow techniques. J. Hydrol. 288, 74–91. http://dx.doi.org/10.1016/j.jhydrol.2003.11.011.

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. Meteorol. Soc. 132, 2127–2155. http://dx.doi.org/10.1256/qj.04.100.

Brandes, E.A., Ryzhkov, A.V., Zrnić, D.S., 2001. An evaluation of radar rainfall estimates from specific differential phase. J. Atmos. Ocean. Technol. 18, 363–375. http://dx.doi.org/10.1175/1520-0426(2001) 018<0363:AEORRE>2.0.CO;2.

Bringi, V.N., Chandrasekar, V., 2001. Polarimetric Doppler Weather Radar Principles and Applications. Cambridge University Press, Cambridge, UK.

Bringi, V.N., Rico-Ramirez, M.A., Thurai, M., 2011. Rainfall estimation with an operational polarimetric C-band radar in the United Kingdom: comparison with a gauge network and error analysis. J. Hydrometeorol. 12, 935–954. http://dx.doi.org/10.1175/JHM-D-10-05013.1.

Brox, T., Bruhn, A., Papenberg, N., Weickert, J., 2004. High accuracy optical flow estimation based on a theory for warping. In: Pajdla, T., Matas, J. (Eds.), Computer Vision – ECCV 2004 SE – 3, Lecture Notes in Computer Science. Springer, Berlin Heidelberg, pp. 25–36. http://dx.doi.org/10.1007/978-3-540-24673-2_3.

Bruhn, A., Weickert, J., Schnorr, C., 2005. Lucas/Kanade M meets Horn/Schunck: combining local and global optic flow methods. Int. J. Comput. Vis. 61, 211–231. http://dx.doi.org/10.1023/B:VISI.0000045324.43199.43.

Chandrasekar, V., Wang, Y., Maki, M., Nakane, K., 2009. Urban flood monitoring using X-band dual-polarization radar network: program of the CASA-NIED partnership. In: The 34th Conference on Radar Meteorology. Williamsburg, USA.

Chen, C.-Y., Lin, L.-Y., Yu, F.-C., Lee, C.-S., Tseng, C.-C., Wang, A.-H., Cheung, K.-W., 2007. Improving debris flow monitoring in Taiwan by using high-resolution rainfall products from QPESUMS. Nat. Hazards 40, 447–461. http://dx.doi.org/10.1007/s11069-006-9004-2.

Cheung, P., Yeung, H.Y., 2012. Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong. In: The 3rd WMO International Symposium on Nowcasting and Very Short-Range Forecasting (WSN12). Rio de Janeiro, Brasil.

Cole, S.J., Moore, R.J., 2008. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. J. Hydrol. 358, 159–181. http://dx.doi.org/10.1016/j.jhydrol.2008.05.025.

Collier, C.G., 1986. Accuracy of rainfall estimates by radar, Part I: calibration by telemetering raingauges. J. Hydrol. 83, 207–223. http://dx.doi.org/10.1016/0022-1694(86)90152-6.

Collier, C.G., 1996. Applications of Weather Radar Systems, second ed. Wiley, Chichester, England.

Collier, C.G., 2009. On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models. Meteorol. Appl. 16, 35–40. http://dx.doi.org/10.1002/met.120.

Deidda, R., 2000. Rainfall downscaling in a space–time multifractal framework. Water Resour. Res. 36, 1779–1794. http://dx.doi.org/10.1029/2000WR900038.

Deidda, R., Benzi, R., Siccardi, F., 1999. Multifractal modeling of anomalous scaling laws in rainfall. Water Resour. Res. 35, 1853–1867. http://dx.doi.org/10.1029/1999WR900036.

Deletic, A., Dotto, C.B.S., McCarthy, D.T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T.D., Rauch, W., Bertrand-Krajewski, J.L., Tait, S., 2012. Assessing uncertainties in urban drainage models. Phys. Chem. Earth, Parts A/B/C 42–44, 3–10. http://dx.doi.org/10.1016/j.pce.2011.04.007.

Delobbe, L., Holleman, I., 2006. Uncertainties in radar echo top heights used for hail detection. Meteorol. Appl. 13, 361. http://dx.doi.org/10.1017/S1350482706002374.

Delobbe, L., Goudenhoofdt, E., Mohymont, B., 2008. Improvement of quantitative precipitation estimates in Belgium. In: Proceedings of the 5th European Conference on Radar in Meteorology and Hydrology.

Ehrhardt, J., Werner, R., Säring, D., Frenzel, T., Lu, W., Low, D., Handels, H., 2007. An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. Med. Phys. 34.

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. J. Hydrol. 299, 186–202. http://dx.doi.org/10.1016/j.jhydrol.2004.08.004.

Einfalt, T., Jessen, M., Mehlig, B., 2005. Comparison of radar and raingauge measurements during heavy rainfall. Water Sci. Technol. 51, 195–201.

Fabry, F., Bellon, A., Duncan, M.R., Austin, G.L., 1994. High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. J. Hydrol. 161, 415–428. http://dx.doi.org/10.1016/0022-1694(94)90138-4.

Foresti, L., Seed, A., 2014. The effect of flow and orography on the spatial distribution of the very short-term predictability of rainfall from composite radar images. Hydrol. Earth Syst. Sci. 18, 4671–4686. http://dx.doi.org/10.5194/hess-18-4671-2014.

Fulton, R.A., Breidenbach, J.P., Dong-Jun, S., Miller, D.A., 1998. The WSR-88D rainfall algorithm. Weather Forecast. 13, 377–395.

Gao, J., Xue, M., Shapiro, A., Droegemeier, K.K., 1999. A variational method for the analysis of three-dimensional wind fields from two doppler radars. Mon. Weather Rev. 127, 2128–2142. http://dx.doi.org/10.1175/1520-0493(1999) 127<2128:AVMFTA>2.0.CO;2.

Germann, U., Zawadzki, I., 2002. Scale-dependence of the predictability of precipitation from continental radar images. Part I: description of the methodology. Mon. Weather Rev. 130, 2859–2873. http://dx.doi.org/10.1175/1520-0493(2002) 130<2859:SDOTPO>2.0.CO;2.

Germann, U., Zawadzki, I., 2004. Scale dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. J. Appl. Meteorol. 43, 74–89. http://dx.doi.org/10.1175/1520-0450(2004) 043<0074:SDOTPO>2.0.CO;2.

Germann, U., Berenguer, M., Sempere-torres, D., Zappa, M., 2009. REAL – ensemble radar precipitation estimation for hydrology in a mountain region. Q. J. R. Meteorol. Soc. 135, 445–456. http://dx.doi.org/10.1002/qj.

Gerstner, E.-M., Heinemann, G., 2008. Real-time areal precipitation determination from radar by means of statistical objective analysis. J. Hydrol. 352, 296–308. http://dx.doi.org/10.1016/j.jhydrol.2008.01.016.

Gill, R.S., Overgaard, S., Bøvith, T., 2006. The Danish weather radar network. In: European Radar Conference 2006. Barcelona, Spain.

Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., Simoes, N., 2012. Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: a case study. J. Hydrol. 442–443, 117–128. http://dx.doi.org/10.1016/j.jhydrol.2012.04.005.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., Berne, A., Lovejoy, S., 2014. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. Atmos. Res. 138, 125–138. http://dx.doi.org/10.1016/j.atmosres.2013.11.008.

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

Goudenhoofdt, E., Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrol. Earth Syst. Sci. 13, 195–203. http://dx.doi.org/10.5194/hess-13-195-2009.

Goudenhoofdt, E., Delobbe, L., 2013. Statistical characteristics of convective storms in Belgium derived from volumetric weather radar observations. J. Appl. Meteorol. Climatol. 52, 918–934. http://dx.doi.org/10.1175/JAMC-D-12-079.1.

Gupta, V.K., Waymire, E.C., 1993. A statistical analysis of mesoscale rainfall as a random cascade. J. Appl. Meteorol. 32, 251–267. http://dx.doi.org/10.1175/1520-0450(1993) 032<0251:ASAOMR>2.0.CO;2.

Harrison, D.L., Driscoll, S.J., Kitchen, M., 2000. Improving precipitation estimates from weather radar using quality control and correction techniques. Meteorol. Appl. 7, 135–144. http://dx.doi.org/10.1017/S1350482700001468.

Harrison, D.L., Scovell, R.W., Kitchen, M., 2009. High-resolution precipitation estimates for hydrological uses. Proc. Inst. Civ. Eng. Water Manage. 162, 125–135.

Herbst, E., Seitz, S., Baker, S., 2009. Occlusion reasoning for temporal interpolation using optical flow. Microsoft Research.

Horn, B.K.P., Schunck, B.G., 1981. Determining optical flow. Arti. Intell. 17, 185–203. http://dx.doi.org/10.1016/0004-3702(81)90024-2.

Jewell, S.A., Gaussiat, N., 2015. An assessment of Kriging based rain-gauge – radar merging techniques. Q. J. R. Meteorol. Soc. http://dx.doi.org/10.1002/qj.2522.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME – J. Basic Eng. 82, 35–45. http://dx.doi.org/10.1115/1.3662552.

Kim, J.-H., Yang, J.-G., Kim, S.-H., Kim, J.-S., 2012. Enhancement of Korean weather radar capability by introducing a dual-pol radar network. In: WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation. Brussels, Belgium.

Koutsoyiannis, D., Onof, C., 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. J. Hydrol. 246, 109–122. http://dx.doi.org/10.1016/S0022-1694(01)00363-8.

Krajewski, W.F., Smith, J.A., 2002. Radar hydrology: rainfall estimation. Adv. Water Resour. 25, 1387–1394. http://dx.doi.org/10.1016/S0309-1708(02)00062-3.

Krämer, S., Fuchs, L., Verworn, H., 2007. Aspects of radar rainfall forecasts and their effectiveness for real time control – the example of the sewer system of the city of Vienna. Water Pract. Technol. 2, 42–49.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97. http://dx.doi.org/10.5194/adgeo-5-89-2005.

Laroche, S., Zawadzki, I., 1995. Retrievals of horizontal winds from single-doppler clear-air data by methods of cross correlation and variational analysis. J. Atmos. Ocean. Technol. 12, 721–738. http://dx.doi.org/10.1175/1520-0426(1995) 012<0721:ROHWFS>2.0.CO;2.

Larsen, R., Hansen, J.D., Ersboll, B.K., Conradsen, K., 1997. Temporal Interpolation in Meteosat Images, Technical University of Denmark.

Leandro, J., Chen, A.S., Schumann, A., 2014. A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave). J. Hydrol. 517, 250–259. http://dx.doi.org/10.1016/j.jhydrol.2014.05.020.

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. http://dx.doi.org/10.1016/j.atmosres.2011.05.004.

Lovejoy, S., Schertzer, D., 1990. Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. J. Geophys. Res. 95. http://dx.doi.org/10.1029/JD095iD03p02021.

Marsan, D., Schertzer, D., Lovejoy, S., 1996. Causal space–time multifractal processes: predictability and forecasting of rain fields. J. Geophys. Res. 101. http://dx.doi.org/10.1029/96JD01840.

Marshall, J.S., Palmer, W.M.K., 1948. The distribution of raindrops with size. J. Meteorol. 5, 165–166. http://dx.doi.org/10.1175/1520-0469(1948) 005<0165:TDORWS>2.0.CO;2.

Mazzetti, C., Todini, E., 2004. Combining raingauges and radar precipitation measurements using a Bayesian approach. In: Sanchez-Vila, X., Carrera, J., Gómez-Hernández, J.J. (Eds.), GeoENV IV – Geostatistics for Environmental Applications. Kluwer Academic Publishers, pp. 401–412. http://dx.doi.org/ 10.1007/1-4020-2115-1_34.

Nielsen, J.E., Thorndahl, S., Rasmussen, M.R., 2014. A numerical method to generate high temporal resolution precipitation time series by combining weather radar measurements with a nowcast model. Atmos. Res. 138, 1–12. http://dx.doi.org/ 10.1016/j.atmosres.2013.10.015.

Ochoa-Rodríguez, S., Wang, L.-P., Gires, A., Pina, R., Reinoso-Rodinel, R., Bruini, G., van Assel, J., Kroll, S., Murlà-Tuyls, D., Ichiba, A., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., Ten Veldhuis, J.A.E., 2015. Impact of spatial and temporal resolution of rainfall inputs on semi-distributed urban drainage modelling outputs: a multi-catchment investigation. J. Hydrol., in press.

Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J., 2006. Highly accurate optic flow computation with theoretically justified warping. Int. J. Comput. Vis. 67, 141–158. http://dx.doi.org/10.1007/s11263-005-3960-y.

Pegram, G.G.S., Clothier, A.N., 2001. High resolution space–time modelling of rainfall: the "String of Beads" model. J. Hydrol. 241, 26–41. http://dx.doi.org/ 10.1016/S0022-1694(00)00373-5.

Radhakrishna, B., Zawadzki, I., Fabry, F., 2012. Predictability of precipitation from continental radar images. Part V: Growth and decay. J. Atmos. Sci. 69, 3336–3349. http://dx.doi.org/10.1175/JAS-D-12-029.1.

Rinehart, R.E., Garvey, E.T., 1978. Three-dimensional storm motion detection by conventional weather radar. Nature 273, 287–289. http://dx.doi.org/10.1038/ 273287a0.

Sadek, R., Ballester, C., Garrido, L., Meinhardt, E., Caselles, V., 2012. Frame interpolation with occlusion detection using a time coherent segmentation. In: International Conference on Computer Vision Theory and Applications (VISAPP 2012). Roma, Italy.

Sadjadi, F., 2000. Radar beam sharpening using an optimum FIR filter. Circuits, Syst. Signal Process. 19, 121–129. http://dx.doi.org/10.1007/BF01212466.

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. http://dx.doi.org/10.1016/j.advwatres.2011.10.012.

Segond, M.-L., Onof, C., Wheater, H.S., 2006. Spatial–temporal disaggregation of daily rainfall from a generalized linear model. J. Hydrol. 331, 647–689. http:// dx.doi.org/10.1016/j.jhydrol.2006.06.019.

Seo, D., Smith, J., 1991. Rainfall estimation using raingages and radar – a Bayesian approach: 1. Derivation of estimators. Stoch. Hydrol. Hydraul. 5, 17–29. http:// dx.doi.org/10.1007/BF01544175.

Sinclair, S., Pegram, G., 2005. Combining radar and rain gauge rainfall estimates using conditional merging. Atmos. Sci. Lett. 6, 19–22. http://dx.doi.org/10.1002/ asl.85.

Singh, V.P., 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrol. Process. 11, 1649–1669. http://dx.doi.org/10.1002/(SICI)1099-1085(19971015)11:12<1649::AID-HYP495>3.0.CO;2-1.

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 Resour. Res. 45, W04430. http://dx.doi.org/10.1029/2008WR006840.

Sugier, J., Tabary, P., 2006. Evaluation of dual polarization technology at C-band for operational weather radars as part of the EUMETNET OPERA programme. In: Fourth European Conference on Radar in Meteorology and Hydrology – ERAD 2006. Barcelona, Spain.

Sun, D., Roth, S., Black, M.J., 2010. Secrets of optical flow estimation and their principles. Comput. Vis. Pattern Recognit. (CVPR), 2010 IEEE Conf. http://dx.doi. org/10.1109/CVPR.2010.5539939.

Tabary, P., 2007. The new French operational radar rainfall product. Part I: methodology. Weather Forecast. 22, 393–408. http://dx.doi.org/10.1175/WAF1004.1.

Tessier, Y., Lovejoy, S., Schertzer, D., 1993. Universal multifractals: theory and observations for rain and clouds. J. Appl. Meteorol. 32, 223–250. http:// dx.doi.org/10.1175/1520-0450(1993) 032<0223:UMTAOF>2.0.CO;2.

Thorndahl, S., Nielsen, J.E., Rasmussen, M.R., 2014. Bias adjustment and advection interpolation of long-term high resolution radar rainfall series. J. Hydrol. 508, 214–226. http://dx.doi.org/10.1016/j.jhydrol.2013.10.056.

Todini, E., 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. Hydrol. Earth Syst. Sci. 5 (187–199). http:// dx.doi.org/10.5194/hess-5-187-2001, 2001.

Ulbrich, C.W., 1983. Natural variations in the analytical form of the raindrop size distribution. J. Clim. Appl. Meteorol. 22, 1764–1775. http://dx.doi.org/10.1175/ 1520-0450(1983) 022<1764:NVITAF>2.0.CO;2.

Van Horne, M.P., 2003. Short-Term Precipitation Nowcasting for Composite Radar Rainfall Fields. Master Thesis, Massachusetts Institute of Technology.

Vasiloff, S., 2012. Evaluation of Dual-Polarization QPE: Initial Results for Spring and Summer 2012 – Final Report, Memorandum of Understanding Task 1.1. NOAA-NSSL, USA.

Vieux, B.E., Bedient, P.B., 2004. Assessing urban hydrologic prediction accuracy through event reconstruction. J. Hydrol. 299, 217–236. http://dx.doi.org/ 10.1016/j.jhydrol.2004.08.005.

Vieux, B.E., Vieux, J.E., Chen, C., Howard, K.W., 2003. Operational deployment of a physics-based distributed rainfall–runoff model for flood forecasting in Taiwan. In: Tachikawa, Y., Vieux, B.E., Georgakakos, K.P., Nakakita, E. (Eds.), International Symposium on Information from Weather Radar and Distributed Hydrological Modeling, IAHS General Assembly, IAHS Red Book Publication, No. 282. IAHS, Sapporo, Japan, pp. 251–257.

Villarini, G., Smith, J.A., Lynn Baeck, M., Sturdevant-Rees, P., Krajewski, W.F., 2010. Radar analyses of extreme rainfall and flooding in urban drainage basins. J. Hydrol. 381, 266–286. http://dx.doi.org/10.1016/j.jhydrol.2009.11.048.

Wallingford, H.R., 1983. Wallingford Procedure for Design and Analysis of Urban Storm Drainage. Wallingford, UK.

Wang, L.-P., Onof, C., Hsieh, S.-H., Maksimović, C., 2010. Reconstruction of sub-daily rainfall sequences using multinomial multiplicative cascades. Hydrol. Earth Syst. Sci. Discuss. 7, 5267–5297.

Wang, L.-P., Simões, N.E., Rico-Ramirez, M., Ochoa, S., Maksimović, C., 2011a. Radar-based pluvial flood forecasting over urban areas: Redbridge case study. In: International Symposium on Weather Radar and Hydrology (WRaH). Exeter, UK.

Wang, Y., Zhang, J., Ryzhkov, A.V., Tang, L., Howard, K.W., 2011b. Quantitative precipitation estimation using specific differential propagation phase in typhoon system in Taiwan. In: 35th Conference on Radar Meteorology. Pittsburgh, USA.

Wang, L.-P., Onof, C., Ochoa-Rodríguez, S., Simoes, N.E., Maksimović, Č., 2012. On the propagation of rainfall bias and spatial variability through urban pluvial flood modelling. In: 9th Int. Work. Precip. Urban Areas Urban challenges rainfall Anal.

Wang, G., Wong, W., Liu, L., Wang, H., 2013a. Application of multi-scale tracking radar echoes scheme in quantitative precipitation nowcasting. Adv. Atmos. Sci. 30, 448–460. http://dx.doi.org/10.1007/s00376-012-2026-7.

Wang, L.-P., Ochoa-Rodríguez, S., Simões, N.E., Onof, C., Maksimović, C., 2013b. Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. Water Sci. Technol. 68, 737–747.

Wang, L.-P., Ochoa-Rodríguez, S., Onof, C., Willems, P., 2015. Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications. Hydrol. Earth Syst. Sci. Discuss. 12, 1855–1900. http://dx.doi.org/ 10.5194/hessd-12-1855-2015.

WaPUG, 2002. Code of Practice for the Hydraulic Modelling of Sewer Systems. Wastewater Planning Users Group WaPUG, UK.

Weickert, J., Schnörr, C., 2001. Variational optic flow computation with a spatio-temporal smoothness constraint. J. Math. Imag. Vis. 14, 245–255. http:// dx.doi.org/10.1023/A:1011286029287.

Wilson, J.W., Ebert, E.E., Saxen, T.R., Roberts, R.D., Mueller, C.K., Sleigh, M., Pierce, C.E., Seed, A., 2004. Sydney 2000 forecast demonstration project: convective storm nowcasting. Weather Forecast. 19, 131–150. http://dx.doi.org/10.1175/ 1520-0434(2004) 019<0131:SFDPCS>2.0.CO;2.

Zhang, G., Yu, T.-Y., Doviak, R.J., 2005. Angular and range interferometry to refine weather radar resolution. Radio Sci. 40, RS3013. http://dx.doi.org/10.1029/ 2004RS003125.

CrossMark

# Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation

Susana Ochoa-Rodriguez [a,*], Li-Pen Wang [b], Auguste Gires [c], Rui Daniel Pina [a], Ricardo Reinoso-Rondinel [d], Guendalina Bruni [e], Abdellah Ichiba [c,f], Santiago Gaitan [e], Elena Cristiano [e], Johan van Assel [g], Stefan Kroll [g], Damian Murlà-Tuyls [b], Bruno Tisserand [h], Daniel Schertzer [c], Ioulia Tchiguirinskaia [c], Christian Onof [a], Patrick Willems [b], Marie-Claire ten Veldhuis [e]

[a] Urban Water Research Group, Department of Civil and Environmental Engineering, Imperial College London, Skempton Building, London SW7 2AZ, UK
[b] Hydraulics Laboratory, KU Leuven, 3001 Heverlee (Leuven), Belgium
[c] Université Paris-Est, École des Ponts ParisTech, LEESU, 6-8 Av Blaise Pascal Cité Descartes, Marne-la-Vallée, 77455 Cx2, France
[d] Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, PO Box 5048, 2600 GA Delft, The Netherlands
[e] Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, PO Box 5048, 2600 GA Delft, The Netherlands
[f] Conseil Général du Val-de-Marne, Direction des Services de l'Environnement et de l'Assainissement (DSEA), Bonneuil-sur-Marne, 94381, France
[g] Aquafin NV, Dijkstraat 8, 2630 Aartselaar, Belgium
[h] Veolia Environment Research and Innovation, Chemin de la Digue, BP 76, 78603 Maisons Laffitte Cedex, France

## ARTICLE INFO

## SUMMARY

Urban catchments are typically characterised by high spatial variability and fast runoff processes resulting in short response times. Hydrological analysis of such catchments requires high resolution precipitation and catchment information to properly represent catchment response. This study investigated the impact of rainfall input resolution on the outputs of detailed hydrodynamic models of seven urban catchments in North-West Europe. The aim was to identify critical rainfall resolutions for urban catchments to properly characterise catchment response. Nine storm events measured by a dual-polarimetric X-band weather radar, located in the Cabauw Experimental Site for Atmospheric Research (CESAR) of the Netherlands, were selected for analysis. Based on the original radar estimates, at 100 m and 1 min resolutions, 15 different combinations of coarser spatial and temporal resolutions, up to 3000 m and 10 min, were generated. These estimates were then applied to the operational semi-distributed hydrodynamic models of the urban catchments, all of which have similar size (between 3 and 8 km²), but different morphological, hydrological and hydraulic characteristics. When doing so, methodologies for standardising model outputs and making results comparable were implemented. Results were analysed in the light of storm and catchment characteristics. Three main features were observed in the results: (1) the impact of rainfall input resolution decreases rapidly as catchment drainage area increases; (2) in general, variations in temporal resolution of rainfall inputs affect hydrodynamic modelling results more strongly than variations in spatial resolution; (3) there is a strong interaction between the spatial and temporal resolution of rainfall input estimates. Based upon these results, methods to quantify the impact of rainfall input resolution as a function of catchment size and spatial–temporal characteristics of storms are proposed and discussed.

* Corresponding author. Tel.: +44 (0)20 7594 6018.
E-mail addresses: s.ochoa-rodriguez@imperial.ac.uk (S. Ochoa-Rodriguez), Lipen.Wang@bwk.kuleuven.be (L.-P. Wang), auguste.gires@leesu.enpc.fr (A. Gires), r.pina13@imperial.ac.uk (R.D. Pina), r.r.reinosorondinel@tudelft.nl (R. Reinoso-Rondinel), G.Bruni@tudelft.nl (G. Bruni), abdellah.ichiba@leesu.enpc.fr (A. Ichiba), S.Gaitan@tudelft.nl (S. Gaitan), E.Cristiano@tudelft.nl (E. Cristiano), johan.vanassel@aquafin.be (J. van Assel), stefan.kroll@aquafin.be (S. Kroll), Damian.MurlaTuyls@bwk.kuleuven.be (D. Murlà-Tuyls), bruno.tisserand@veoliaeau.fr (B. Tisserand), Daniel.Schertzer@enpc.fr (D. Schertzer), ioulia.tchiguirinskaia@enpc.fr (I. Tchiguirinskaia), c.onof@imperial.ac.uk (C. Onof), Patrick.Willems@bwk.kuleuven.be (P. Willems), j.a.e.tenveldhuis@tudelft.nl (M.-C. ten Veldhuis).

## 1. Introduction

The impact of spatial–temporal variability of rainfall on catchment response and the sensitivity of hydrological models to the spatial–temporal resolution of rainfall inputs have been active topics of research over the last few decades (e.g. Singh, 1997; Berndtsson and Niemczynowicz, 1988; Lobligeois et al., 2014). Several studies have shown that the spatial–temporal variability of rainfall fields can translate into large variations in flows; as a result, it is necessary to account for this variability in order to

properly characterise hydrological response (Tabios and Salas, 1985; Berndtsson and Niemczynowicz, 1988; Krajewski et al., 1991; Obled et al., 1994; Singh, 1997; Chaubey et al., 1999; Arnaud et al., 2002; Syed et al., 2003; Smith et al., 2004; Kavetski et al., 2006). This is particularly the case in small urban catchments, which are characterised by fast runoff processes and short response times, and are therefore very sensitive to the spatial and temporal variability of precipitation (this variability was found to be significant even at the small scales of urban catchments (Emmanuel et al., 2012; Gires et al., 2014b)). In order to well represent urban runoff processes, high resolution precipitation information is therefore needed (Schilling, 1991; Faurès et al., 1995; Shah et al., 1996; Aronica and Cannarozzo, 2000; Einfalt, 2005; Tetzlaff and Uhlenbrook, 2005; Segond et al., 2007; Vieux and Imgarten, 2012; Schellart et al., 2012). This need has been further fuelled by recent developments in, and increasing use of, higher-resolution urban hydrological models (e.g. Fewtrell et al., 2011; Giangola-Murzyn et al., 2012; Pina et al., 2014), which allow incorporation of detailed rainfall, surface and runoff information. With regards to rainfall monitoring, significant progress has been made over the last few decades, including widespread increase in the use of weather radar rainfall estimates, generally provided by national meteorological services at 1 km/5–10 min resolutions. Multiple studies have been conducted in recent years aimed at analysing urban hydrological/hydraulic model sensitivity to the spatial–temporal resolution of rainfall inputs and at establishing required rainfall input resolutions for urban hydrological applications. However, there is not as yet a consensus on these topics.

A theoretical study undertaken by Schilling (1991) suggested that, for urban drainage modelling, rainfall data of at least 1–5 min and 1 km resolutions should be used. Another study undertaken by Fabry et al. (1994) suggested that finer resolution data (i.e. 1–5 min in time and 100–500 m in space) are required for urban hydrological applications. This however may vary according to the application (Einfalt et al., 2004; Einfalt, 2005); for detailed sewer system simulation, for example, it is believed that the spatial–temporal resolutions suggested in Fabry et al. (1994) are essential.

Berne et al. (2004) analysed the relation between catchment size and minimum required spatial and temporal resolutions or rainfall measurements in a study involving very high resolution precipitation data (∼7.5 m/4 s) and runoff records from six urban catchments on the French Mediterranean coast (but not models were used). Their study suggests that for small urban catchments, of the order of 3 ha, ∼1.5 km/1 min resolution, rainfall estimates are recommended, whereas for larger catchments, of the order of 500 ha, ∼3 km/5 min estimates may suffice. Slightly more stringent resolution requirements were identified by Notaro et al. (2013): using high spatial–temporal resolution rain gauge records as input to the semi-distributed urban drainage model of a 700 ha urban catchment in Italy, the authors investigated the uncertainty in runoff estimates resulting from coarser resolution rainfall inputs and concluded that temporal resolutions below 5 min and spatial resolutions of ∼1.7 km are generally required for urban hydrological applications.

Using a semi-distributed urban drainage model of a small urban catchment in London, and stochastically-downscaled rainfall estimates, Gires et al. (2012) and Wang et al. (2012) showed that the unmeasured small-scale rainfall variability, i.e. occurring below the typically available resolutions of 1 km in space and 5 min in time, may have a significant impact on simulated flows, with the impact decreasing as the drainage area of interest increases. A similar study was undertaken by Gires et al. (2014a), but this time using a fully-distributed model of a small catchment in Paris; similar results were obtained, but the fully-distributed model displayed higher sensitivity to the resolution of rainfall inputs.

More recently, Bruni et al. (2015) analysed the relationship between spatial and temporal resolution of rainfall input, storm and catchment scales, urban hydrodynamic model properties and modelling outputs. This was done using high resolution (100 m/1 min) rainfall data provided by polarimetric weather radar and a semi-distributed urban drainage model of a subcatchment in Rotterdam, the Netherlands. They showed that for a densely built, highly impervious urban catchment, modelling outputs are sensitive to high resolution rainfall variability and that deviations in model outputs significantly increase as rainfall inputs are aggregated to coarser scales, particularly at very small drainage areas (<1 ha).

As can be seen, few studies have analysed measured spatial–temporal variability of rainfall at the 1 min and 100 m scales and those which have not always involved hydrological/hydraulic models and/or are limited to single catchment studies. Hence, evidence to prove the added value of higher resolution rainfall estimates and to provide an answer about actual resolution requirements for urban hydrological applications is still insufficient. With the purpose of providing additional evidence in this direction, the present study investigates the impact of rainfall input variability for a range of spatial and temporal resolutions on the hydrodynamic modelling outputs of seven urban catchments located in each of the partner countries of the European Interreg RainGain project (http://www.raingain.eu/) (i.e. UK, France, Netherlands and Belgium). Rainfall estimates of nine storm events were derived from a polarimetric X-band radar located in Cabauw (The Netherlands). The original radar estimates, at 100 m and 1 min resolutions, were aggregated to spatial resolutions of 500, 1000 and 3000 m, and were sampled at temporal resolutions of 1, 3, 5 and 10 min. These estimates were then applied to high-resolution semi-distributed hydrodynamic models of the seven urban catchments, all of which have similar size (between 3 and 8 km²), but different morphological, land use and model structure characteristics. Within the catchments, outputs were analysed at different nodes along the main flow path to investigate the effect of drainage areas of different sizes. Methodologies for standardising rainfall inputs and hydrological outputs were implemented to make results comparable. The impact of varying spatial–temporal resolutions of rainfall input on hydrodynamic model outputs was analysed in the light of storm and catchment characteristics. Based upon these results, current research needs and future work are discussed.

The paper is organised as follows. In Section 2, the pilot catchments, hydrodynamic models and radar-rainfall datasets are introduced. Methodologies for selecting relevant spatial–temporal resolution combinations and characterising spatial–temporal characteristics of the nine storms events are explained in Section 3, as well as methodologies used for feeding the rainfall inputs into the hydrodynamic models of the pilot catchments and for extracting and analysing the hydrodynamic modelling results. Results are presented and discussed in Section 4, followed by conclusions and recommendations in Section 5.

## 2. Pilot catchments and datasets

### 2.1. Pilot urban catchments

Seven urban catchments, located in four North-West European countries, were adopted as pilot locations in this study. With the aim of facilitating inter-comparison of results, catchment areas of similar size (3–8 km²) were selected for testing. The main characteristics of the selected pilot catchments are summarised in Table 1. Moreover, images of the boundaries and sewer layouts of all pilot catchment can be found in Fig. 1. More detailed

**Table 1**
Summary characteristics of selected pilot urban catchments.

| | Cranbrook (UK) | Torquay Town Centre (UK) | Morée-Sausset (FR) | Sucy-en-Brie (FR) | Herent (BE) | Ghent (BE) | Kralingen (NL) |
|---|---|---|---|---|---|---|---|
| Catchment ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Area (ha) | 865 | 570 | 560 | 269 | 512 | 649 | 670 |
| Catchment length and width (km)[*] | 6.10/1.42 | 5.35/1.06 | 5.28/1.06 | 4.02/0.67 | 8.16/0.63 | 4.74/1.37 | 2.12/3.16 |
| Catchment shape factor (−)[**] | 0.23 | 0.20 | 0.20 | 0.17 | 0.08 | 0.29 | 1.49 |
| Slope (m/m)[***] | 0.0093 | 0.0262 | 0.0029 | 0.0062 | 0.0083 | 0.0001 | 0.0003 |
| Main flow direction (°) | 239 | 270 | 198 | 138 | 40 | 235 | 152 |
| Type of drainage system | Mostly separate, branched | Mostly combined, branched | Mostly separate, branched | Separate, branched | Mostly combined, branched | Mostly combined, branched | Mostly combined, looped |
| Is flow mainly driven by gravity? | Yes | Yes | Yes | Yes | Yes | Yes | No |
| Control elements | 3 storage lakes | 3 storage tanks, 1 pumping station | 2 storage tanks | 1 storage basin, 1 pumping station | 5 main CSO's with control | 15 pumping stations | 20 pumping stations |
| IMP (%)[****] | 52% | 26% | 37% | 34% | 27% | 41% | 48% |
| Predominant land use[*****] | R&C | R&C | R&C | R&C | R | R | R&C |
| Population density (per/ha) | 47 | 60 | 70 | 95 | 20 | 24 | 154 |

[*] Length = Length of longest flow path (through sewers) to catchment outfall; Width = Catchment Area/Catchment Length.
[**] Shape factor = Width/Length (this parameter is lower for elongated catchments).
[***] Catchment slope = Difference in ground elevation between upstream most point and outlet/catchment length.
[****] IMP: total proportion of impervious areas in relation to total catchment area.
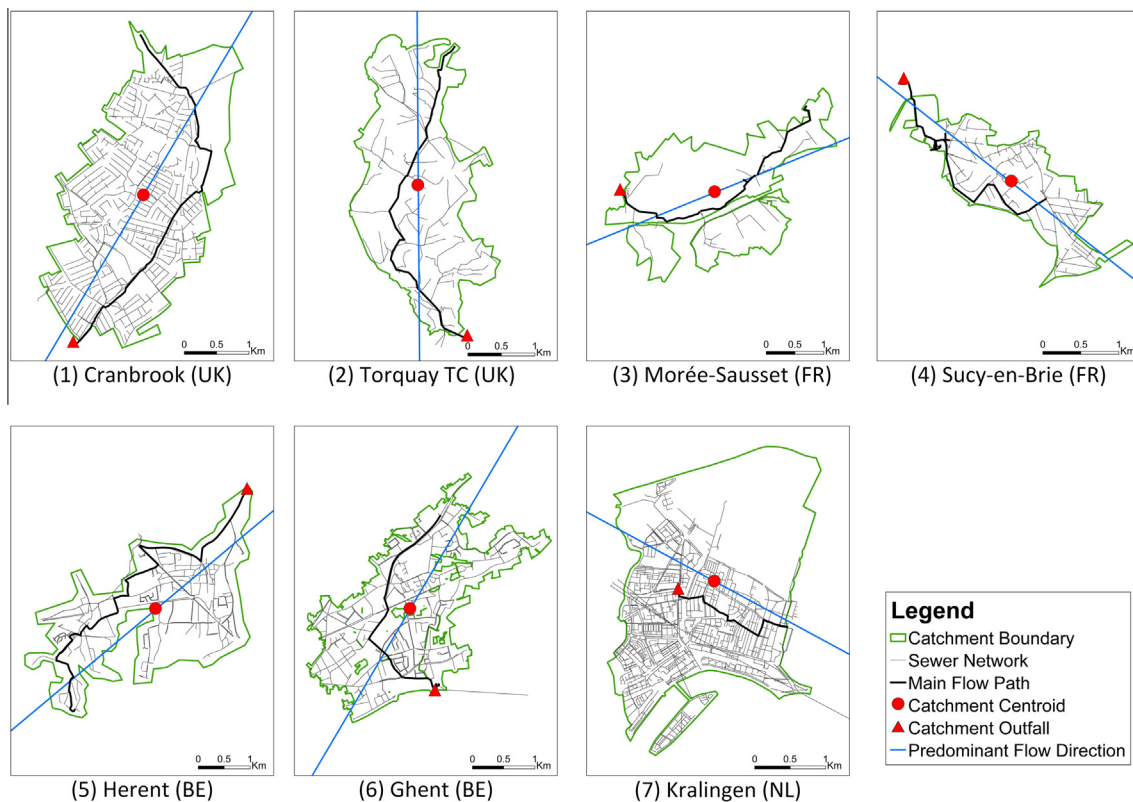[*****] Predominant land use: R = residential; C = commercial.



**Fig. 1.** Catchment boundary and sewer layout for the pilot urban catchments.

**Table 2**
Summary characteristics of the hydrodynamic models of the seven pilot catchments.

| | Cranbrook, UK | Torquay Town Centre, UK | Morée-Sausset, FR | Sucy-en-Brie, FR | Herent, BE | Ghent, BE | Kralingen, NL |
|---|---|---|---|---|---|---|---|
| Total pipe length* (km) | 98.05 | 41.29 | 15.30 | 4.02 | 67.42 | 83.44 | 142.65 |
| Number of SC** | 1765 | 492 | 47 | 9 | 683 | 1424 | 2435 |
| Mean/Median/STD of SC size (ha) | 0.49/0.37/0.71 | 1.16/0.93/1.09 | 11.92/8.00/10.34 | 29.89/13.07/27.47 | 0.71/0.34/1.27 | 0.46/0.23/0.89 | 0.12/0.09/0.13 |
| Mean/Median/STD of SC slopes (m/m) | 0.0560/0.0540/0.0334 | 0.0670/0.0470/0.0973 | 0.0056/0.005/0.0037 | 0.017/0.00575/0.0298 | 0.0221/0.0100/0.0282 | 0.0349/0.002/0.0459 | 0.0513/0.00812/0.0283 |
| Mean/Median/STD of pipe slopes (m/m) | 0.0103/0.0062/0.0138 | 0.0765/0.0488/0.1180 | 0.0062/0.0034/0.0141 | 0.0116/0.006/0.0145 | 0.0155/0.0073/0.0241 | 0.0030/0.00179/0.0078 | 0.0003/0.0000/0.0130 |
| Rainfall–runoff volume estimation model | Fixed runoff coefficient for IMP/NewUK for PER | Wallingford model for IMP and PER | PER: Initial loss + runoff coefficient depending on rainfall depth and soil type. IMP: directly connected or not to the sewer network | Single linear reservoir | Fixed runoff coefficient for all surfaces (0.8 for IMP; 0 for PER except some special ones) | Double linear reservoir | Initial loss + fixed runoff coefficient for IMP; Horton's model for PER |
| Runoff routing model at SC | Double linear reservoir | Double linear reservoir | Single linear reservoir | Single linear reservoir | Double linear reservoir | Double linear reservoir | Runoff delay coefficient |
| Pipe flow routing model | Dynamic wave (full de St Venant Equations) | | | | | | |
| Modelling software | InfoWorks CS 14 | InfoWorks CS 14 | Canoe 3.5 | Canoe 3.5 | InfoWorks CS 14 | InfoWorks CS 13 | Sobek-Urban |

\* Pipe length is estimated based upon modelled pipes only.
\*\* SC = sub-catchment.

information on each of these catchments can be found in the RainGain project website: http://www.raingain.eu/en/actualite/learn-more-about-ten-locations-where-raingain-solutions-will-be-implemented. As can be seen, the selected pilot catchments cover a wide range of morphological, topographic and land use conditions.

### 2.2. Urban drainage models of the pilot catchments

Verified and operational semi-distributed urban drainage models of each catchment were used in this study; their main characteristics are summarised in Table 2. In this type of models the whole catchment surface is split into sub-catchment units through which rainfall is applied. Each sub-catchment unit is treated as a lumped model within which rainfall is assumed to be uniform. Each sub-catchment comprises a mix of pervious (PER) and impervious (IMP) surfaces the runoff of which drains to a common outlet point, which corresponds to an inlet node of the sewer system (i.e. a gully or a manhole). Each sub-catchment is characterised by a number of parameters, including total area, length, slope and proportion of each land use, amongst others. Based upon these parameters, runoff volumes are estimated and routed at subcatchment scale using the rainfall–runoff and concentration models commonly employed in each country (see Table 2). Sub-catchment sizes of the models used in this study typically varied from 0.09 ha to 13.07 ha (median values). Sewer flows in all pilot catchment models are routed using the full de St. Venant equations (i.e. dynamic wave approximation).

### 2.3. High resolution precipitation data

High-resolution rainfall data were obtained by a dual-polarimetric X-band weather radar, IDRA hereafter, located in the CESAR observatory of the Netherlands (Figueras i Ventura, 2009; Leijnse et al., 2010). IDRA is a frequency modulated continuous wave (FMCW) radar working at 9.475 GHz. Its operational range is of 15 km with a range resolution of 30 m, approximately. IDRA is fixed at a height of 213 m from ground level; it scans at a fixed elevation angle of 0.5°, and rotates the antenna over 360° every minute. The technical specifications of IDRA are summarised in Table 3.

The accuracy of radar measurements can be affected by multiple factors, including clutter contamination and signal attenuation. In order to ensure good quality of the final radar product, several correction procedures were implemented; these are summarised next.

Signals of ground and moving clutter were identified and removed, using an optimum filter based on polarimetric spectra (Unal, 2009). Moreover, random fluctuations were separated from weather signals using a threshold of 3 dB above noise level. In addition, areas with linear depolarisation ratio ($L_{dr}$) larger than $-15$ dB were removed to ensure only rain particles are processed. Because IDRA works at X-band frequencies, received signals can

**Table 3**
Specifications of dual-polarimetric X-band weather radar IDRA from which high resolution precipitation data were derived for this study.

| | |
|---|---|
| Radar type | FMCW |
| Polarisation | Dual polarisation |
| Frequency | 9.475 GHz |
| Range resolution | 3–30 m |
| Min range | 230 m |
| Max range | <122 km |
| Max unambiguous radial velocity | 19 m/s |
| Temporal resolution | 1 min |
| Beamwidth | 1.8° |
| Elevation | 0.5° |

experience large attenuation and as a result radar measurements such as reflectivity ($Z$) can be underestimated. However, the specific differential phase ($K_{dp}$) is immune to attenuation and therefore $K_{dp}$ was used to correct reflectivity from attenuation effects as long as the received signals were not totally extinct (Otto and Russchenberg, 2011). Areas with extinct signals are typically located behind regions with heavy precipitation. In the implemented processing routines extinct areas were flagged and excluded from further processing.

$K_{dp}$ is also immune to radar calibration errors and hail contamination. This makes $K_{dp}$ suitable for rainfall rate estimation. However, $K_{dp}$ at X-band frequencies can be contaminated by the backscattering component of the differential phase, which can introduce bias. In addition, with the purpose of maintaining low $K_{dp}$ variability, $K_{dp}$ is typically obtained at spatial resolutions of the order of 2–3 km (Bringi and Chandrasekar, 2001), which can be few times larger than the radar range resolution. Nonetheless, the approach by Otto and Russchenberg (2011), adopted in the present study, addresses both issues. First, the effect of the backscattering component is filtered out by using a theoretical relationship between the backscattering and the differential reflectivity. Second, $K_{dp}$ is obtained at radar spatial resolution by using the self-consistency principle (Scarchilli et al., 1996).

Given the above considerations, for the present study rainfall rate ($R$) is estimated using $K_{dp}$ for areas with $Z > 30$ dBZ, otherwise the corrected reflectivity is used according to Otto and Russchenberg (2012):

$$R = 13K_{dp}^{0.75} \tag{1}$$

$$z = 243R^{1.24} \tag{2}$$

where $R$, $K_{dp}$, and $z$ are given in mm h$^{-1}$, degrees km$^{-1}$, and mm$^6$ m$^{-3}$, respectively. Although these steps improve the estimation of rainfall rate, there remain issues such as insect echoes, melting-layer contamination, and multi-trip echoes. Each of these echoes has a familiar pattern which can be detected through visual inspection: insects are noticeable at short ranges, at which radar reflectivity is highly sensitive; melting-layer contamination leads to strong echoes in the form of a ring around the radar; and multi-trip echoes can be identified in the reflectivity field by lengthened and weak echo lines. The data used in the present study were visually inspected to ensure that the effect of contamination by undesired echoes was minimal.

Rainfall estimates from IDRA were initially available in polar coordinates at temporal and spatial resolutions of 1 min and 30 m by 1.8° (i.e. radar beamwidth), respectively. However, to facilitate handling of the data, it had to be converted from polar to Cartesian coordinates. In this work, data were initially mapped to a regular grid of 100 m by 100 m; this is therefore the finest spatial resolution used as input for the urban drainage models in the present study. From the available IDRA dataset, eight storm events recorded between 2011 and 2014 were selected for this study. The selected events correspond to the most intense events recorded during these years and can be considered characteristic of North-West Europe. Nonetheless, it is worth mentioning that, being a research radar, IDRA does not operate continuously; therefore, not all intense storm events which occurred between 2011 and 2014 were recorded by the radar and the selected events include a combination of high intensity as well as moderate and low intensity storms. For each storm event a square area of 36 km$^2$, which is large enough to circumscribe the eight pilot catchments (considering their different shapes), was clipped from the total area covered by the radar and was used as input for the models of the pilot catchments. The area for analysis was selected such that it comprised the main rainfall cell(s) observed within the radar domain. The dates and main statistics of the selected storm events within the clipped (6 km × 6 km) area are summarised in Table 4. It is important to note that during the storm event on 18/01/2011, strong storm cells were observed in different areas of the radar domain. Given the high intensities and depths associated with the different areas, it was deemed appropriate to select two different areas within the radar domain for analysis. Consequently, for this storm event two sub-events were selected for analysis (i.e. E1 and E2). Storm profiles, snapshot images during the time of peak areal intensity as well as images of the rainfall depth accumulations for each storm event within the clipped area are shown in Fig. 2.

## 3. Methodology

### 3.1. Selection of rainfall input resolutions for analysis

To study the impact of spatial–temporal resolution of rainfall inputs on hydrodynamic model outputs, sixteen combinations of spatial–temporal resolutions were selected. The highest resolution of 100 m in space and 1 min in time was used as reference. Additionally, 15 resolution combinations were adopted based on the following considerations (the rationale behind the selected resolution combinations, as well as the selected combinations, are summarised in Fig. 3):

- In the framework of the simplest space–time scaling model that relies on a scaling anisotropy coefficient $H_t$ (Deidda, 2000; Gires et al., 2011): when the spatial scale of the data is changed by a ratio of $\lambda_{xy}$, the temporal scale should be changed by a factor of $\lambda_t = \lambda_{xy}^{1-H_t}$. By combining the scale invariance property of Navier–Stokes equations with Kolmogorov's (1962) formulation, and assuming that the properties established for the atmosphere remain valid for rainfall, it is possible to show that $H_t$ is expected to be equal to 1/3 (Marsan et al., 1996). This means that when the spatial scale is multiplied by 3, the temporal scale should be multiplied by 2 (i.e. $3^{1-1/3} \approx 2.08$) (Biaou et al., 2005;

**Table 4**
Characteristics of selected storm events (estimated based upon 1 min/100 m resolution estimates for the clipped (6 km × 6 km) area). Time is in UTC.

| Event ID | Date | Duration | Total depth (areal average/pixel min/pixel max) (mm) | Max intensity over 1 min (areal average/individual pixel) (mm/h) |
|---|---|---|---|---|
| E1 | 18/01/2011 | 05.10–08.00 h | 31.48/17.89/45.82 | 31.67/1120.20 |
| E2 | 18/01/2011 | 05.10–08.00 h | 36.12/16.48/47.17 | 26.48/124.00 |
| E3 | 28/06/2011 | 22.05–23.55 h | 8.94/4.46/17.64 | 28.42/241.82 |
| E4 | 18/06/2012 | 05.55–07.10 h | 10.12/8.03/11.76 | 11.62/24.11 |
| E5 | 29/10/2012 | 17.05–19.00 h | 5.34/1.20/13.64 | 7.05/82.83 |
| E6 | 02/12/2012 | 00.05–03.00 h | 4.94/2.39/7.86 | 6.59/38.57 |
| E7 | 23/06/2013 | 08.05–11.30 h | 4.19/0.73/13.39 | 9.41/306.55 |
| E8 | 09/05/2014 | 18.15–19.35 h | 4.48/1.40/8.88 | 12.98/66.76 |
| E9 | 11/05/2014 | 19.05–23.55 h | 5.99/1.22/12.65 | 10.53/246.74 |
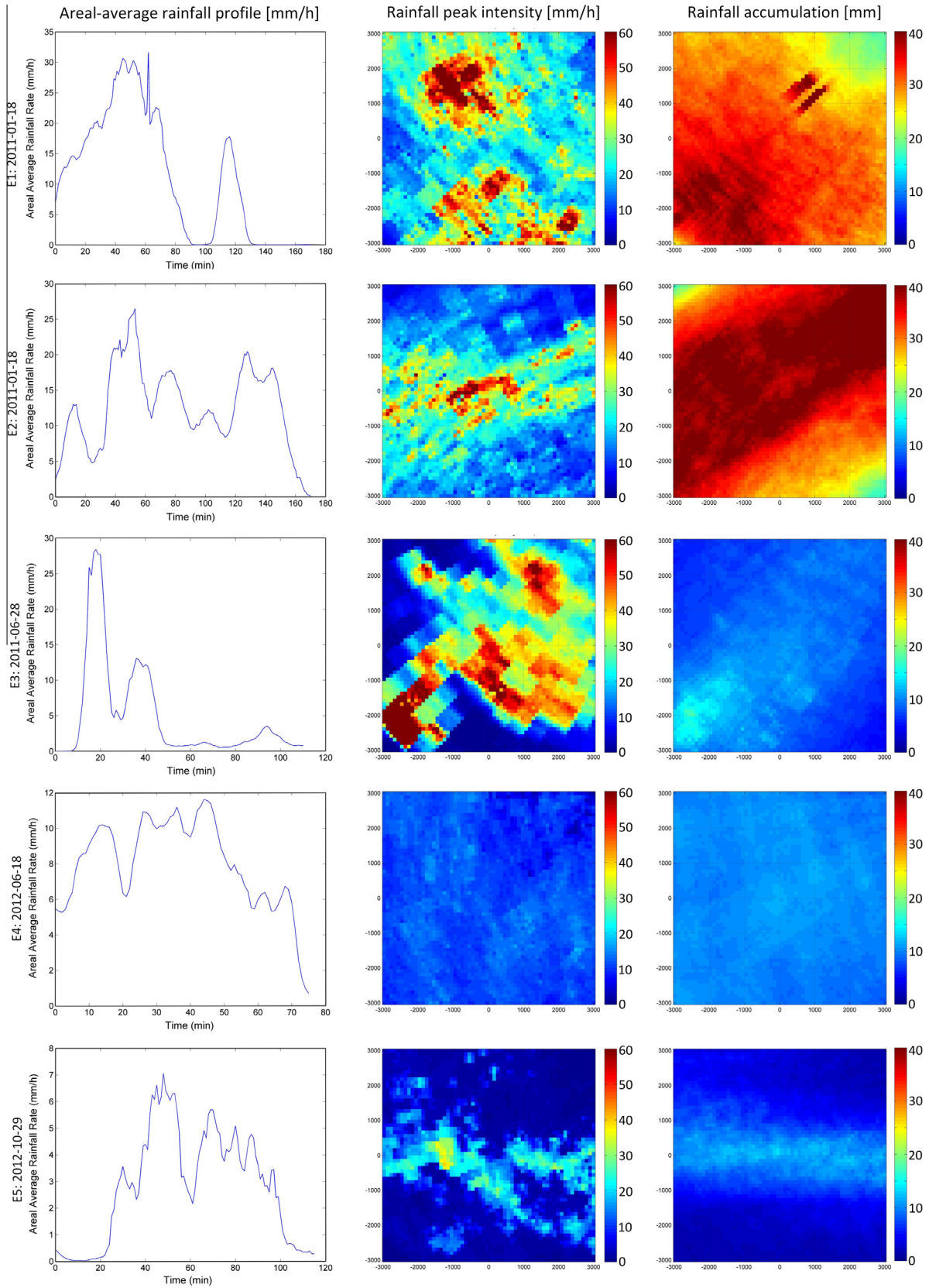
**Fig. 2.** Areal average storm intensity profile (left column), snapshot image during the peak intensity period of the storm (middle column) and total event accumulations for the storm events under consideration.
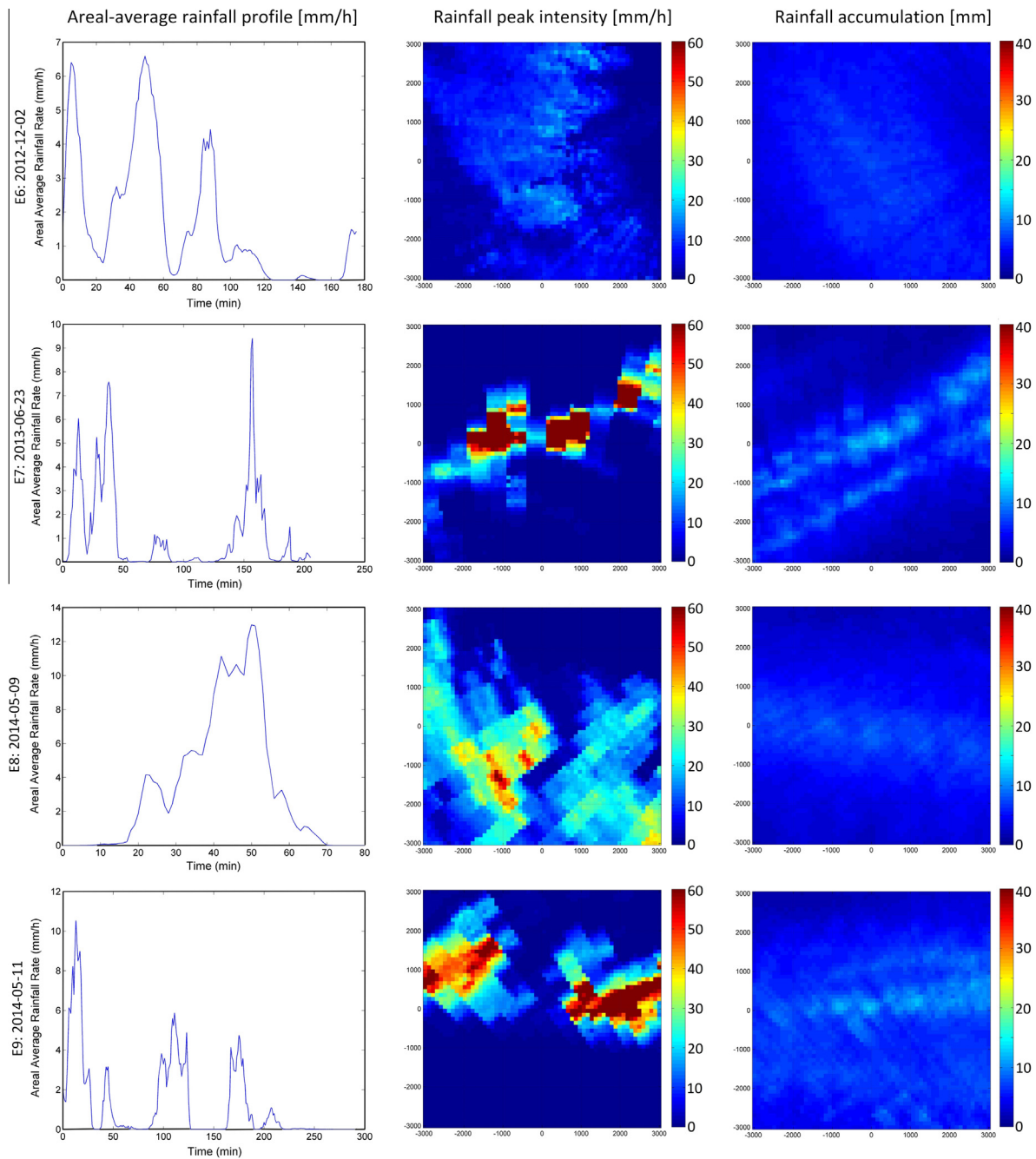
**Fig. 2** (*continued*)

Gires et al., 2012). This leads to the following resolution combinations (indicated in blue in Fig. 3), upscaling from the reference resolution 100 m/1 min: 500 m/3 min; 1000 m/5 min; 3000 m/10 min.

- Operational resolutions: it is of interest to relate the results of this study to resolutions typically available from operational radar networks. The most common resolutions are 1000 m/5 min for national weather radar networks (e.g. in the UK, France, Netherlands, US). Other operational resolutions include: 1000 m/10 min (Malaysia), ~ 500 m/5 min (Belgium). Moreover, the equivalent resolutions of operational urban rain gauge networks are often of the order of several km in space and 1–15 min in time (WAPUG, 2002; Wang et al., 2013). The operational resolutions are indicated in yellow in Fig. 3.

- Berne et al. (2004) identified characteristic temporal and spatial scales relevant to describe the hydrological behaviour of urbanised catchments. They used a simple power law relationship to link lag time to the surface area of catchments. Based on this power law and on the characteristic spatial and temporal dimensions of storms typical of Mediterranean regions, the following approximate characteristic spatial–temporal resolutions were derived (indicated in green in Fig. 3): 1 min/1500 m (for catchment areas ~2.6 ha); 3 min/2600 m (for catchment areas ~100 ha); 5 min/3300 m (for catchment areas ~560 ha); 10 min/4700 m (for catchment area ~5600 ha).

- In addition to the resolution combinations mentioned in the literature based on atmospheric processes and catchment response characteristics, all remaining combinations of the
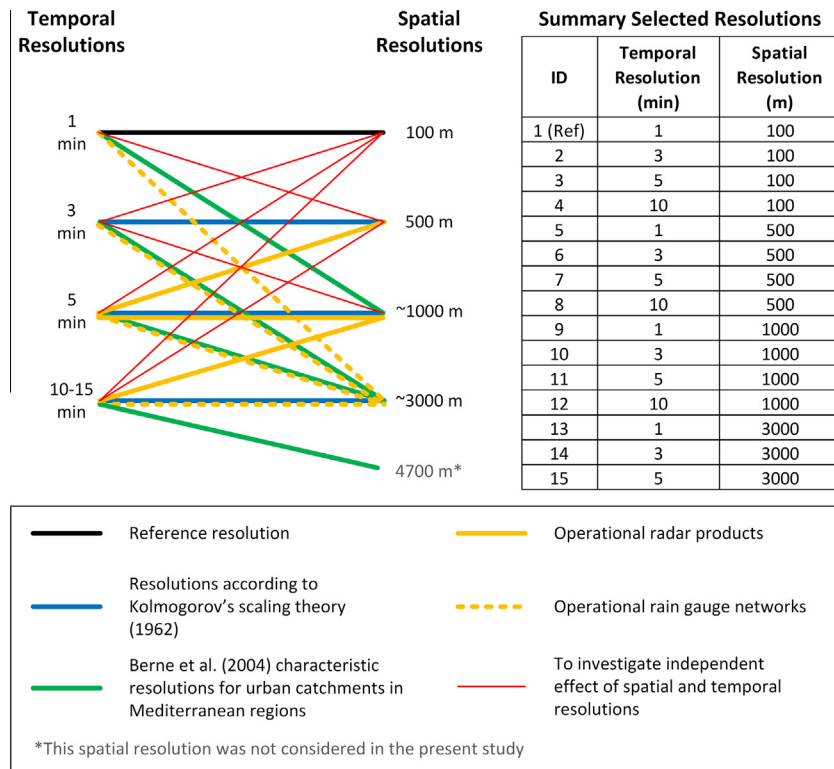
**Fig. 3.** Combinations of space and time resolutions of rainfall inputs investigated in this study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article).

selected space and time scales were investigated, so as to enable the analysis of the 'marginal' as well as the combined effect of the different temporal and spatial resolutions (these are indicated in red in Fig. 3).

Using the finest resolution rainfall estimates (i.e. 100 m/1 min) as starting point, coarser spatial resolutions of up to 3000 m were generated through aggregation (i.e. averaging in space), and coarser temporal resolutions of up to 10 min were obtained by sampling a radar image at the desired time interval. The strategy to generate coarser temporal resolution estimates was chosen so as to replicate radar scanning strategies.

### 3.2. Spatial and temporal characterisation of storm events

Based upon the finest resolution rainfall data (i.e. 100 m/1 min), the following parameters were estimated which provide a measure of the spatial and temporal characteristics of the storm events under consideration. These parameters are used in Section 4 to analyse the observed impact of rainfall input resolution on hydrodynamic modelling results.

In the estimation of these parameters, only the (manually-selected) radar images over the peak period of the storm (i.e. period during which the core of the storm passes through the 6 km $\times$ 6 km clipped area) were considered. Including all radar images in the estimation would result in smooth parameters which do not reflect the dynamic and critical spatial–temporal features of the storm events, hence the analysis was conducted over the peak period only. It is worth noting that some of the storm events under consideration comprised more than one peak; when this was the case, each of the peaks was analysed separately and the peak with the most stringent characteristics and resolution requirements was adopted as representative of the storm event.

### 3.2.1. Spatial structure of storms and theoretically-required spatial resolution of rainfall inputs

A climatological variogram (Bastin et al., 1984; Berne et al., 2004; Bruni et al., 2015) was employed in this study to characterise the average spatial structure of rainfall fields over the peak storm period. Based upon the range of the variogram ($r$), which represents the limit of spatial dependence (Atkinson and Aplin, 2004), the integral range measure ($A$) (Lantuéjoul, 1991, 2002) was derived which can be considered as the mean area of the spatial structure captured by the radar images over the area of interest. Based upon $A$ and following recommended signal/response requirements from communication theory (Shannon, 1948; Garrigues et al., 2006), a theoretically-required spatial resolution was estimated for each storm event under consideration.

The specific steps that were followed to obtain these parameters are the following:

(1) An empirical isotropic (semi-) variogram ($\gamma(h)$) was computed at each time step as:

$$\gamma(h) = \frac{1}{2n} \sum_i^n \left[ (Z(\mathbf{x}) - Z(\mathbf{x}+h))^2 \right] \tag{3}$$

where $n$ is the number of all pairs of radar pixels separated by a distance $h$, $Z$ are the rainfall rate values at the respective pixels and $\mathbf{x}$ corresponds to the centre of a given radar pixel.

(2) Each empirical variogram was normalised by dividing it by the sample variance.

(3) The normalised variograms obtained for each time step were averaged over the time period of analysis; this yields a climatological empirical variogram.

(4) An exponential variogram model was fitted to the empirical climatological variogram using weighted least square fitting (WLS). The exponential variogram function is the following:

$$\gamma(h) = C_0 + C\left[1 - \exp\left(-\frac{3 \cdot |h|}{r}\right)\right] \tag{4}$$

where $C_0$ is the nugget, $C$ is the sill, and $r$ is the (practical) spatial range at which 95% of the sill is reached. It is worth noting that the two classical models that are used to fit climatological variograms are the exponential and spherical ones. For the storms under consideration both models were tested and a better fitting was generally obtained for the exponential one, hence it was adopted to describe the structure of the variogram.

(5) The integral range measure $(A)$ was estimated as (Lantuéjoul, 1991, 2002):

$$A = \int_{h \in \mathcal{R}^2}\left(1 - \frac{\gamma(h)}{\sigma^2}\right)dh \tag{5}$$

where $\sigma^2$ is the variance and $\mathcal{R}^2$ is the 2-dimensional domain over which the variogram was derived. In simple terms, $A$ corresponds to the area under the correlogram curve. For an exponential variogram model $A$ is given by:

$$A = \frac{2\pi r^2}{9} \tag{6}$$

This measure summarises the (spatial) structural information of the variogram provided by the range and the fraction of total variance. As mentioned above, $A$ can be considered as the mean area of the spatial structure captured by the radar images over the area of interest.

(6) The characteristic length scale of the storm event $(r_c)$, which represents the mean extent of the spatial structure captured by the data (Garrigues et al., 2008), was estimated as the square root of $A$. For an exponential variogram model, $r_c$ is given by:

$$r_c = \left(\frac{\sqrt{2\pi}}{3}\right)r \approx 0.836r \tag{7}$$

(7) In a study focusing on the quantification of the spatial heterogeneity of landscapes, Garrigues et al. (2006) demonstrated that by adopting a maximum pixel size equal to half of the characteristic length of the landscape image (i.e. $r_c/2$), it is possible to capture the major part of the spatial variability of land use. Their derivation followed Shannon's (1948) theorem, according to which the proper sampling frequency of a signal must be higher than twice the maximal frequency of this signal. Following Garrigues et al. (2006) approach, the coarsest spatial resolution $(\Delta s_r)$ that is required to properly characterise a given storm event is therefore given by half of the characteristic length scale. For an exponential variogram:

$$\Delta s_r = \frac{r_c}{2} \approx 0.418r \tag{8}$$

In the case of a spherical variogram model, such as that used by Berne et al. (2004), $\Delta s_r = \frac{r_c}{2} \approx 0.396r$, where the ratio 0.396 is similar to the 1/3 ratio adopted by Berne et al. (2004), though it was derived with a different rationale.

### 3.2.2. Storm direction and velocity

Storm motion was estimated using the TREC (TRacking Radar Echoes by Correlation) method (Rinehart and Garvey, 1978), which is widely used in rainfall nowcasting (Tuttle and Foote, 1990; Laroche and Zawadzki, 1995; Horne, 2003; Li and Lai, 2004). This method analyses the cross-correlation of each two consecutive rainfall fields in order to derive a field of movement vectors (i.e. the displacements in easting and northing directions). Given that

**Table 5**
Estimated spatial and temporal characteristics and required rainfall input resolution for the storm events under consideration.

| Event ID | Spatial range $(r)$ (m) | Mean velocity $(|\overline{v}|)$ (m/s) | Max. observable singularity (Small/Large)[*] $(\gamma_s)$ (−) | Required spatial resolution $(\Delta S_r)$ (m) | Required temporal resolution $(\Delta t_r)$ (min) |
|---|---|---|---|---|---|
| E1 | 4056.69 | 9.76 | 0.33/0.23 | 1694.77 | 5.79 |
| E2 | 3524.76 | 9.91 | 0.33/0.23 | 1472.54 | 4.95 |
| E3 | 4655.10 | 14.04 | 0.53/0.27 | 1944.77 | 4.62 |
| E4 | 3218.91 | 11.71 | 0.62/0.37 | 1344.77 | 3.83 |
| E5 | 2061.98 | 14.11 | 0.66/0.44 | 861.43 | 2.03 |
| E6 | 3737.52 | 11.68 | 0.59/0.33 | 1561.43 | 4.46 |
| E7 | 1702.93 | 13.95 | 0.92/0.50 | 711.43 | 1.70 |
| E8 | 3644.43 | 18.40 | 0.55/0.24 | 1522.54 | 2.76 |
| E9 | 2354.53 | 16.97 | 0.80/0.36 | 983.66 | 1.93 |

[*] $\gamma_s$ values were estimated for two ranges of scales, for which scale invariance was found through multifractal analysis of rainfall images. These are: 100–600 m (small scales) and 600 m–6 km (large scales).

the study area was rather small (i.e. 6 km × 6 km), the domain was analysed as a whole (i.e. it was not divided into sub-domains, as is often done when large areas are analysed). A single movement vector representing the main velocity (both magnitude and direction) was thus obtained at each time step. The series of vectors obtained for the multiple time steps of the peak storm period were then averaged in order to obtain the mean velocity during this period (estimated velocity magnitudes are indicated in Table 5 of the results section).

### 3.2.3. Theoretically-required temporal resolution of rainfall inputs

The coarsest temporal resolution $(\Delta t_r)$ that is required to reflect the spatial structure of a storm as captured by data can be defined as the time needed to 'pass' the mean extent of the spatial structure (defined above). Based upon this definition, $\Delta t_r$ can be computed as:

$$\Delta t_r = r_c/|\overline{v}| \tag{9}$$

where $|\overline{v}|$ is the magnitude of the mean velocity of the storm over the peak period.

### 3.2.4. Maximum observable singularity $(\gamma_s)$

While the geostatistical approach used to compute $\Delta s_r$ provides a tangible estimate of the spatial features of a storm, it has the limitation of being a second-order approximation which means that it cannot properly reflect non-linear features (Schertzer and Lovejoy, 1987; Wang et al., 2015). With the purpose of further quantifying the spatial variability of rainfall fields, including higher-order statistical features, the concept of maximum observable singularity was used (Hubert et al., 1993; Douglas and Barros, 2003; Royer et al., 2008). This concept relies on the Universal Multifractal (UM) framework (see Schertzer and Lovejoy (2011) for a recent review) and quantifies the extremes one can expect to observe on a given sample of data according to its intrinsic variability. $\gamma_s$ is estimated not at a single resolution, but across a range of resolutions over which scale invariance or scaling behaviour is detected (i.e. fluctuations at small scales are related to larger ones by the same scaling law). More precisely, a multifractal analysis is first conducted on the rainfall images for a given storm event, based upon which UM parameters are retrieved and scaling across different resolutions, as well as breaks in scaling, are identified. $\gamma_s$ is subsequently computed from the UM parameters across the resolutions for which scale invariance is detected. By comparing $\gamma_s$ over different scaling regimes, it is possible to detect changes in the spatial variability of rainfall fields as a result of resolution coarsening.

## 3.3. Application of rainfall inputs to models

Rainfall estimates at the selected temporal and spatial resolutions were applied as input to the hydraulic models of the seven urban catchments in such a way that the resulting modelling outputs were as comparable as possible. Firstly, rainfall estimates were applied such that the centroid of the clipped rainfall area (see Section 2.3) coincides with the centroid of each catchment (see Section 2.1). Moreover, rainfall inputs were applied in two relative directions: parallel and perpendicular to the main flow direction at each catchment. As explained in Section 3.2, storm direction was estimated based on the TREC method. The predominant flow direction at each pilot catchment was estimated based upon the slope of the linear regression of the $(x, y)$ coordinates of the nodes located along the longest pipe flow path of each of the catchment models (Fig. 1). By applying rainfall inputs in the same relative direction to each catchment, variations in response due to differences in relative storm/flow direction (Singh, 1997) are avoided, thus making the results more comparable. Moreover, by applying rainfall inputs in these two relative directions it is possible to study variations in response due to differences in relative storm/flow direction.

## 3.4. Retrieval of hydraulic modelling results

For each of the hydraulic simulations carried out for each catchment (i.e. 9 storm events $\times$ 16 resolution combinations $\times$ 2 storm directions = 288), simulated flow time series at the downstream end of 8 pipes were retrieved for analysis. The 8 pipe locations were chosen such that the area that they drain (DA = drainage area) was approximately the following:

- 2 locations with DA $\sim$ 1 ha (i.e. characteristic length ($L = \sqrt{DA}) \sim 100$ m)
- 2 locations with DA $\sim$ 25 ha (i.e. $L \sim 500$ m)
- 1 locations with DA $\sim$ 100 ha (i.e. $L \sim 1000$ m)
- 1 locations with DA $\sim$ 300 ha (i.e. $L \sim 1700$ m)
- 1 locations with DA $\sim$ 500 ha (i.e. $L \sim 2200$ m)
- 1 locations with DA $\sim$ 600 ha (i.e. $L \sim 2500$ m)

These points for analysis were selected so as to assess the impact of rainfall input resolution in relation to the DA, which in previous studies has shown to play a dominant role in the requirements/impacts of rainfall input resolutions (e.g. Berne et al., 2004; Gires et al., 2012).

For the smallest catchments (e.g. Sucy-en-Brie (FR)), locations with the largest DAs do not exist. In these cases, simulation results for fewer points were retrieved. Conversely, in the case of catchments with total area >600 ha, results at an additional point corresponding to the downstream end of the catchment were retrieved. It is important to mention that the looped nature of the Kralingen catchment and the fact that flows may change direction throughout a storm event make it difficult to determine and estimate the area drained by a given pipe. For this catchment drainage areas were determined following the approach proposed by Bruni et al. (2015).

## 3.5. Evaluation of hydraulic modelling results

Using the hydraulic simulation results associated to the finest resolution rainfall estimates (i.e. 100 m/1 min) as reference, the following statistics were computed to quantify the impact of rainfall input resolution on the outputs of the hydraulic models of the seven urban. In order to allow inter-comparison of results from different catchments, storm events and points of analysis, only dimensionless statistics, which characterise different aspects of the simulated hydrographs, were used in this study.

- **Relative error (RE) in peak flow:**

$$RE_{st} = (Qmax_{st} - Qmax_{ref})/Qmax_{ref} \qquad (10)$$

where $RE_{st}$ is the relative error in the flow peak ($Qmax_{st}$) corresponding to a rainfall input of spatial resolution $s$ and temporal resolution $t$, in relation to the reference (100 m/1 min) flow peak, $Qmax_{ref}$. Positive $RE$ values indicate overestimation by the peak flow associated to the rainfall input $st$ (i.e. $Qmax_{st}$), and vice versa. The RE has the advantage of being a 'tangible' statistic which evaluates the performance of a critical parameter as is the peak flow. It is important to note that very large RE values can be obtained when low flows are evaluated, even if the absolute difference in peak flows is small. Hence RE values must be analysed with caution.

- **Coefficient of determination ($R^2$)** and **regression coefficient ($\beta$)** resulting from a simple linear regression analysis applied between each simulated flows time series ($Qst$, resulting from a rainfall input of spatial resolution $s$ and temporal resolution $t$) and the reference flow time series ($Qref$, resulting from the 100 m/1 min rainfall input). These two statistics provide an indication of how well the reference flows $Qref$ are replicated by the 'simulated' $Qst$ flows, both in terms of pattern and accuracy. The $R^2$ measure ranges from 0 to 1 and describes how much of the 'observed' variability in the $Qref$ time series is explained by the 'simulated' one (i.e. $Qst$). In practical terms, $R^2$ provides a measurement of the similarity between the patterns of the reference flow time series ($Qref$) and the 'simulated' ($Qst$) flow time series. However, biases in modelled estimates cannot be detected from this measure (Murphy, 1988; Krause et al., 2005; Gupta et al., 2009). The regression coefficient, $\beta$, is therefore employed to provide this supplementary information to the $R^2$. $\beta \approx 1$ represents good agreement in the magnitude of $Qref$ and $Qst$ time series; $\beta > 1$ means that the simulated flows ($Qst$) are higher in the mean (by a factor of $\beta$) than the reference flows ($Qref$); and $\beta < 1$ means the opposite (i.e. $Qst$ are lower in the mean than $Qref$). The $R^2$ and $\beta$ statistics have the advantage of taking into account the entire time series (as opposed to RE, which only provides an assessment of Qmax), as well as of being relatively insensitive to the magnitudes of the flows under consideration.

## 4. Results and discussion

### 4.1. Spatial/temporal characteristics of storm events

The estimated spatial and temporal characteristics of the storm events, as defined in Section 3.2, are summarised in Table 5. As can be seen, the mean velocity of the nine storms analysed in this study varies from 9.8 m/s to 18.4 m/s. The combination of storm velocity and catchment dimensions (namely length and width) provides an indication of the time that it takes for a given storm cell to cross a catchment. Given that the length and width of the pilot catchments range between $\sim$0.6 km and 8.2 km (see Table 1) and considering the minimum and maximum storm velocities, the time that it takes for the storms under consideration to cross the pilot catchments varies between $\sim$0.6 min and 13.9 min.

With regards to the minimum required resolutions, it can be seen that the required temporal scales for all storm events are rather small and generally below the 5 min temporal resolution of rainfall estimates provided by most meteorological services based on national weather radar networks. Considering the fine requirements in terms of temporal resolution, significant changes in hydraulic performance would be expected when switching from

the finest temporal resolution of 1 min to coarser resolutions of 3, 5 and 10 min, which quickly exceed the minimum required temporal resolution for most storm events. In contrast, the required spatial resolutions are less stringent. In fact, the typical spatial resolution of rainfall estimates provided by national networks (i.e. 1000 m) matches the required spatial resolution for 6 out of the 9 storms under consideration. Given that all of the theoretically-required spatial resolutions are coarser than 500 m and most of them are coarser than 1000 m, little impact is to be expected in the hydraulic outputs associated to rainfall input resolutions of 500 m and 1000 m, as compared to those associated to the finest 100 m estimates. However, a drop in performance would be expected for hydraulic outputs corresponding to rainfall input resolutions of 3000 m, as this spatial resolution largely exceeds the theoretically-required resolution of all storm events.

Storm events 5, 7 and 9 have the 'finest' requirements, both in terms of temporal and spatial resolutions. Therefore, the impact of resolution coarsening for these three events is expected to be larger than for other events.

The scaling analysis prior to the computation of the maximum observable singularity ($\gamma_s$) suggests that the studied storms generally exhibit a scaling behaviour on two ranges of scales: 100–600 m (small scales) and 600 m–6 km (large scales). The actual location of the scaling break varies from approximately 400 m to 800 m, depending on the event. With the purpose of allowing inter-comparison of $\gamma_s$ values, these are reported for the same ranges of scales for all storm events (i.e. 100–600 m and 600 m–6 km; see Table 5), Fig. 4 shows plots of the theoretically required spatial resolution ($\Delta s_r$) as a function of $\gamma_s$. Before proceeding to the analysis of these parameters, it is important to note that the limited range of scales available for the scaling analysis and computation of $\gamma_s$ (due to the small domain of the X-band radar) means that results are not very robust and should be interpreted as trends. Nevertheless, they provide useful and complementary insights into the intrinsic variability of the rainfall fields under consideration. The first interesting finding of this analysis is the identification of two different scaling regimes, which highlights the importance of measuring rainfall at high resolution (i.e. below the identified scaling break) in order to properly capture extremes, which cannot be extrapolated from coarser scale measurements. Secondly, from Fig. 4 it can be seen that, for both scaling regimes, the theoretically required spatial resolution ($\Delta s_r$) decreases with increasing $\gamma_s$. This means that data at higher spatial resolution are required to well characterise storms which display higher intrinsic variability. This is logical and indicates that the outputs of the two analysis
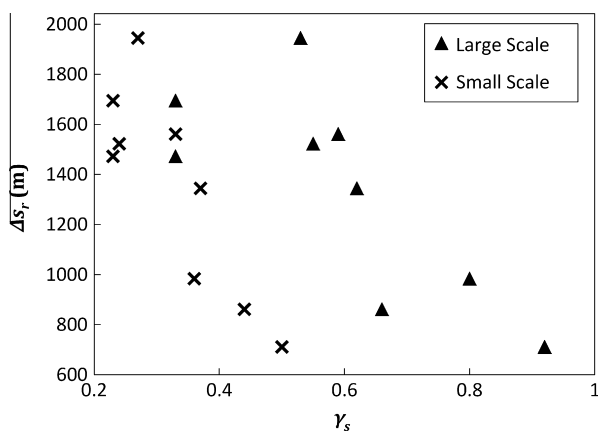
approaches used in the present study (i.e. geostatistical and multifractal) provide consistent results with regards to observed rainfall variability and extremes. However, it is worth noting that the required spatial resolutions ($\Delta s_r$) estimated with the geostatistical approach are mostly within the larger scale regime identified from the fractal analysis. This suggests that the geostatistical approach may be insufficient to characterise small scale, non-linear spatial features present in rainfall fields. This highlights the complementarity between the information provided by the two approaches, though more work is needed to better understand their relationship and optimise the way in which this information is used.

The way in which these spatial–temporal characteristics of rainfall relate to the impact of rainfall input resolution on hydrodynamic modelling results is investigated in the next section.

## 4.2. Hydrodynamic modelling results

Hydrodynamic modelling outputs are analysed based upon the dimensionless statistics introduced in Section 3.5: relative error in flow peaks (RE), coefficient of determination ($R^2$) and regression coefficient ($\beta$). In this section general trends observed in the hydraulic outputs are first identified. Afterwards, a detailed analysis is conducted to better understand the relationship between storm characteristics, catchment drainage area and the impact of rainfall input resolution on hydrodynamic modelling results.

### 4.2.1. General trends observed in hydrodynamic modelling results

In Fig. 5 performance statistics for all rainfall inputs are plotted as a function of drainage area (DA) size, for storms applied parallelly and perpendicularly to the catchments' main flow direction. At a glance and as was expected, a general trend can be identified of the impact of rainfall input resolution to decrease as drainage area increases. Moreover, the coarsening of temporal resolution generally appears to have a stronger influence as compared to the coarsening of spatial resolution; this is especially the case for small drainage areas. The stronger impact of temporal resolution over spatial resolution is in agreement with the estimated required temporal and spatial resolutions discussed in Section 4.1, as well as with previous studies (Krajewski et al., 1991; Meselhe et al., 2009; Notaro et al., 2013). The strong impact of temporal resolution coarsening can be partly explained by the way in which coarser temporal resolutions were obtained (i.e. by sampling radar images at the desired time resolution, in order to replicate radar scanning strategies); this is further discussed in Section 5.

In terms of magnitudes, as captured by RE and $\beta$ statistics, a general underestimation tendency is observed as space and time resolutions of rainfall inputs become coarser (notice general trend of RE < 0 and $\beta$ < 1). Noteworthy is the fact that coarser spatial resolutions systematically lead to underestimation of flows (notice behaviour of 3000 m resolutions denoted by red to yellow triangular markers), while coarser temporal resolutions have a more random effect and occasionally lead to large overestimation of flows. The underestimation associated with coarser spatial resolutions can be partly due to the smoothing of peak rainfall intensities which occurs when rainfall is averaged in space. In addition, it can also be explained by the fact that the cores of the storms were centred on the catchments; thus, as the spatial resolution of rainfall inputs approaches catchment size, storm water may be transferred outside of the catchment boundaries (Ogden and Julien, 1994; Bruni et al., 2015). The random effect of the coarsening of temporal resolution on flow magnitudes can in part be explained by the way in which the varying temporal resolutions were obtained (i.e. by sampling). It is interesting to note that, as DA increases, the random effect of temporal resolution on flow magnitudes decreases and a systematic underestimation tendency becomes clearer. In terms of $R^2$, it can be seen that the coarsening of
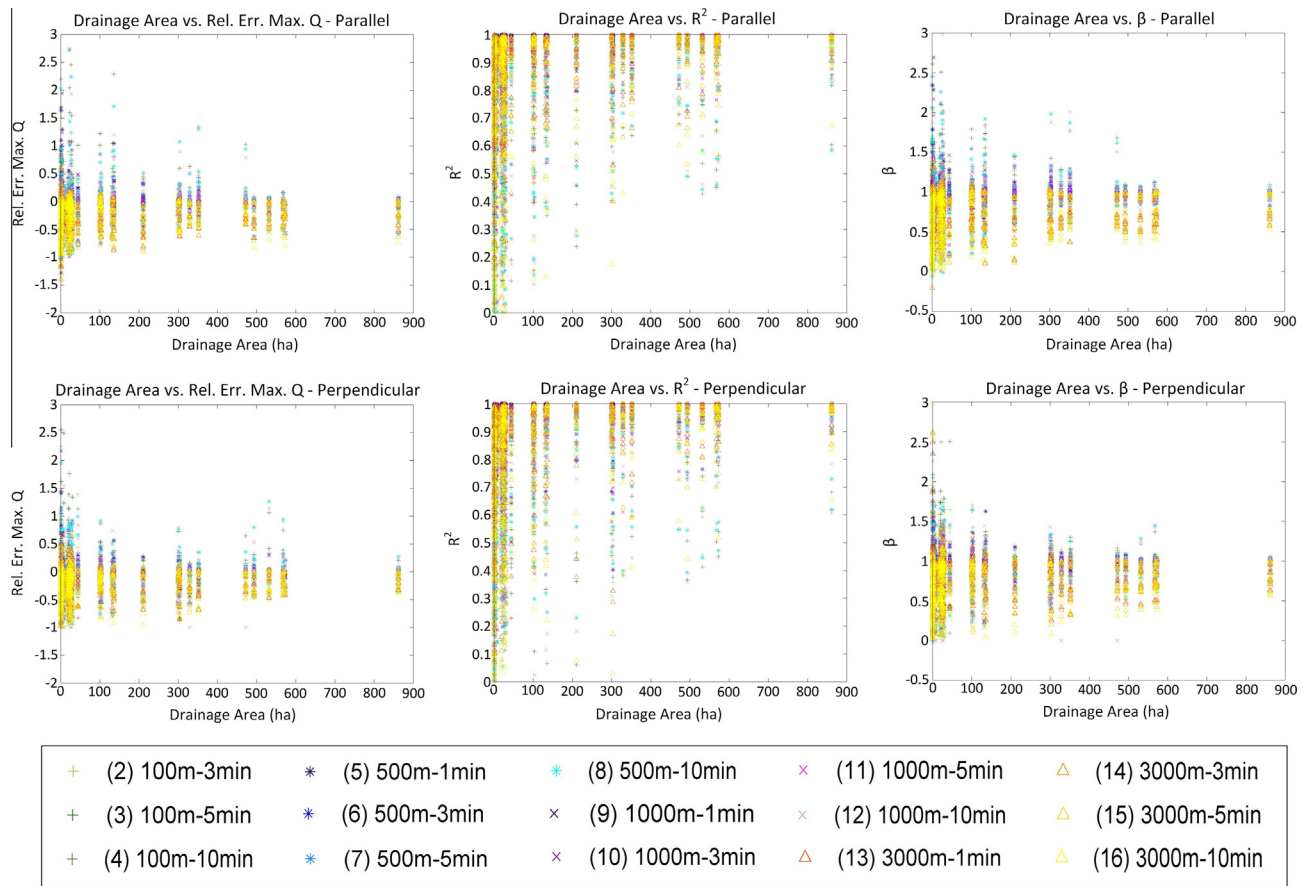


**Fig. 4.** Theoretically required spatial resolution ($\Delta S_r$) as a function of maximum observable singularity ($\gamma_s$), for small (100–600 m) and large (600 m–6 km) scale ranges.

**Fig. 5.** Scatterplots of performance statistics relative error in maximum flow peak, $R^2$ and $\beta$ versus drainage area sizes for 15 resolution combinations relative to the reference resolution of 1 min/100 m.

temporal resolution can easily alter the pattern of flow hydrographs: the lowest $R^2$ values are associated to the coarsest temporal resolutions, even when the associated spatial resolution is relative fine. Large drops in $R^2$ are also observed at spatial resolutions of 3000 m, which are significantly larger than the theoretically required spatial resolutions estimated for the storm events under consideration.

Regarding storm direction, similar trends are observed when storms are applied parallel and perpendicular to the predominant flow direction in the catchments (top and bottom plots in Fig. 5, respectively). Differences in response behaviour in relation to rainfall input resolutions for different storm direction would be expected particularly for elongated catchments. Such differences can be seen in some cases at the level of individual storms and catchments (plots not shown here), but these are rather small and do not have a significant impact on the general trends observed in summary statistics over all events and catchments. Given that a similar behaviour is observed for both relative storm directions, from now onwards only results for the parallel storm direction will be displayed and discussed. A detailed investigation of the impact of storm direction and individual catchment behaviour remains a topic for future study.

It is important to mention that some of the points of analysis at the different pilot catchments are subject to strong hydraulic controls (see Table 1). These controls influence flow behaviour and may lead to different sensitivity to rainfall input resolutions. To investigate this effect, the summary statistics shown in Fig. 5 were plotted separately for points with and without control elements. The resulting plots showed similar trends, indicating that control elements do not induce significantly different sensitivity to rainfall

input resolution for the investigated storms, catchments and drainage area sizes.

Fig. 6 shows boxplots of the performance statistics by spatial–temporal resolution, per group of drainage area (DA) sizes. These boxplots allow direct comparison of the performance of different rainfall inputs. Moreover, the separation by DA sizes allows for a partial removal from the analysis of the impact of catchment parameters on hydraulic outputs. The following groups of DA sizes were defined, corresponding with the spatial resolutions investigated in this study:

- DA1: 0.7–1.3 ha (i.e. characteristic length ($L = \sqrt{DA}$) $\sim$ 100 m)
- DA2: 20–30 ha ($L \sim 500$ m)
- DA3: 85–135 ha ($L \sim 1000$ m)
- DA4: 300–800 ha (1000 m < $L$ < 3000 m)

From these boxplots it can clearly be seen that the temporal resolution of rainfall input has a bigger impact on simulated flows than spatial resolution, thus confirming the initial findings derived from Fig. 5 and from the analysis of spatial–temporal characteristics of storms (Section 4.1). The results show that coarse temporal resolutions of 5–10 min can lead to large errors, even if spatial resolution is high. This also affects hydrograph shape, as reflected by low $R^2$ values. In agreement with Fig. 5, it can be seen that sensitivity to rainfall input resolution decreases with drainage area size: drainage areas of spatial scales of 100–500 m show high sensitivity to temporal resolution coarsening and comparatively moderate sensitivity to spatial resolution coarsening. Drainage areas of spatial scale above 1000 m display lower sensitivity to space and time resolution. Large errors due to spatial resolution coarsening occur
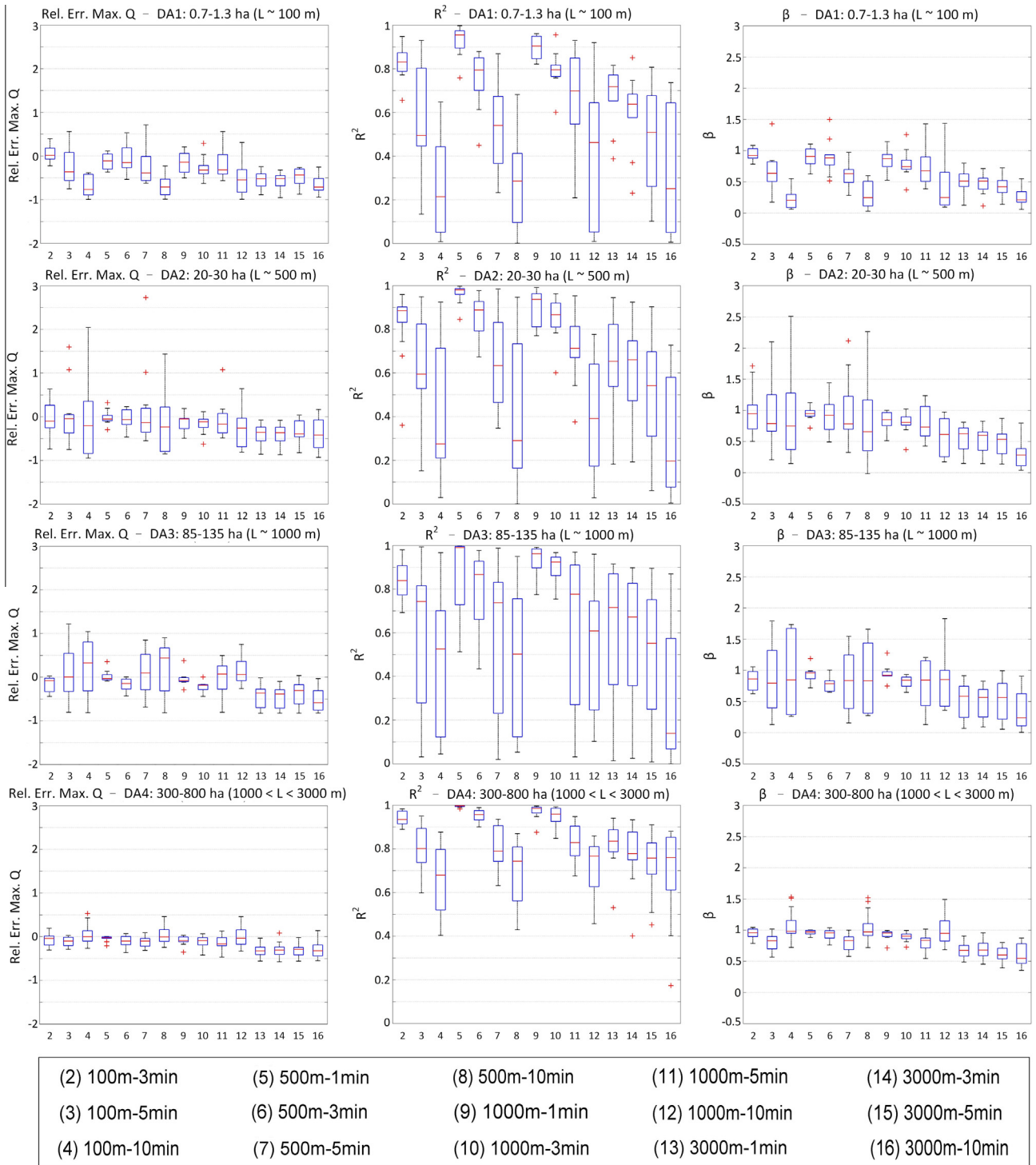
**Fig. 6.** Box plots of performance statistics relative error in maximum flow peak, $R^2$ and $\beta$ per rainfall input resolution, per group of drainage area sizes. Note that the boxplots' whiskers extend 1.5 times the interquartile range below the first quartile (Q1) and above the third quartile (Q3), respectively. Points beyond this distance are represented as outliers.

at 3000 m resolution, for all drainage area sizes. The trends observed in Fig. 6 corroborate previous findings from Fig. 5 and provide confirmation that the theoretically-derived required spatial and temporal resolutions are sound.

An interesting feature that can be observed in Fig. 6 is the interaction and mutual dependence between temporal and spatial resolutions. Notice, for instance, that the 1000 m/5 min (one of

the resolution combinations derived from Kolmogorov – see Fig. 3) associated outputs generally display a better performance than the 100 m/5 min ones, thus confirming the need for agreement between spatial and temporal resolution. The dependence between spatial and temporal resolutions has been widely discussed (e.g. Kolmogorov, 1962; Schertzer and Lovejoy, 1987; Marsan et al., 1996; Deidda, 2000; Gires et al., 2012), but there is

not as yet much evidence in urban hydrology to corroborate this hypothesis. The results of this study do provide evidence to support it.

The findings from Figs. 5 and 6 are generally in agreement with the findings and recommendations of Berne et al. (2004), but some differences are found. Berne et al. (2004) derived a relationship between space and time resolution of rainfall input required for urban hydrological analysis, based on catchment sizes of 10–10,000 ha in the Mediterranean region. The relationship they derived corresponds to a minimum rainfall resolution of 1 min/1.5 km for catchments smaller than 10 ha; 6 min/3.7 km for catchments of about 1000 ha. The temporal resolution they suggest for small drainage areas is in agreement with the findings of the present study; however, in relation to the spatial resolution, the present study suggests that for small drainage areas significant differences in flow estimates can be caused by changes in spatial resolution between 100 m, 500 m and 1000 m, at 1 min time resolution. In addition, the present study suggests that even for larger basins, relevant information is lost at time resolutions below 5 min.

### 4.2.2. Analysis of rainfall input resolution versus resolution requirements based on characteristic space–time scale of storm events

To investigate the impact of the spatial–temporal characteristics of storms on the observed variability in runoff estimates resulting from different rainfall input resolutions, performance statistics were plotted as a function of the following spatial and temporal scaling factors, as well as a function of a combined spatial–temporal factor which accounts for spatial–temporal scaling anisotropy (described in Section 3.1):

$$\theta_s = \left(\frac{\Delta S_r}{\Delta S}\right) \tag{11}$$

$$\theta_t = \left(\frac{\Delta t_r}{\Delta t}\right) \tag{12}$$

$$\theta_{st} = \left(\frac{\Delta S_r}{\Delta S}\right)\left(\frac{\Delta t_r}{\Delta t}\right)^{\frac{1}{1-H_t}} \tag{13}$$

where $\theta$ is a spatial–temporal scaling factor, $\Delta S_r$ and $\Delta t_r$ are the required spatial and temporal resolutions estimated based upon storm characteristics (see Table 5), $\Delta S$ and $\Delta t$ are the space and time resolutions of the rainfall inputs applied in model simulation and $H_t$ is the scaling anisotropy factor, defined in Section 3.1, which theoretically has a value of 1/3.

Figs. 7 and 8 show performance statistics $R^2$ and $\beta$ as a function of the scaling factor $\theta$ for scaling in space, scaling in time and combined spatial–temporal scaling, accounting for anisotropy. Relative errors (RE) plots were not included due to space constraints and given that these display a very similar behaviour to that of the $\beta$ plots. Same as in Fig. 6, in Figs. 7 and 8 plots are displayed per group of drainage area (DA) sizes, in order to partially remove from the analysis the impact of catchment parameters on hydraulic outputs. In Figs. 7 and 8, for $\theta$ values above 1, the applied rainfall input resolution is finer than the theoretically required spatial–temporal resolution, estimated based upon storm characteristics (see Table 5). In the case of the spatial scaling factor ($\theta_s$) alone (first row in Figs. 7 and 8), significant dispersion is observed in the plots and although performance statistics generally improve as $\theta_s$ increases, the improvement is not significant and the trend is rather unclear. In contrast, in the case of the temporal scaling factor ($\theta_t$) (middle row in Figs. 7 and 8) a more clear pattern can be observed in the plots, with performance statistics visibly improving at larger values of $\theta_t$. In the case of the combined factor ($\theta_{st}$) (bottom row in Figs. 7 and 8) a significantly clearer pattern can be identified, with performance consistently improving for higher $\theta_{st}$ values, whereby small drainage areas remain more sensitive. While some dispersion can still be seen in the plots of combined factor ($\theta_{st}$) vs. performance statistics, the fact that a significantly clearer pattern is observed in the $\theta_{st}$ plots, in comparison to the plots of the independent factors $\theta_s$ and $\theta_t$, suggests that in order to properly represent the effect of temporal and spatial resolution
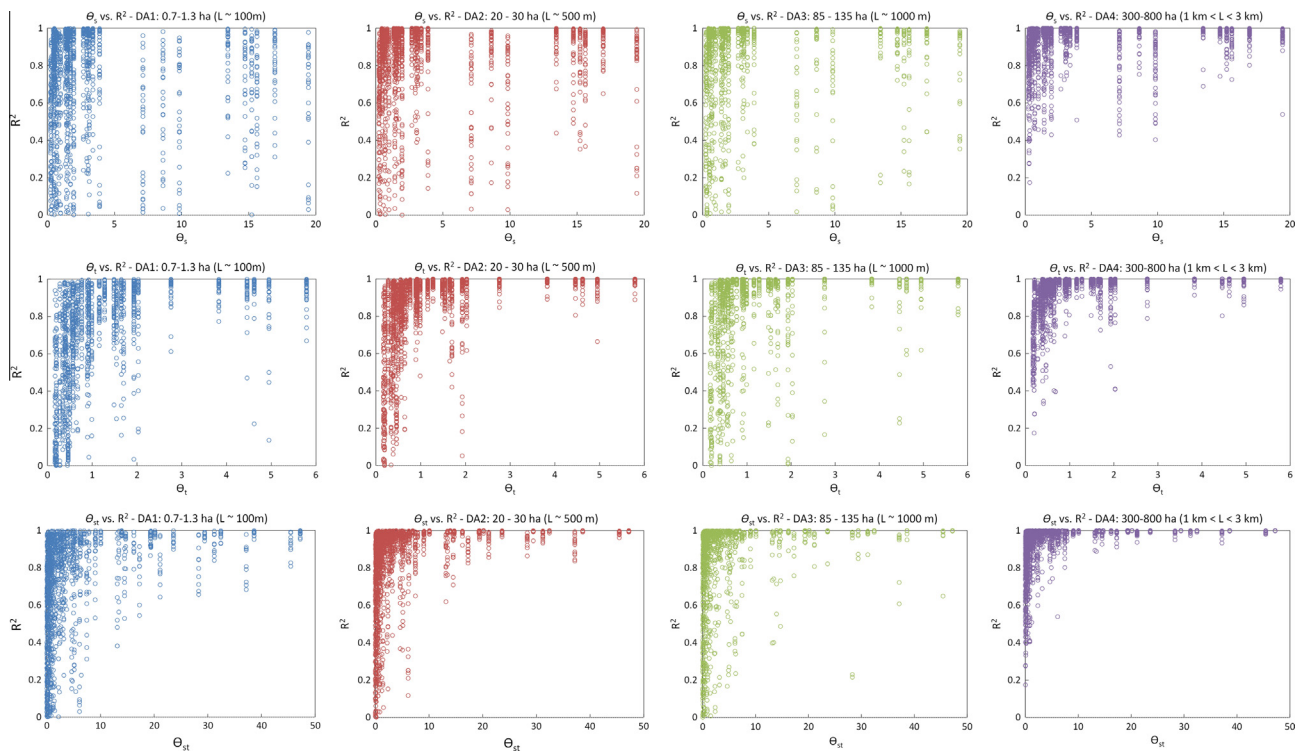


**Fig. 7.** Scatterplots of performance statistic $R^2$ as a function of scaling factors $\theta_s$ (top row), $\theta_t$ (middle row) and $\theta_{st}$ (bottom row), for 4 groups of drainage area sizes.
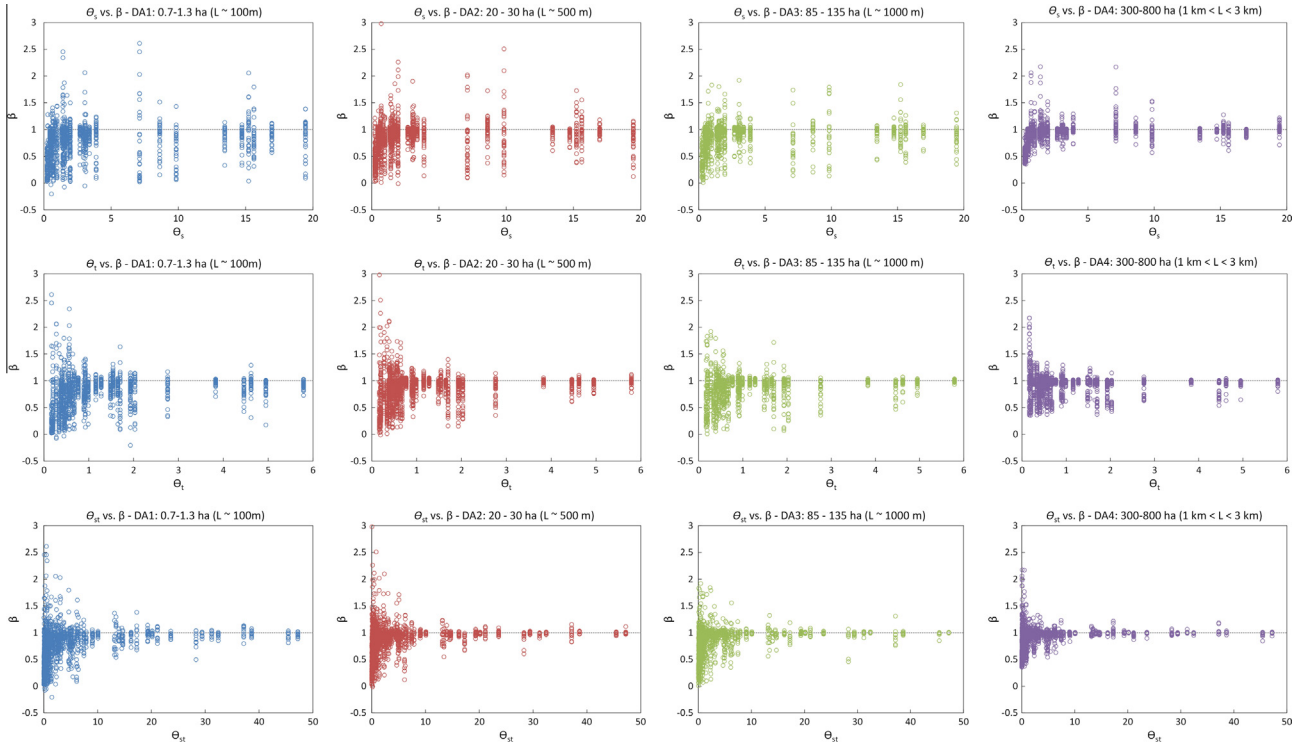
**Fig. 8.** Scatterplots of performance statistic $\beta$ as a function of scaling factors $\theta_s$ (top row), $\theta_t$ (middle row) and $\theta_{st}$ (bottom row), for 4 groups of drainage area sizes.
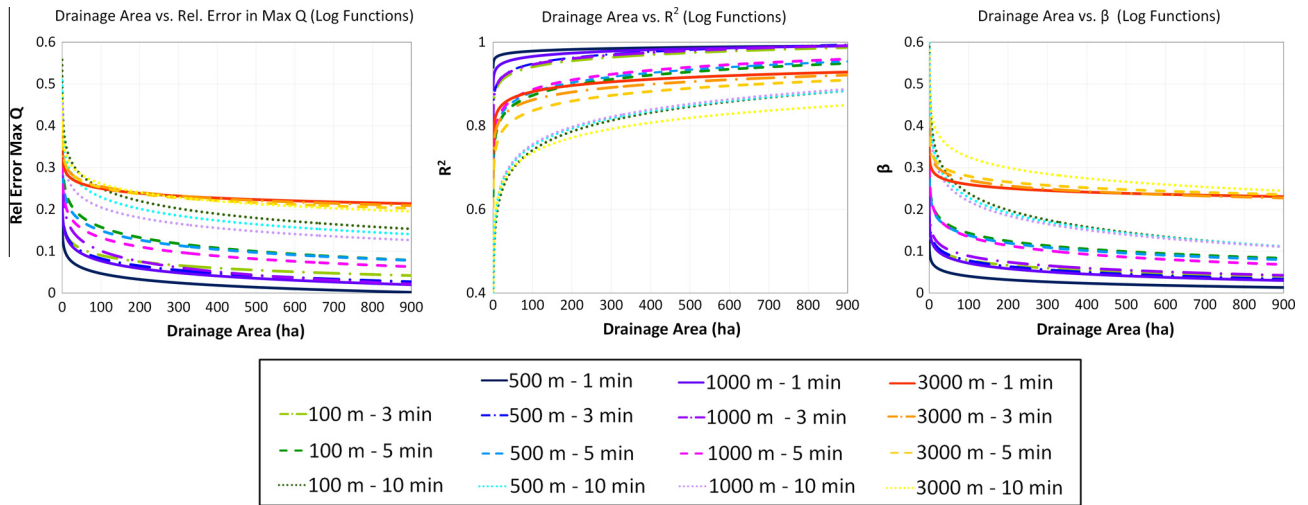


**Fig. 9.** Logarithmic functions fitted to data of performance statistics relative error in maximum flow peak, $R^2$ and $\beta$ as a function of drainage area size, for different space–time resolution combinations. Line type denotes different temporal resolutions (1 min = solid; 3 min = dash-dot; 5 min = dashed; 10 min = dotted) and colour range denotes different spatial resolutions (100 m = green; 500 m = blue; 1000 m = purple; 3000 m = orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

of rainfall inputs, these must be considered together. This corroborates the interaction that exists between the two resolutions. Future work will focus on further investigating these interactions, along with other catchment and model factors which influence the results and may be responsible for the remaining dispersion observed in the $\theta_{st}$ plots.

### 4.2.3. Analysis of hydrodynamic response statistics in relation to rainfall input resolution and drainage area size

In Fig. 9, performance statistics were plotted as a function of drainage area size, for different spatial–temporal resolution combinations. A logarithmic function was fitted to the resulting plots using the least squares method. The function structure was defined as

$$Performance\ Stat = a \cdot \ln DA + b \qquad (14)$$

The obtained $a$ and $b$ parameters and the associated mean square errors (MSE) of the fitting are summarised in Table 6.

The logarithmic functions provide a rough estimate of what hydrodynamic modelling performance can be expected for a given rainfall input resolution and catchment drainage area. For instance, for drainage area size of 100 ha, relative errors in maximum flow peak are expected to be below 0.1 for resolution combinations of 1 min/100–1000 m, while errors above 0.2 are expected for combinations of 10 min/100–1000 m and 1–10 min/3000 m resolution.

Based on the logarithmic functions plotted in Fig. 9, operational resolution of 5 min/1000 m provided by many national weather

**Table 6**
Parameters $a$ and $b$ and MSE-values for logarithmic function fitting, for performance statistics relative error (RE) in maximum flow peak, $\beta$ and $R^2$.

| Res ID | $a$ Abs RE | $b$ Abs RE | MSE Abs RE | $a$ $R^2$ | $b$ $R^2$ | MSE $R^2$ | $a$ Abs $\beta$ | $b$ Abs $\beta$ | MSE Abs $\beta$ |
|---|---|---|---|---|---|---|---|---|---|
| 2 | −0.0221 | 0.1920 | 0.0255 | 0.0211 | 0.8439 | 0.0143 | −0.0172 | 0.1577 | 0.0181 |
| 3 | −0.0361 | 0.3242 | 0.0694 | 0.0349 | 0.7120 | 0.0334 | −0.0274 | 0.2696 | 0.0361 |
| 4 | −0.0444 | 0.4557 | 0.2265 | 0.0648 | 0.4432 | 0.0605 | −0.0583 | 0.5073 | 0.0812 |
| 5 | −0.0205 | 0.1412 | 0.0182 | 0.0056 | 0.9528 | 0.0054 | −0.0123 | 0.0969 | 0.0090 |
| 6 | −0.0253 | 0.1988 | 0.0246 | 0.0208 | 0.8525 | 0.0093 | −0.0201 | 0.1702 | 0.0163 |
| 7 | −0.0325 | 0.2991 | 0.0802 | 0.0336 | 0.7254 | 0.0309 | −0.0266 | 0.2605 | 0.0329 |
| 8 | −0.0404 | 0.4158 | 0.1642 | 0.0604 | 0.4738 | 0.0605 | −0.0534 | 0.4745 | 0.0675 |
| 9 | −0.0259 | 0.1959 | 0.0276 | 0.0115 | 0.9130 | 0.0082 | −0.0189 | 0.1585 | 0.0175 |
| 10 | −0.0361 | 0.2674 | 0.1594 | 0.0207 | 0.8510 | 0.0111 | −0.0211 | 0.1860 | 0.0172 |
| 11 | −0.0317 | 0.2786 | 0.0598 | 0.0333 | 0.7334 | 0.0294 | −0.0303 | 0.2744 | 0.0307 |
| 12 | −0.0358 | 0.3698 | 0.1023 | 0.0603 | 0.4779 | 0.0616 | −0.0498 | 0.4489 | 0.0610 |
| 13 | −0.0166 | 0.3265 | 0.0539 | 0.0214 | 0.7832 | 0.0310 | −0.0132 | 0.3207 | 0.0501 |
| 14 | −0.0209 | 0.3512 | 0.0532 | 0.0271 | 0.7372 | 0.0386 | −0.0194 | 0.3596 | 0.0530 |
| 15 | −0.0227 | 0.3574 | 0.0793 | 0.0334 | 0.6821 | 0.0451 | −0.0200 | 0.3719 | 0.0579 |
| 16 | −0.0304 | 0.4021 | 0.0939 | 0.0517 | 0.4979 | 0.0699 | −0.0368 | 0.4947 | 0.0676 |

radar networks is expected to result in relative errors in flow peak of about 0.2 for small drainage areas (1–10 ha) down to about 0.1 for drainage area sizes of up to 800 ha. $R^2$ and $\beta$ values are expected to vary between 0.8 and 0.95 and between 0.3 and 0.7 for drainage area ranging from 1 to 800 ha.

While these results provide an indication of expected performance for varying rainfall input resolutions, they should be interpreted with caution. As values in Table 6 show, MSE-values are generally low for temporal resolutions of 1–3 min, but tend to decrease for lower temporal resolution and for spatial resolution above 1000 m.

Besides providing a practical estimate of the performance that can be expected for a given rainfall input resolution, the fitted logarithmic functions provide useful insights into the impact and interaction of spatial and temporal resolutions. In the case of relative error in peak flow and most evidently in the case of $\beta$, it can be seen that the fitted curves are grouped into four main sets: three of them corresponding to a given temporal resolution (and varying spatial resolutions from 100 m to 1000 m), and a fourth group corresponding to all the curves of spatial resolution 3000 m and varying temporal resolutions. The first three sets of curves further confirm the predominant effect of temporal resolution, which determines the performance of a given rainfall input, regardless of its spatial resolution, so long as the latter is kept close to the estimated required resolution. The fourth set of curves, corresponding to spatial resolutions of 3000 m and varying temporal resolutions, confirms that the 3000 m resolution largely exceeds the required spatial resolution, thus causing a general drop in performance for all rainfall inputs at this spatial resolution, regardless of their temporal resolution. A similar behaviour is observed in the case of $R^2$, although for this statistic the 3000 m estimates curves are not grouped together, suggesting that in terms of the pattern of flow hydrographs, as measured by $R^2$, temporal resolution plays an ever more predominant role, which even overshadows the effect of the coarsest (3000 m) resolution.

## 5. Summary, conclusions and outlook

The aim of this paper was to quantify the impact of rainfall input resolutions on operational urban drainage modelling outputs and, based upon it, to identify critical resolutions which enable a proper characterisation of urban catchment hydrological response. Using X-band radar-rainfall estimates for nine storm events, initially at 100 m and 1 min resolution, 16 different combinations of spatial and temporal resolutions, up to 3000 m and 10 min, were generated. Coarser spatial resolutions were generated by averaging in space, whereas coarser temporal resolutions were generated by sampling radar images at the desired temporal resolution, thus replicating radar scanning strategies. The resulting rainfall estimates were applied as input to the operational semi-distributed hydrodynamic models of seven urban catchments in North-West Europe, all of which have similar size (between 3 and 8 km$^2$), but different morphological, hydrological and hydraulic characteristics.

The spatial–temporal characteristics of the storm events, including theoretically required spatial and temporal resolutions given the observed rainfall variability, were derived using geostatistical analysis and storm cell tracking. In addition, the concept of maximum observable singularity, which relies on the framework of Universal Multifractals and allows quantifying higher-order statistical features, was used to quantify the intrinsic variability of rainfall fields at different spatial scales. Hydrodynamic response behaviour was summarised using dimensionless performance statistics and was analysed in the light of drainage area and critical spatial–temporal resolutions computed for each of the storm events.

The main findings and implications of this study are the following:

- Results of the geostatistical analysis and storm cell tracking showed that very fine temporal resolutions, usually below 5 min, are required to properly capture the variability observed in the rainfall data. This requirement is seldom met by rainfall estimates available from national weather radar networks, usually at temporal resolutions of 5 or 10 min. In contrast, the theoretically required spatial resolutions (derived from the geostatistical analysis) appear to be less stringent, with required resolutions ranging between 700 m and 2 km, which are generally met by the radar products provided by national weather services (usually at 1 km resolution). Nonetheless, the multifractal analysis of rainfall fields revealed a break in scaling behaviour between 400 m and 800 m which suggests that rainfall should be measured at sub-kilometric scales, in order to capture structures and extremes which cannot be extrapolated from measurements at coarser resolutions.

- In agreement with previous studies (e.g. Berne et al., 2004; Gires et al., 2012; Lobligeois et al., 2014), the impact of rainfall input resolution on hydraulic outputs was shown to decrease significantly as catchment drainage area increases. For drainage areas of the order of 1 ha errors in peak discharges of up to 250% were observed as a result of rainfall input resolution coarsening, whereas for drainage areas of ~800 ha maximum errors in peak discharge were of the order of 50%.

- Across the entire range of drainage areas under investigation (1–800 ha), the coarsening of temporal resolution of rainfall inputs was shown to have a bigger effect upon hydrodynamic modelling results than the coarsening of spatial resolution. These results are in agreement with the independent

(geostatistical and cell tracking) analysis of the storms and corroborate the need for rainfall input temporal resolutions below 5 min for urban hydrological applications. The strong and dominant effect of temporal resolution coarsening can be partially explained by the way in which coarser temporal resolutions were derived, i.e., by sampling radar images at the desired time step, thus replicating radar scanning strategies. A study focusing on investigating the impact of rainfall temporal resolution coarsening through aggregation (i.e. averaging in time, which resembles the functioning of rain gauges) is currently underway. Initial results indicate that the impact of temporal resolution decreases significantly when coarser temporal resolution estimates are generated through aggregation as opposed to sampling. When aggregation is used, coarsening of spatial and temporal resolutions leads to impacts of comparable magnitudes on estimated runoff, although the latter still has a bigger effect. Similar findings regarding the dominant effect of temporal resolution over spatial resolution have been obtain from other studies, both in rural and urban catchments (Krajewski et al., 1991; Meselhe et al., 2009; Notaro et al., 2013).

- With regards to required rainfall input spatial resolution, this is strongly dependent upon the drainage area of interest. For very small drainage areas, below 1 ha, rainfall input resolutions of ~100 m are required. For drainage areas between 1 ha and ~100 ha, rainfall inputs at a spatial resolution of 500 m appear to be sufficient; for these areas no significant improvement is observed when using finer spatial resolution rainfall estimates and acceptable hydraulic performance is still obtained for rainfall estimates at 1 km/1 min resolution. For drainage areas larger than 100 ha rainfall input spatial resolutions of 1 km appear to be sufficient, leading to high values of performance statistics, as long as the accompanying temporal resolution is fine enough (<5 min). For all drainage areas, rainfall input spatial resolution of 3 km, which may be compared to common distances between rain gauges, appears to be insufficient, leading to very poor hydraulic performance statistics. It can be seen that, in general (except for very small drainage areas) and in agreement with the results of the storm analysis described above, ~1 km resolution rainfall estimates appear to be sufficient for urban hydrological modelling. However, it is important to mention that these results are bound to the storm events under consideration and to the type of models employed in this study; it is, operational semi-distributed, albeit high-resolution models (see subcatchment sizes in Table 2), calibrated using rain gauge records of coarse spatial resolution as input, which may lead to spatially homogeneous model parametrisation (Finnerty et al., 1997). Higher-resolution fully-distributed models, implemented and calibrated using high resolution datasets, are likely to be more sensitive to the spatial resolution of rainfall inputs and may therefore require sub-kilometric resolution rainfall estimates as input (Schertzer et al., 2010; Gires et al., 2014a,b; Pina and Ochoa-Rodriguez, 2014; Ichiba et al., 2015).
- Despite the dominant effect of temporal resolution, the hydraulic results show that there is a strong interaction and dependence between the spatial and temporal resolution of rainfall input estimates. As such, in order to avoid losing relevant information from the rainfall fields, the two resolutions must be in agreement with each other. For example, the hydraulic outputs associated with rainfall inputs at 1000 m/5 min resolution display a better performance than those associated with 100 m/5 min ones. The dependence between spatial and temporal resolutions has been widely discussed (e.g. Kolmogorov, 1962; Schertzer and Lovejoy, 1987; Marsan et al., 1996; Deidda, 2000; Gires et al., 2012), but there is not as yet much evidence in urban hydrology to corroborate this hypothesis. The results of this study do provide evidence to support it.

- The theoretically derived minimum spatial–temporal resolution of rainfall inputs, estimated on the basis of the sole analysis of rainfall images, are consistent with the results of the hydraulic analysis. This validates the proposed approach to characterise storm events and suggests that, in addition to drainage area, a big part of the impact of rainfall input resolution on urban runoff estimates can be explained by the spatial–temporal characteristics of the storm events. The influence of other factors such as catchment and model characteristics was not investigated in detail and remains a topic for future study.

While the present study has several limitations, the results provide useful insights into rainfall input resolution requirements for urban hydrological applications, considering currently available data and models. Evidently, higher spatial and temporal resolution rainfall estimates are desirable. However, resolution comes at a cost and resources are limited. According to the results of this study, rainfall monitoring strategies may consider prioritising improvements in temporal resolution (e.g. modifying radar scanning strategies, using local X-band radars which have higher rotation rates, employing temporal interpolation techniques), while keeping in mind the dependence between temporal and spatial resolutions, as well as the fact that measuring rainfall at higher resolutions can lead to improvements in accuracy. Future research should focus on gathering high resolution rainfall datasets alongside high resolution local urban runoff records and implementation of higher resolution urban drainage models, which enable a better assessment of the added value of high resolution rainfall data and models. Further work is also needed to better understand factors affecting model sensitivity to rainfall input resolution, including storm spatial–temporal characteristics, as well as catchment and model characteristics (e.g. slope, degree of imperviousness, presence of control elements, spatial homogeneity/heterogeneity, amongst others).

## Acknowledgements

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2015.05.035.

## References

Arnaud, P., Bouvier, C., Cisneros, L., Dominguez, R., 2002. Influence of rainfall spatial variability on flood prediction. J. Hydrol. 260 (1–4), 216–230.

Aronica, G., Cannarozzo, M., 2000. Studying the hydrological response of urban catchments using a semi-distributed linear non-linear model. J. Hydrol. 238 (1–2), 35–43.

Atkinson, P.M., Aplin, P., 2004. Spatial variation in land cover and choice of spatial resolution for remote sensing. Int. J. Remote Sens. 25 (18), 3687–3702.

Bastin, G., Lorent, B., Duqué, C., Gevers, M., 1984. Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. Water Resour. Res. 20 (4), 463–470.

Berndtsson, R., Niemczynowicz, J., 1988. Spatial and temporal scales in rainfall analysis — Some aspects and future perspectives. J. Hydrol. 100 (1–3), 293–313.

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

Biaou, A., Chauvin, F., Royer, J.F., Schertzer, D., 2005. Analyse multifractale des précipitations dans un scénario GIEC du CNRM. Note de centre GMGEC, CNRM, 101, 45 pp.

Bringi, V.N., Chandrasekar, V., 2001. Polarimetric Doppler Weather Radar Principles and Applications. Cambridge University Press, Cambridge, UK.

Bruni, G., Reinoso, R., van de Giesen, N.C., Clemens, F.H.L.R., ten Veldhuis, J.A.E., 2015. On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. Hydrol. Earth Syst. Sci. 19 (2), 691–709.

Chaubey, I., Haan, C.T., Grunwald, S., Salisbury, J.M., 1999. Uncertainty in the model parameters due to spatial variability of rainfall. J. Hydrol. 220 (1–2), 48–61.

Deidda, R., 2000. Rainfall downscaling in a space-time multifractal framework. Water Resour. Res. 36 (7), 1779–1794.

Douglas, E.M., Barros, A.P., 2003. Probable maximum precipitation estimation using multifractals: application in the Eastern United States. J. Hydrometeorol. 4 (6), 1012–1024.

Einfalt, T., 2005. A hydrologists' guide to radar use in various applications. In: 10th International Conference on Urban Drainage, Copenhagen, Denmark.

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

Emmanuel, I., Andrieu, H., Leblois, E., Flahaut, B., 2012. Temporal and spatial variability of rainfall at the urban hydrological scale. J. Hydrol. 430–431, 162–172.

Fabry, F., Bellon, A., Duncan, M.R., Austin, G.L., 1994. High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. J. Hydrol. 161, 415–428.

Faurès, J.-M., Goodrich, D.C., Woolhiser, D.A., Sorooshian, S., 1995. Impact of small-scale spatial rainfall variability on runoff modeling. J. Hydrol. 173 (1–4), 309–326.

Fewtrell, T.J., Duncan, A., Sampson, C.C., Neal, J.C., Bates, P.D., 2011. Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. Phys. Chem. Earth, Parts A/B/C 36 (7–8), 281–291.

Figueras i Ventura, J., 2009. Design of a High Resolution X-Band Doppler Polarimetric Radar. PhD Dissertation, Delft University of Technology, Delft, The Netherlands.

Finnerty, B.D., Smith, M.B., Seo, D.-J., Koren, V., Moglen, G.E., 1997. Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs. J. Hydrol. 203 (1–4), 21–38.

Garrigues, S., Allard, D., Baret, F., Weiss, M., 2006. Quantifying spatial heterogeneity at the landscape scale using variogram models. Remote Sens. Environ. 103 (1), 81–96.

Garrigues, S., Allard, D., Baret, F., Morisette, J., 2008. Multivariate quantification of landscape spatial heterogeneity using variogram models. Remote Sens. Environ. 112 (1), 216–230.

Giangola-Murzyn, A., Gires, A., Hoang, C.T., Tchiguirinskaia, I., Schertzer, D., 2012. Multi-Hydro modelling to assess flood resilience across scales, case study in the Paris region. In: 9th International Conference on Urban Drainage Modelling, Belgrade, Serbia.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Lovejoy, S., 2011. Multifractal and spatio-temporal analysis of the rainfall output of the Meso-NH model and radar data. Hydrol. Sci. J. 56 (3), 380–396.

Gires, A., Onof, C., Maksimović, Č., Schertzer, D., Tchiguirinskaia, I., Simoes, N., 2012. Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: a case study. J. Hydrol. 442, 117–128.

Gires, A., Giangola-Murzyn, A., Abbes, J.-B., Tchiguirinskaia, I., Schertzer, D., Lovejoy, S., 2014a. Impacts of small scale rainfall variability in urban areas: a case study with 1D and 1D/2D hydrological models in a multifractal framework. Urban Water J., 1–11.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., Berne, A., Lovejoy, S., 2014b. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. Atmos. Res. 138, 125–138.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91.

Horne, M.P., 2003. Short-Term Precipitation Nowcasting for Composite Radar Rainfall Fields. MSc Dissertation, Massachusetts Institute of Technology, Cambridge, USA.

Hubert, P., Tessier, Y., Lovejoy, S., Schertzer, D., Schmitt, F., Ladoy, P., Carbonnel, J.P., Violette, S., Desurosne, I., 1993. Multifractals and extreme rainfall events. Geophys. Res. Lett. 20 (10), 931–934.

Ichiba, A., Gires, A., Tchguirinskaia, I., Bompard, P., Schertzer, D., 2015. High resolution modeling in urban hydrology: comparison between two modeling approaches and their sensitivity to high rainfall variability. In: EGU General Assembly 2015, Vienna, Austria.

Kavetski, D., Kuczera, G., Franks, S.W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour. Res. 42 (3), W03407.

Kolmogorov, A.N., 1962. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. J. Fluid Mech. 13 (01), 82–85.

Krajewski, W.F., Lakshmi, V., Georgakakos, K.P., Jain, S.C., 1991. A Monte Carlo Study of rainfall sampling effect on a distributed catchment model. Water Resour. Res. 27 (1), 119–128.

Krause, P., Boyle, D.P., Base, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97.

Lantuéjoul, C., 1991. Ergodicity and integral range. J. Microsc. 161 (3), 387–403.

Lantuéjoul, C., 2002. Geostatistical Simulation – Models and Algorithms. Springer, Berlin, Germany.

Laroche, S., Zawadzki, I., 1995. Retrievals of horizontal winds from single-doppler clear-air data by methods of cross correlation and variational analysis. J. Atmos. Oceanic Technol. 12 (4), 721–738.

Leijnse, H., Uijlenhoet, R., van de Beek, C.Z., Overeem, A., Otto, T., Unal, C.M.H., Dufournet, Y., Russchenberg, H.W.J., Figueras i Ventura, J., Klein Baltink, H., Holleman, I., 2010. Precipitation measurement at CESAR, the Netherlands. J. Hydrometeorol. 11 (6), 1322–1329.

Li, P.W., Lai, E.S.T., 2004. Short-range quantitative precipitation forecasting in Hong Kong. J. Hydrol. 288 (1–2), 189–209.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., Loumagne, C., 2014. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. Hydrol. Earth Syst. Sci. 18 (2), 575–594.

Marsan, D., Schertzer, D., Lovejoy, S., 1996. Causal space-time multifractal processes: predictability and forecasting of rain fields. J. Geophys. Res.: Atmos. 101 (D21), 26333–26346.

Meselhe, E., Habib, E., Oche, O., Gautam, S., 2009. Sensitivity of conceptual and physically based hydrologic models to temporal and spatial rainfall sampling. J. Hydrol. Eng. 14 (7), 711–720.

Murphy, A.H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Weather Rev. 116 (12), 2417–2424.

Notaro, V., Fontanazza, C.M., Freni, G., Puleo, V., 2013. Impact of rainfall data resolution in time and space on the urban flooding evaluation. Water Sci. Technol. 68 (9), 1984–1993.

Obled, C., Wendling, J., Beven, K., 1994. The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. J. Hydrol. 159 (1–4), 305–333.

Ogden, F.L., Julien, P.Y., 1994. Runoff model sensitivity to radar rainfall resolution. J. Hydrol. 158 (1–2), 1–18.

Otto, T., Russchenberg, H.W.J., 2011. Estimation of specific differential phase and differential backscatter phase from polarimetric weather radar measurements of rain. Geosci. Remote Sens. Lett., IEEE 8 (5), 988–992.

Otto, T., Russchenberg, H.W.J., 2012. Rainfall rate retrieval with IDRA, the polarimetric X-band radar at Cabauw, NL. In: 7th European Conference on Radar Meteorology and Hydrology, Toulouse, France.

Pina, R., Ochoa-Rodriguez, S., 2014. Fully-distributed vs. Semi-distributed urban drainage models. In: RainGain International Workshop on Urban Pluvial Flood Models, Exeter, UK.

Pina, R., Ochoa-Rodríguez, S., Simoes, N., Mijic, A., Sa Marques, A., Maksimovic, Č., 2014. Semi-distributed or fully distributed rainfall-runoff models for urban pluvial flood modelling? In: 13th International Conference on Urban Drainage, Sarawak, Malaysia.

Rinehart, R.E., Garvey, E.T., 1978. Three-dimensional storm motion detection by conventional weather radar. Nature 273 (5660), 287–289.

Royer, J.-F., Biaou, A., Chauvin, F., Schertzer, D., Lovejoy, S., 2008. Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. C.R. Geosci. 340 (7), 431–440.

Scarchilli, G., Gorgucci, E., Chandrasekar, V., 1996. Self-consistency of polarization diversity measurements of rainfall. IEEE Trans. Geosci. Remote Sens. 34 (1), 22–26.

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.

Schertzer, D., Lovejoy, S., 1987. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. J. Geophys. Res. 92 (D8), 9693–9714.

Schertzer, D., Lovejoy, S., 2011. Multifractals, generalized scale invariance and complexity in geophysics. Int. J. Bifurcation Chaos 21 (12), 3417–3456.

Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., Hubert, P., 2010. No monsters, no miracles: in nonlinear sciences hydrology is not an outlier! Hydrol. Sci. J. 55 (6), 965–979.

Schilling, W., 1991. Rainfall data for urban hydrology: what do we need? Atmos. Res. 27, 5–21.

Segond, M.L., Neokleous, N., Makropoulos, C., Onof, C., Maksimović, Č., 2007. Simulation and spatio-temporal disaggregation of multi-site rainfall data for urban drainage applications. Hydrol. Sci. J. – J. Des Sci. Hydrol. 52 (5), 917–935.

Shah, S.M.S., O'Connell, P.E., Hosking, J.R.M., 1996. Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. J. Hydrol. 175 (1–4), 89–111.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (3), 379–423.

Singh, V.P., 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrol. Process. 11 (12), 1649–1669.

Smith, M.B., Koren, V.I., Zhang, Z., Reed, S.M., Pan, J.-J., Moreda, F., 2004. Runoff response to spatial variability in precipitation: an analysis of observed data. J. Hydrol. 298 (1–4), 267–286.

Syed, K.H., Goodrich, D.C., Myers, D.E., Sorooshian, S., 2003. Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. J. Hydrol. 271 (1–4), 1–21.

Tabios, G.Q., Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. JAWRA J. Am. Water Resour. Assoc. 21 (3), 365–380.

Tetzlaff, D., Uhlenbrook, S., 2005. Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data. Hydrol. Earth Syst. Sci. 9, 29–41.

Tuttle, J.D., Foote, G.B., 1990. Determination of the boundary layer airflow from a single doppler radar. J. Atmos. Oceanic Technol. 7 (2), 218–232.

Unal, C., 2009. Spectral polarimetric radar clutter suppression to enhance atmospheric echoes. J. Atmos. Oceanic Technol. 26 (9), 1781–1797.

Vieux, B.E., Imgarten, J.M., 2012. On the scale-dependent propagation of hydrologic uncertainty using high-resolution X-band radar rainfall estimates. Atmos. Res. 103, 96–105.

Wang, L.-P., Onof, C., Ochoa-Rodríguez, S., Simoes, N.E., Maksimović, Č., 2012. On the propagation of rainfall bias and spatial variability through urban pluvial flood modelling. In: 9th International Workshop on Precipitation in Urban Areas: Urban challenges in rainfall analysis, Saint Moritz, Switzerland.

Wang, L.-P., Ochoa-Rodríguez, S., Simoes, N., Onof, C., Maksimović, Č., 2013. Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. Water Sci. Technol. 68 (4), 737–747.

Wang, L.P., Ochoa-Rodríguez, S., Onof, C., Willems, P., 2015. Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications. Hydrol. Earth Syst. Sci. Discuss. 12 (2), 1855–1900.

WAPUG, 2002. Code of Practice For The Hydraulic Modelling of Sewer Systems, Wastewater Planning Users Group.
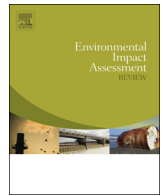
# Multivariate pluvial flood damage models ☆

Luc Van Ootegem [a,b,\*], Elsy Verhofstadt [b], Kristine Van Herck [a], Tom Creten [a]

[a] HIVA — University of Louvain, Belgium
[b] SHERPPA — Ghent University, Belgium

## ABSTRACT

Depth–damage-functions, relating the monetary flood damage to the depth of the inundation, are commonly used in the case of fluvial floods (floods caused by a river overflowing). We construct four multivariate damage models for pluvial floods (caused by extreme rainfall) by differentiating on the one hand between ground floor floods and basement floods and on the other hand between damage to residential buildings and damage to housing contents. We do not only take into account the effect of flood-depth on damage, but also incorporate the effects of non-hazard indicators (building characteristics, behavioural indicators and socio-economic variables). By using a Tobit-estimation technique on identified victims of pluvial floods in Flanders (Belgium), we take into account the effect of cases of reported zero damage. Our results show that the flood depth is an important predictor of damage, but with a diverging impact between ground floor floods and basement floods. Also non-hazard indicators are important. For example being aware of the risk just before the water enters the building reduces content damage considerably, underlining the importance of warning systems and policy in this case of pluvial floods.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

In Europe floods are currently among the natural catastrophes that cause the largest economic damage (European Environment Agency, 2012). In the future, the importance and occurrence of floods are expected to increase even further as predictions on the impact of climate change indicate that the winters in Western Europe will be wetter (Bruwier et al., 2015). As a result, flood risk management is likely to become more important in the coming years. An essential component of flood risk management and risk assessment will be to accurately predict the damage caused by the flood as this is an essential element in undertaking cost–benefit analyses to decide about e.g. government investments in specific infrastructure or for setting up flood warning systems. Most of the currently used damage predictions make use of depth–damage functions that relate the monetary damage to residential buildings and/or contents or other 'receptors at risk' to the depth of an inundation (Jonkman et al., 2008; Messner and Meyer, 2006). These functions are typically using data from floods caused by rivers overflowing (called fluvial floods). Their generalizability to other floods is unclear (Kellens et al., 2013).

In this paper, the focus is on the prediction of damage caused by a type of flood that is rarely studied, i.e. floods in urban areas that are caused by extreme rainfall events during which the water cannot be sufficiently processed by existing urban drainage systems (called pluvial floods). Fluvial floods are more devastating and spectacular, yet they do not occur that often. Pluvial floods come with less damage, but occur frequently and the cumulative damage over the years can be just as high as with fluvial events (ten Veldhuis, 2011). Currently, the lack of adequate pluvial damage models is a major bottleneck in estimating damage and therefore in calculating costs in pluvial risk assessments (Zhou et al., 2012). An exception is Blanc et al. (2012) who estimate pluvial flood damage using rainfall simulations to compute flood depth. The investigation of other than fluvial and coastal floods is for that reason also one of the main challenges for the implementation of the EU flood directive by EU member states (Kellens et al., 2013). In a study on the monetization of urban flood damage, ten Veldhuis (2011) shows that the depths and water velocities in the case of pluvial floods are usually rather low. It is quite probable that with low inundation depth and flow velocities, the damage during pluvial floods will depend on other predictors than fluvial flood damage does (Spekkers et al., 2012a). Research by Spekkers et al. (2011), (2012a), (2012b) and by Kreibich et al. (2009) also points in that direction.

We conducted a questionnaire among victims of pluvial floods. So the information we use, also about the flood, is self-reported. Our respondents reported on the replacement value of goods, while insurance companies pay only the depreciated value of the damaged goods. In addition, insurance companies may include deductibles in their contracts and not all victims are always able to deal effectively with the insurance companies (Michel-Kerjan and Kousky, 2010). As a result, compensations paid by insurance companies are on average below the damage reported by the households (e.g. Thieken et al., 2006 for Germany). This

**Table 1**
Socio-economic indicators of respondents.

| Socio-economic indicator | Percent |
|---|---|
| **Occupational situation** | |
| Full time working | 35.2% |
| Part time working | 8.0% |
| Retired | 46.2% |
| Unemployed with benefits | 1.5% |
| Disabled with benefits | 2.5% |
| Non-paid homeworking | 2.7% |
| Other | 3.8% |
| **Education** | |
| Basic education | 10.9% |
| Lower secondary education | 17.0% |
| Higher secondary education | 28.6% |
| Bachelor | 25.1% |
| Master or post-university | 18.3% |
| **Net income/month** | |
| Below 1000 euros | 6.13% |
| 1000–1499 euros | 24.52% |
| 1500–1999 euros | 19.52% |
| 2000–2500 euros | 12.58% |
| Above 2500 euros | 10.97% |
| No answer[a] | 26.29% |

[a] In the regression analysis later on we treat this group as a separate group so that these observations are not excluded from the analysis.

**Table 2**
Reported damage figures based on replacement values and based on payments by the insurance company and/or the Rampenfonds and zero damages (euro corrected for inflation).

| Damage to | Observations | Mean | Zero damage | Max damage |
|---|---|---|---|---|
| **Based on replacement value (€)** | | | | |
| Building | 363 | 6244.5 | 37 | 77,499.3 |
| Contents | 413 | 2979.9 | 184 | 65,576.3 |
| Total | 453 | 7720.6 | 87 | 91,657.6 |
| **Based on insurance company/Rampenfonds** | | | | |
| Total | 212 | 1997.7 | 0 | 19,307.8 |

can lead to a downward bias of more 'objective' damage figures. As such, self-reported damage figures by victims can be as appropriate (or as problematic) as insurance figures are. Therefore, using the more optimistic point of view, we explore and emphasize the opportunities that can be created by using survey data. This is possible because we have additional information (about zero damage) and also more detailed information (including non-hazard data and information concerning the specific places that are flooded).

The contribution of our research is threefold. Firstly, we are able to include 'no damage cases' of people who were flooded but did not suffer any damage. Those people would not appear in insurance or disaster fund records. As such, in that case, an unintentional truncation of the data takes place (i.e. some data are omitted albeit being relevant). Our questionnaire revealed that quite some flood victims suffer no damage to buildings and/or contents. Such an observation is important, especially for a pluvial flood event. It is possible that people are able to remove the water immediately before/during the flood or they are able to protect their belongings in some way (for instance by moving them to another place). We use a Tobit model estimation technique (Tobin, 1958) to deal with this issue of reported zero damage.

Secondly, we extend the traditional bivariate depth–damage relation by constructing more detailed multivariate flood damage models. For this model construction, thanks to the use of survey data, we can use various non-hazard indicators that are mostly unavailable otherwise. We include building characteristics (type of building, size of the dwelling), behavioural predictors related to the behaviour of victims before and during the flood (recurrence, risk awareness, emergency measures) and income as socio-economic indicator. Depth–damage functions usually relate the monetary damage only to the depth of an inundation (Jonkman et al., 2008; Messner and Meyer, 2006). In a related approach, several depth–damage functions are constructed for different categories of a specific hazard indicator other than depth or in some rare cases also for a non-hazard indicator. Such an approach is still quite common and is for instance used in the "British Multi-coloured Manual" by Penning-Rowsell et al. (2005) by constructing depth–damage curves for different types of buildings (terraced, semi-detached, detached, flats, bungalows and prefab buildings). Obviously, such a method becomes

**Table 3**
Description of the outcome and explanatory variables included in the different model specifications.

| | Description | Mean | Std. dev |
|---|---|---|---|
| **Damage variables** | | | |
| Damage to building | Natural logarithm of the monetary damage to the households' building increased by one euro[a], corrected for inflation (in euros) | 7.13 | 2.79 |
| Damage to content | Natural logarithm of the monetary damage to the households' content increased by one euro, corrected for inflation (in euros) | 4.26 | 3.97 |
| **Depth variables** | | | |
| Depth: basement | Depth in the basement (in cm) | 87.86 | 77.76 |
| Depth: ground floor | Depth at the ground floor (in cm) | 14.08 | 21.37 |
| **Control variables** | | | |
| Part of building affected | Dummy variable that takes a value of one if the basement was flooded and zero otherwise[b] | 0.70 | 0.46 |
| | Dummy variable that takes a value of one if the ground floor was flooded and zero otherwise[c] | 0.46 | 0.50 |
| | Dummy variable that takes a value of one if the garage was flooded and zero otherwise | 0.60 | 0.49 |
| Type of building[d] | Categorical variable distinguishing between three types of buildings: (1) Detached; (2) Semi-detached; and (3) Terraced. | – | – |
| Size of dwelling | Size of the dwelling (in squared metres) | 143.52 | 199.23 |
| Recurrence | Number of times that the household has been flooded (times) | 3.39 | 4.67 |
| Risk awareness | Dummy variable that takes a value of one if the household was aware of their house being at risk just before the water entered the building; zero otherwise | 0.42 | 0.49 |
| Emergency measure: elevating contents to another floor | Dummy that takes a value of one if the household moved some of their belongings to a higher floor | 0.12 | 0.32 |
| Income | Categorical variable distinguishing between five income categories: (1) Below 1000 euros; (2) Between 1000 and 1499 euros; (3) Between 1500 and 1999; (4) Between 2000 and 2499; and (5) Above 1500 euros. | – | – |

[a] In order to include the observations where zero damage was reported, we increased the level of damage by one euro before taking the natural logarithm.
[b] Included in those model specifications were the main explanatory variable of interest was the depth of the flood at the ground floor.
[c] Included in those model specifications were the main explanatory variable of interest was the depth of the flood in the basement.
[d] This variable is not included in the models that use the damage to the contents as the explanatory variable of interest.

unworkable if one would like to include many indicators, as it would lead to an endless amount of damage functions. Therefore, a shift to more complex multivariate models has taken place in the literature in recent years. In 2005, Thieken et al. (2005) investigated in addition to the depth of the flood also the influence of other factors (such as flood duration, contamination and preparedness for flood damage) based on survey data for the 2002 floods in Germany. In another paper, Thieken et al. (2008) developed the FLEMOps(+) model. The model includes, apart from flood depth, also building type, building quality, water contamination and private precautionary measures as explanatory variables for damage to private sector buildings. The authors demonstrated that the model had a much stronger explanatory power compared to simple depth–damage functions, a result that was also found by others such as Apel et al. (2009). In 2010 the FLEMOps+ model was extended by Elmer et al. (2010) who also included recurrence as a predicting variable. Recurrence can be important because of its impact on private precautionary measures (long term protection) and emergency measures (during the flood) that can strongly decrease the damage suffered (Kreibich et al., 2011; Siegrist and Gutscher, 2006). A last example of a multivariate model is the one developed by Merz et al. (2013) who used an innovative tree-based damage model to evaluate the impact of no less than 28 indicators. The construction of those more complex models has been achieved only for fluvial flood research. Pluvial flood research is only getting started and is still limited to predicting damage by using one or two hazard indicators (Spekkers et al., 2011, 2012a, 2012b). In the case of pluvial floods, other indicators and specifically non-hazard indicators might be very relevant. For instance, some authors have pointed out that private precautionary measures such as flood adapted building renovations are more useful in the context of small, frequent floods than in the case of huge fluvial floods when the water depth is often too high and velocity too strong to avoid the water from entering anyway (Kreibich et al., 2005). Moreover, in areas with frequent flooding (as another way of precaution) residents sometimes store their valuable objects in the attics and not in the basement or lower floors preventing them from an imminent loss (Stulc, 2007). The effect of emergency measures might also be different with pluvial floods, because pluvial events take place very quickly and often there is almost no time to act.

Thirdly, thanks to the detailed survey, we can construct differentiated models. In current flood damage research, no consideration is made w.r.t. the place where the water enters the building. For residential houses for instance, flood damage to buildings and contents can be expected to differ strongly depending upon which rooms are flooded. Water in the living room will have different (damage) effects from water in the basement or the garage. Our survey distinguished between basement floods and ground floor floods. Also, we made a distinction between damage to the building and damage to contents. This combination enables us to estimate four differentiated models. We pursue as follows. Section 2 describes the data (collection) and the Tobit-methodology. In Section 3 the damage models are discussed. Results (and the limitations) are presented in Section 4 (and Section 5). Finally, conclusions and policy implications are drawn in Section 6.

## 2. Survey data and Tobit estimation technique

We developed a survey that was conducted in 2013 among private households that were presumed to be affected by one or more pluvial floods. We specifically focused on pluvial floods caused by summer storms, as those are the most common and are easy to explain to people in a questionnaire. The survey was sent to 3963 addresses all over Flanders (the northern region of Belgium). The majority of those addresses came from a database of the Belgian national disaster fund that contains addresses of pluvial flood victims. An important setback of the national disaster fund data is the fact that no flood events after 2007 are reported. The reason for this is a change in the legislation that year, obliging insurance companies to provide fire insurance and flood insurance

in one package. This implies that most people now have private insurance, thereby ending the need for government compensations that were provided by the disaster fund (Portaal Belgische overheid, 2012). The lack of recent data was tackled in two ways. Firstly, we included 260 addresses from records of fire and police departments or from local authorities in villages and cities that were flooded in recent years during a pluvial event. Secondly, we asked people to fill in the questionnaire for the most severe flood since 2000 on their address. Many of the addresses in the national disaster fund were expected to be quite prone to pluvial floods, for instance because they are close to malfunctioning sewer systems or in lower parts of the village or city. As such, it was probable that quite some victims become flooded on a regular basis. By asking to take the worst flood as a reference, people could report also about floods after 2007 (21% of the reported floods occurred after the year 2007).

973 households filled in the survey (24.6% of our sample). After a first investigation, we deleted the files of 260 respondents that claimed never to have suffered from pluvial floods. There are two probable reasons for this. Firstly, a number of respondents suffered from damage caused by hail or winds. That type of damage was not separated from flood damage in the disaster fund database. Secondly, some people moved to the address we found in the disaster fund database after the recorded pluvial event took place at that address. The data of another 93 respondents were deleted, mainly because they turned out to be small shopkeepers, farmers or other self-employed businessmen reporting not only about the damage to their home, but also about the damage to their business. Finally, we have the data of 620 respondents on which we perform further analyses in this paper. The average age of 60.03 of our respondents shows that many elderly sent back the questionnaire. Somewhat more men (55.4%) than woman (44.6%) filled in the questionnaire. Table 1 shows some other socio-economic characteristics of the sample. The higher average age is reflected in the fact that over 45% of the respondents indicate being retired. The summary statistics on income indicate that our sample contains victims from all socio-economic classes.

We asked people to report how much damage they suffered because of the flood. First we asked them to report the financial damage to several parts of the building. The reported amounts for the different parts of the building are added to construct a total building damage variable.[1] For damage to the contents, we did not differentiate between different belongings and we asked victims to report immediately on the total damage to their contents.[2] Victims reported about floods in many different years, so the damage data have to be corrected for inflation. For this correction, we used the Belgian Consumer Price Index. 108 respondents did not answer the question about the year of the incident. Most of the remaining 512 respondents (89%) reported on flood incidents that happened after 2001.[3] Table 2 shows the summary statistics: 363 respondents reported a damage figure for their building (and also provide information on the flood year). For contents, there are 413 respondents with a damage figure (that can be corrected for inflation). Clearly the

---

[1] For building damage subcategories, respondents had to report either to have no damage or stipulate the amount of damage. It happens that none of both options is used. For the 240 respondents (used in the regressions later on) flooded in the basement, 55 persons had a missing on at least one subcategory. For the 141 respondents flooded at the ground floor, 31 had a least one missing. When summing up to construct total building damage, the missing amount had to be set at zero value. This could imply a downward bias for the total building damage.

[2] We provided the opportunity for respondents to report also 'other' damage. In case they report damage in this 'other' category that clearly should have been reported as damage to buildings or to contents, this was added to the total building or total contents damage.

[3] The distribution over time of the year in which respondents reported to have experienced their worst flood experience is as follows: 11% of the respondents reported on a flood that dated from before 2002, 16% of the respondents reported on a flood that dated from the period 2002–2004; 46% of the respondents reported on a flood that dated from the period 2005–2006; 9% of the respondents reported on a flood that dated from the period 2007–2009 and 19% of the respondents reported on a flood that dated from the period 2010–2012.

**Table 4**
Missing values for the outcome and explanatory variables and reported zero damages.

| Kind of flood/damage to | Total floods | Missing damage figure | Missing year | Flood damage data | Missing explanatory variables | Observations in regression | Zero damage |
|---|---|---|---|---|---|---|---|
| Basement/building | 430 | 133 | 36 | 261 | 21 | 240 | 26 |
| Basement/content | 430 | 111 | 28 | 291 | 31 | 260 | 108 |
| Ground floor/building | 283 | 94 | 26 | 163 | 22 | 141 | 5 |
| Ground floor/content | 283 | 82 | 15 | 186 | 34 | 152 | 55 |

damage to the building is on average much higher than to the contents. This can be partly explained by the higher share of zero damage observations for the category "damage to the contents". In addition, we find that the total damage reported by the households is substantially higher than the payment they received from their insurance company or the Belgian Rampenfonds.

For both building and content damage, and especially for content damage, victims also reported zero damage. These respondents were the victim of a pluvial flood (cfr. the selection of the sample) but have no reported damage to their building or their contents. It can also be the case that they do not report damage that is very small. This means that in case the damage is lower than c (which we assume not necessarily to be zero) the individuals report this as a zero damage case. This is relevant if we want to estimate the impact of e.g. risk awareness or emergency measures. A large number of zero observations lead to severely biased regression coefficients when using OLS. One solution is simply to delete the zero values, i.e. truncating the sample. However, this would mean ignoring relevant information and therefore would also lead to inconsistent estimations of the parameters (Wooldridge, 2010). The problem of many zero values is well-known in the field of economics. A specific model has been developed by James Tobin (1958) to deal with this issue, known as the Tobit model (James Tobin's Probit model). The methodology is widely applicable for any situation where observations are 'censored' above or below a certain border value. The Tobit model is therefore also known as a specific type of 'censored regression model' (Wooldridge, 2010) and is used widely for analysing diverse micro-economic topics such as the hours worked by different socio-economic groups (see McDonald and Moffitt (1980)) or explaining charity donations. To our knowledge, the Tobit model has never been used before in flood damage research. A mathematical elaboration of the Tobit model can for example be found in Long (1997) or Sigelman and Zeng (1999). In our case of flood damage, it seems reasonable to claim that factors such as flood depth, flood duration, building characteristics, behavioural predictors (such as emergency measures) and socio-economic indicators all will contribute in a very comparable way to the amount of positive damage on the one hand and to whether or not there is any damage on the other hand. This is a necessary condition to use the Tobit model (see for example Wooldridge, 2010).

The Tobit coefficients refer to the latent variable, and this is not the effect we are really interested in. We do want to take into account the zero values in the estimations, but in the end we only want to know the effect of the explanatory variables on the actual observed damage (comparable to the interpretation of an OLS coefficient). Therefore, in the next section, we calculate the average marginal effects of the hazard and non-hazard indicators on the observed (or censored) damage. The marginal effect is equal to the Tobit coefficient multiplied by the probability of obtaining an uncensored observation. It is calculated by the following formula (Collis et al., 2010; Sigelman and Zeng, 1999):

$$\frac{\partial E(y|x)}{\partial x} = \beta \Phi\left(\frac{x\beta}{\sigma}\right). \tag{1}$$

The left hand side of this formula is the derivative of y with respect to x, which is the marginal effect by definition. The right hand side shows that the marginal effect equals the original Tobit parameter of x ($\beta$) multiplied by the estimated probability of obtaining an uncensored observation $\Phi(x\beta / \sigma)$ where $\Phi$ is the cumulative density function from

the standard normal distribution and $\sigma$ is the variance (Collis et al., 2010). It is clear from this formula that the marginal effect depends upon the value of x. To deal with this, we prefer to calculate average marginal effects instead of marginal effects at the mean, which is in line with modern practices (for a discussion on the topic of marginal effects at the mean and average marginal effects, see Bartus (2005)).

## 3. Differentiated flood models

The questionnaire made a distinction between basement floods, ground floor floods and garage floods. Water in the garage is not investigated in itself, it is taken up as a control variable. As we also make a distinction between damage to the building and damage to the contents (see Table 1), we will construct four different (refined) models[4]:

- Damage to the building for respondents that were flooded in the basement
- Damage to the contents for respondents that were flooded in the basement
- Damage to the building for respondents that were flooded at the ground floor
- Damage to the contents for respondents that were flooded at the ground floor.

Table 3 presents the outcome and explanatory variables included in the different model specifications.

In Table 4, we summarize the information about the number of observations used in the regressions and about the number of reported zero damages. Starting with 620 respondents, 56 of them reported being flooded only in the garage or in 'another place' (with different specifications). This brings the number of useful respondents at 564. Out of these 564 respondents, 134 respondents were flooded at the ground floor but not in the basement, 281 were flooded in the basement but not at the ground floor and 149 were flooded in both. As a result, 283 (134 + 149) respondents can be included in ground floor models and 430 (281 + 149) can be included in basement models.

For many respondents, it seemed difficult to estimate and report the damage caused by the flood. That is both the cases for content damage and for building damage. Those respondents did not report a positive nor a zero damage figure. Out of 430 people that reported being flooded in the basement, only 297 reported a positive or a zero value for damage to the building and 319 for damage to contents. For ground floor floods, 189 people reported damage figures to the building and 201 reported damage figures for contents. Also, a number of the respondents (that do report a valid damage value) did not fill in the year the flood took place. For those respondents it was impossible to calculate the damage figures corrected for inflation.

Missing values on the flood variable (damage or year) lead to a fall-out between 32% (basement flood/content damage-model) and 42% (ground floor flood/building damage-model). The main reason clearly is that people have difficulties in estimating or reporting the damage caused by the floods. We performed a non-response analysis making a

---

[4] Note that we have tried to include the depth of the flood in the basement and the depth of the flood at the ground floor as two explanatory variables in one regression. However, both variables are strongly correlated (corr. 0.3942) and including them in one regression causes problems related to multicollinearity.

**Table 5**
Tobit model results — original coefficients.

| Variables | Tobit<br>Basement flood — building damage (ln) | Tobit<br>Basement flood — content damage (ln) | OLS<br>Ground floor flood — building damage (ln) | Tobit<br>Ground floor flood — content damage (ln) |
|---|---|---|---|---|
| Constant | 4.6119*** | 1.9542* | 6.6090*** | −0.1048 |
| | (0.6475) | (1.1687) | (0.5663) | (0.4092) |
| Depth basement/gr floor | 0.0114*** | 0.0129** | 0.0331*** | 0.0775** |
| | (0.0026) | (0.0052) | (0.0078) | (0.0227) |
| Gr floor/basement flooded | 0.8880** | 0.6690 | 0.8599** | 1.4913 |
| | (0.8880) | (0.8418) | (0.3623) | (0.9251) |
| Garage flooded | 0.3596 | 1.2074 | −0.1333 | −0.8475 |
| | (0.4176) | (0.8038) | (0.3199) | (1.0062) |
| Type: semi-detached | 0.3045 | | 0.3039 | |
| | (0.4810) | | (0.3896) | |
| Type: terraced | −0.3604 | | −1.0380** | |
| | (0.4756) | | (0.4215) | |
| Size of dwelling | | | 0.0017 | 0.0040 |
| | | | (0.0010) | (0.0040) |
| Recurrence | 0.0172 | −0.2633** | 0.0996 | 0.2627 |
| | (0.0430) | (0.1086) | (0.0861) | (0.2361) |
| Risk awareness | 0.3383 | −3.6354*** | −0.1341 | −2.0096* |
| | (0.4145) | (1.0207) | (0.3276) | (1.0619) |
| Emergency measure: elevating contents to another floor | | 4.2229*** | | 1.3787 |
| | | (1.3211) | | (1.5646) |
| Income below 1000 | 1.6063* | −1.7429 | 0.0599 | 0.0112 |
| | (0.8143) | (1.6472) | (0.8180) | (2.1828) |
| Income 1500–1999 | 0.8185 | 0.5940 | −0.0588 | 3.7098*** |
| | (0.5691) | (1.1300) | (0.4526) | (1.3548) |
| Income 2000–2499 | 0.9777 | 2.9990** | −0.2734 | 2.6687* |
| | (0.61108) | (1.2166) | (0.5075) | (1.4875) |
| Income 2500+ | 0.4250 | −0.4502 | 0.6704 | 2.5716 |
| | (0.7161) | (1.3874) | (0.5594) | (1.7098) |
| Obs/Pos | 240/214 | 260/152 | | 152/97 |
| Obs/Cens | 240/26 | 260/108 | | 152/55 |
| Log lik. | −563.267 | −564.440 | | −343.300 |
| Prob > chi² | 0.000 | 0.000 | | 0.001 |
| (Corr y ~ ^y)² (Tobit)/ R-squared (OLS) | 0.1815 | 0.1516 | 0.2836 | 0.2055 |

Standard errors in parentheses.
  * Significant at 0.10 level.
 ** Significant at 0.05 level.
*** Significant at 0.01 level.

comparison between respondents with missing values and those with valid values on the inflation corrected damage figures. We did this for both damage to building and damage to contents for a number of important characteristics (see Appendix A). The non-response analysis shows that there exist some differences between the respondents included in our models and the respondents that failed to answer a number of crucial questions. The differences are however restricted and any bias would thus be quite limited.

## 4. Results and discussion

For the two basement flood models and for the ground floor flood-content damage model, we used a Tobit-estimation because a high number of reported damage figures are equal to zero. In contrast, given the limited number of zero damage observations, we estimated the model for ground floor flood-building damage using an OLS regression.[5] The estimation results of the four models are presented in Table 5.

Our findings show that a number of the explanatory variables, including the depth of the flood, are highly significant in explaining the

damage. Moreover, the estimated models explain a reasonable share of the variation in the damage variables. The squared correlation between the predicted values by the models and the actually observed values is a measure for explained variation and can be used for comparison with $R^2$ in OLS regressions (Wooldridge, 2010). For the Tobit models, the squared correlation ranges between 15% and 21%. The OLS model (to estimate the ground floor model on building damage) explains 28% of the variance.[6] The marginal effects of the Tobit and OLS models are presented in Table 6.[7]

Looking at Table 6, it is clear that the depth of the flood is an important predictor for damage, also in this case of pluvial flood analysis. The effect of depth is highly significant and also important in terms of magnitude for all four models. For the basement models, there is an increase

---

[5] Although not necessary given the low number of zero values, it is possible to calculate Tobit coefficients for the analysis of building damage during ground floor floods instead of using the OLS analysis. When doing this, we find that the Tobit coefficients are very similar to the OLS coefficients. This is very logical given that the Tobit model uses OLS for the non-censored values and almost all observations (except for five zeros) are uncensored in that case. This is not the case when comparing OLS-coefficients with Tobit-coefficients for the three other models. The results of the alternative estimations are available from the authors upon request.

[6] Note that by estimating multivariate models in which in addition to the depth of the flood also a number of control variables are included, a substantial larger share of the variation in the damage is explained as compared to a restricted model in which only the depth of the flood is included. In fact, in case only the depth of the flood would be included as an explanatory variable, the explained variation ranges between only 5% and 10% for our four model specifications. This underlines the importance of estimating multivariate depth–damage models.

[7] Because of the logarithmic transformation of the outcome variable, the exponent of the coefficients was taken, decreased by one and multiplied by 100 to obtain the effect of an increase of one unit in the explanatory variables on the outcome variable in percentage changes. When doing a logarithmic transformation of the outcome variable, coefficients should be recalculated by the following formula to obtain the effects in percent changes:

$$\%\Delta y = 100 * (e^{\beta_i} - 1).$$

**Table 6**
Differentiated models, based on average marginal Tobit effects and OLS results (percentage effects).

| Variables | Tobit Basement flood — building damage | Tobit Basement flood — content damage | OLS Ground floor flood — building damage | Tobit Ground floor flood — content damage |
|---|---|---|---|---|
| Depth basement/ground floor | 1.14%*** | 0.88%** | 3.37%*** | 5.88%*** |
| Part: Ground floor/basement flooded | 140.70%** | 57.43% | −12.48% | 197.96% |
| Part: garage flooded | 42.59% | 123.56% | 136.29%** | −46.83% |
| Type: Semi-detached | 35.22% | | 35.51% | |
| Type: terraced | −29.95% | | −64.58%** | |
| Size of dwelling | | | 0.17% | 0.29% |
| Recurrence | 1.70% | −16.21%** | 10.47% | 21.36% |
| Risk awareness | 39.65% | −90.15%*** | −12.55% | −77.08%* |
| Emergency measure: elevating contents to another floor | | 2264.61%*** | | 186.74% |
| Income below 1000 | 389.40%** | −64.84% | 6.17% | 0.72% |
| Income 1500–1999 | 123.92% | 48.99% | −5.71% | 1548.91%*** |
| Income 2000–2499 | 162.16% | 809.02%** | −23.92% | 596.29%* |
| Income 2500+ | 51.83% | −25.04% | 95.50% | 544.17% |

  * Significant at 0.10 level.
 ** Significant at 0.05 level.
*** Significant at 0.01 level.

of 1.14% in building damage and of 0.88% in content damage for each extra centimetre of water depth. The depth coefficients are much higher in the ground floor flood models, with an increase in damage to the building of 3.37% and to contents of 5.88% for each additional centimetre of water depth. It is no surprise that the same increase in the depth level causes more severe damage (to buildings and contents) in case of ground floor floods compared to basement floods. On the ground floor, the contents will be much more valuable and more valuable building elements can be damaged. Fig. 1 presents the depth–damage relationship for the different model specifications graphically. Panel A for the damage to the building caused by basement floods, Panel B for the damage to the content caused by basement floods, Panel C for the damage to the building caused by ground floor floods and Panel D for the damage to the content caused by ground floor floods. For each panel, there are two graphs. The first graph presents the relationship between the depth of the flood and the natural logarithm of the damage, which is linear. The second graph presents the relationship between the depth of the flood and the damage, which is exponential.

Our findings illustrate the importance of using differentiated models, here making a distinction based upon where the water enters the building. It can be expected that people that were flooded also in the garage or people that were flooded in both the basement and at the ground floor will suffer more damage. Therefore we include a dummy variable in case people were also flooded in the garage and another dummy in case people were flooded in the other of the two options. This latter dummy variable takes a value of one when also the ground floor was flooded (in case the explanatory variable of interest was the depth of the flood in the basement) or when the basement was flooded (in case the explanatory variable of interest was the depth of the flood in the ground floor). The results show that these control variables have no effect on content damage. They are significant in two cases for building damage. First, in case of a basement flood, the fact that the ground floor was also flooded increased the damage to the building by 141% compared to a situation in which the ground floor was not flooded. Second, in case of a ground floor flood, the fact that the garage was also flooded increased the damage to the building by 136% compared to a situation in which the garage was not flooded.

Survey information provides the opportunity to include more non-hazard indicators. We incorporated building characteristics and socio-economic and behavioural information about the flood victims. Again differentiating, the type of the building is considered to be relevant for damage to the building, not for damage to contents. The reference category of detached houses is compared to semi-detached houses and terraced buildings. The type of building has no significant effect on damage

to buildings in the case of basement floods. For ground floor floods the building damage for terraced houses is on average 65% lower than the building damage to detached houses. We have information about the size of the dwelling at the ground floor, but this turns out not to be a significant predictor for any of the two ground floor models, yet the P-value is at the margin in the case of building damage (0.105).

Another category of variables are predictors related to the behaviour of victims before and during the floods. There are important effects in the models predicting damage to contents in the case of basement flooding. Recurrence refers to the amount of times people have been flooded. People seem to learn from those experiences: in the case of basement floods the content damage decreases on average by 16% for each new flood. Risk awareness is captured by a dummy variable in case people were aware of their house being at risk just before the water entered the building. Being aware of the risk just before the water enters the building reduces content damage on average by 90% in case of basement floods and by 77% in case of ground floor floods. Such a finding underlines the importance of warning systems and policy in this case of pluvial floods. We included an emergency measure in our models that predicts damage to contents, namely elevating contents to another floor. Unexpectedly, victims moving their contents from the basement to a higher level seem to suffer significantly more damage. Perhaps those people desperately try to protect their belongings by moving them to a higher floor.[8] Finally, we included income, which is correlated with the level of education and the labour market situation of the individual. The reference category here is respondents with a net monthly income between 1000 and 1500 euros. It seems, in general, that people with a higher income suffer more content damage. In the case of building damage during basement floods, people earning less than 1000 euros net a month suffer more damage compared to people in the reference category.

## 5. Limitations and scope for future research

The present study extends the traditional bivariate depth–damage relation by constructing more detailed multivariate flood damage models including no-damage cases. However, there are a number of

---

[8] The number of respondents taking this measure is also limited. On a total of 620 respondents, only 68 mention to take the measure. For the specific model on content damage during basement floods, 35 respondents took this measure. For the model on content damage during ground floor floods, only 19 respondents responded positively on the question whether they took the measure on a total of 152 observations in the model. Leaving out the variable from the model changes very little to the coefficients of other variables, although the overall explanatory power of the models does decrease somewhat.

limitations that provide scope for future research. First, the present study reflects only a part of the total damage caused by floods and additional information is required to be incorporated in any cost–benefit analysis on flood management. Other authors have pointed to damage to commercial and public property (Kreibich et al., 2010) and damage caused by other types of floods, such as ground water floods (Kreibich and Thieken, 2008). Also, pluvial floods can have a large impact on non-monetary damage as they may have a significant impact on the health of those individuals that have been flooded. In addition to a physical non-monetary impact pluvial floods can also affect the well-being of individuals. In Van Ootegem and Verhofstadt (2015) we analyse the impact of pluvial floods on well-being and in particular on life satisfaction and capabilities.

Second, the inclusion of non-hazard indicators is an important strength of our paper. However, it may make the model less suitable for making predictions of the damage as this requires information on the non-hazard indicators and often this type of information is not easily available. This issue becomes particularly problematic when the population for which predictions are made is very different from our sample in terms of its non-hazard properties. In addition, it is unclear whether the existing hydraulic models will be able to simulate the water depth as it has been interpreted in our models, namely the subjective reported depth at the ground floor and in the basement.

Third, the models that we present explain between 15% and 28% of the variation in the damage caused by the flood. These figures are in line with other economic models. Nevertheless, the explanatory power of the model could be further improved by including additional control variables that could explain the damage caused by the flood, such as variables related to the characteristics of the flood (e.g. velocity or the level of sediment of the water that entered the property) or variables related to the socio-economic characteristics of the households affected by the flood (e.g. total value of the building or contents owned or rented by the household or detailed data on the wealth of individuals instead of income data).

Fourth, the data are collected based on recall from an event that may have happened more than ten years ago. This may raise questions regarding the accuracy of some of the variables. In particular, it is unclear to what extent respondents are able to remember in detail the amount of the damage or the depth of the flood. In order to test for the robustness of our results, we restricted the sample to the past ten years (period 2002–2012) since we expect that respondents are better able to remember recent events. This did not change the results significantly (results are available from the authors upon request).

Finally, in our model specifications the main explanatory variable of interest is the depth of the flood. Other authors have estimated a direct relationship between the rainfall intensity and the monetary damage (e.g. Spekkers et al., 2014). Such rainfall-damage models have several advantages. Rainfall data are completely exogenous to
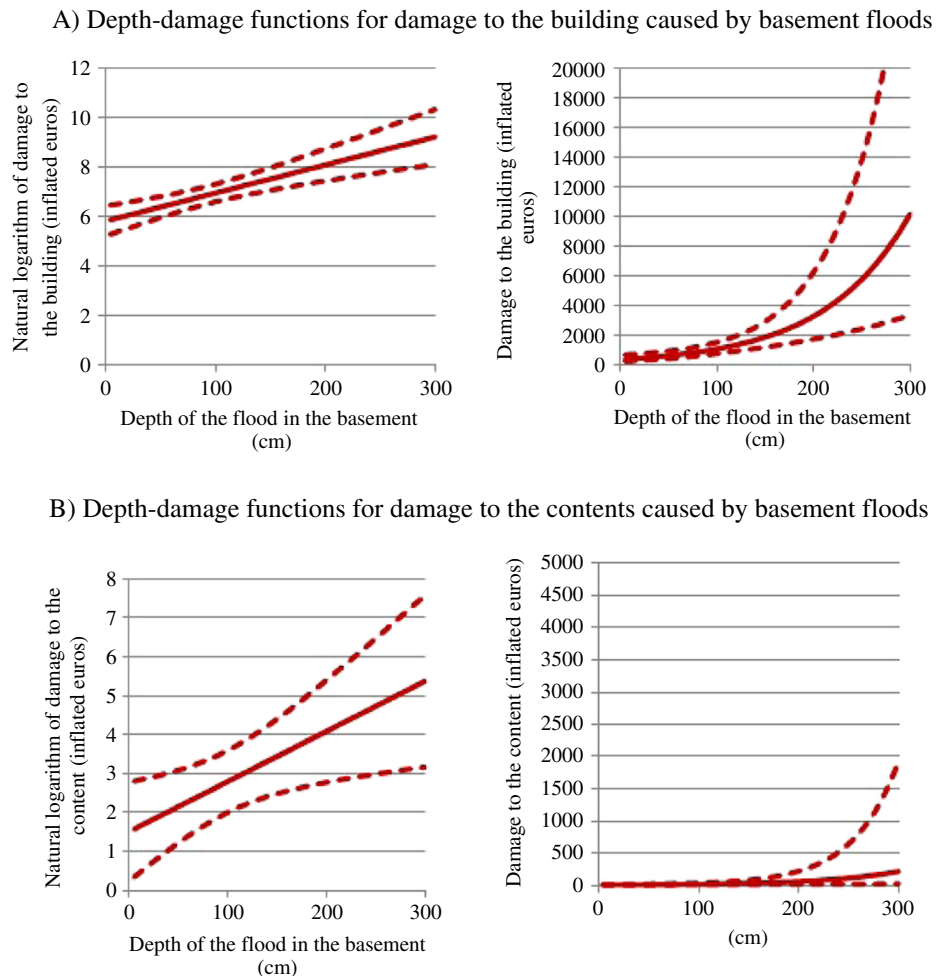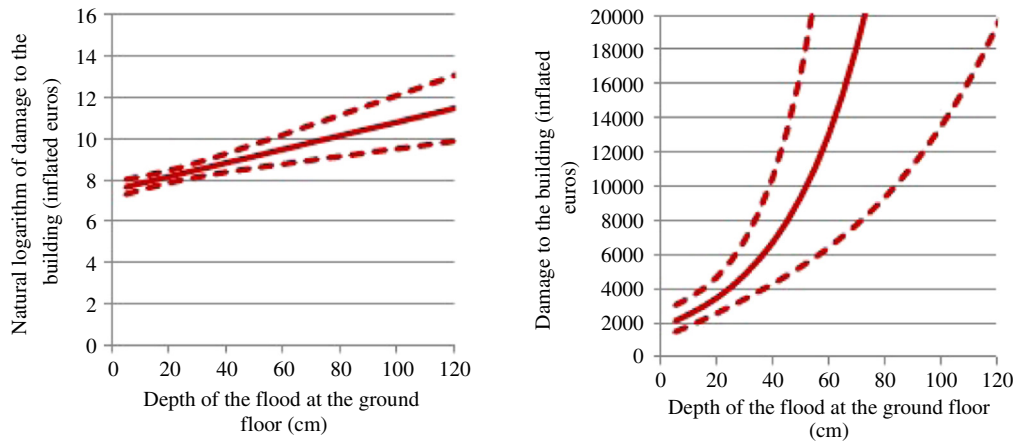
A) Depth-damage functions for damage to the building caused by basement floods



B) Depth-damage functions for damage to the contents caused by basement floods



**Fig. 1.** Graphical representation of the estimated depth–damage functions.

C) Depth-damage functions for damage to the building caused by ground floor floods



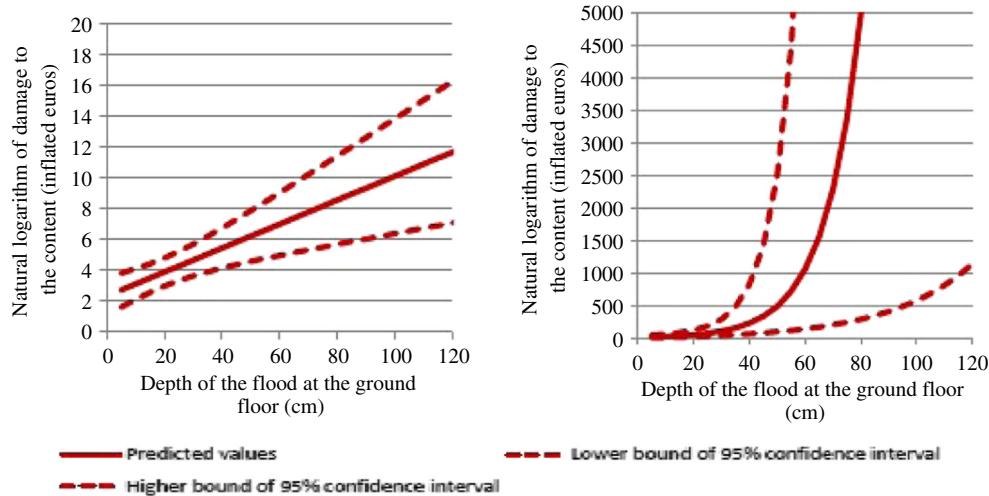D) Depth-damage functions for damage to the contents caused by ground floor floods



**Fig. 1** (*continued*).

the other control variables while depth is not necessarily exogenous. In fact, the depth of the flood could be affected by some of the control variables, which could bias the estimated coefficient on the depth of the flood. In addition, rainfall-damage could be rapidly implemented in early warning systems, while for the depth–damage models the depth first needs to be simulated. This is in particularly important with respect to pluvial floods, which occur rather unexpectedly and have in general a short lead time. Only few authors have estimated "rainfall-damage" models. For example, Zhou et al. (2013) analysed the impact of the rainfall on 1000 insurance claims related to pluvial floods in Aarhus (Denmark). They find that the level of rainfall did not significantly affect the cost per claim, but it did affect the total costs per day. Spekkers et al. (2013) find that rainfall intensity has a significant impact on the monetary damage. However, overall the explained variance in case of the rainfall-damage models is rather low, which emphasizes the need to further study these types of models and include additional explanatory variables, such as the socio-economic characteristics of the households that have been flooded.

## 6. Conclusions

We explore a number of approaches and issues that are not customary in flood damage research. First of all, we investigate pluvial floods caused by heavy rainfalls, this instead of the more traditional focus on fluvial or coastal floods. We use self-reported data making use of a questionnaire conducted among flood victims in the northern Dutch speaking part of Belgium. The data-base offers possibilities to investigate a number of novelties because we have information about cases with reported zero damage, we have detailed non-hazard information (building characteristics, behavioural responses and socio-economic data) and we have information about the specific places that are flooded (basement, ground floor or garage) and the kind of damage that is suffered (damage to contents or damage to the building). This allows one to develop (four) differentiated multivariate flood damage models. We differentiate between ground floor floods and basement floods and between damage to the building and to contents. We apply a Tobit estimation technique that specifically deals with the issue of reported zero damage.

The combination of the differentiated models with also zero damage information shows that the damage effect of an increase in the flood depth is significantly stronger for ground floor floods than for basement floods. For the basement flood models, there is an increase of about 1% damage to the building or the contents for each extra centimetre of water depth. For the ground floor flood models, damage to the building increases with 3% and damage to the contents with 6% with each additional centimetre of water. This is because the contents will be much more valuable on the ground floor and more valuable building elements can be damaged on the ground floor. This result is also due to the fact that there is a high number of reported zero damages for content damage in the case of basement floods. The Tobit estimation technique makes it possible to include that kind of information. For three of the four models, an important number of no damage cases occur and the Tobit-models yield results different from OLS-models.

Our results show that, although flood depth remains an important predictor of flood damage (as it is also the case for fluvial floods), also non-hazard indicators are important for explaining pluvial flood damage. We find important effects on damage of building characteristics, behavioural predictors and income. Risk awareness and other behavioural predictors only affect the damage to contents and are more important in the case of basement floods than in the case of ground floor floods. Being aware of the risk before the water enters the building reduces content damage on average by 90% in case of basement floods and by 77% in case of ground floor floods. These results point to the potential gain of policies that increase the awareness or the knowledge of flood risks. Then, people can protect their belongings before the flood takes place, and so before the flood (depth) can cause damage. Also, people learn from past flooding. In the case of basement floods the content damage decreases for each new flood.

We can summarize that the most important hazard indicator (depth of the flood) is more important because of its damage-increasing effect in case of ground floor floods, while the non-hazard indicators are more important for their potential damage-decreasing effect; this is especially for the content damage and specifically for basement floods. Non-hazard indicators can partially be influenced by (warning) policies. A cost benefit analysis could be conducted using these results.

## Appendix A. Non-response analysis

The non-response analyses in Tables A1 and A2 show us that there are few significant differences in water depth levels between respondents with valid and with missing damage figures. Only for the ground floor model predicting content damage, there is a statistically important difference as respondents with missing damage values report significantly higher water levels compared to respondents that were able to estimate damage figures. As depth and damage are highly correlated, the results are an indication that respondents with a lot of damage are more likely not to report any figure, possibly because the damage was so severe that they are not able to estimate it anymore. As a consequence, the ground floor model for content damage is likely to lead to somewhat underestimated damage figures as a number of respondents with a lot of damage are probably excluded.

As for socio-economic indicators, respondents with missing values are significantly (up to five years) older than those who are capable of estimating their damage. It is probably no surprise that older people have more difficulties in estimating damage figures. Also, respondents with missing values are overrepresented among the lower income groups and even more among respondents that are not willing to report their income. The latter result is also quite logical as a number of people did not fill in the questionnaire very rigorously and as such they have missing values on many variables, among which also both the damage figures and income. Women and lower educated groups are also overrepresented among respondents that do not report damage figures for buildings.

Finally, we wanted to check whether there is a time effect. Indeed, people with missing values on damage figures for buildings report on floods that occurred further in the past, about 10 months on average. Again, the result is not a surprise. As time passes, people will very probably find it harder and harder to recall the damage they suffered.

Table A1. Non-response analysis for building damage.

| Variable | Test | Test statistic | P-value | Conclusion |
|---|---|---|---|---|
| Depth Ground floor | T-test | 0.771 | 0.411 | No significant difference |
| Depth Basement | T-test | −0.994 | 0.321 | No significant difference |
| Age | T-test | 4.165 | 0.000 | Respondents with missing values are significantly older (+/−5 years) than other respondents |
| Gender | Chi$^2$ | 5.001 | 0.025 | Respondents with missing values are significantly more often women than other respondents |
| Education | Chi$^2$ | 13.374 | 0.010 | Respondents with missing values are overrepresented among the lower educated groups |
| Income | Chi$^2$ | 25.603 | 0.000 | Respondents with missing values are overrepresented among the lower income groups and even more among respondents not reporting income |
| Month passed | T-test | 2.067 | 0.039 | Respondents with missing values report on floods that happened earlier (+/−10 months) |

Table A2. Non-response analysis for content damage.

| Variable | Test | Test statistic | P-value | Conclusion |
|---|---|---|---|---|
| Depth Ground floor | T-test | 2.027 | 0.043 | Respondents with missing values have significantly higher water levels than other respondents (+/−5 cm) |
| Depth Basement | T-test | −0.852 | 0.395 | No significant difference |
| Age | T-test | 0.011 | 0.002 | Respondents with missing values are significantly older (+/−4 years) than other respondents |
| Gender | Chi$^2$ | 1.853 | 0.173 | No significant difference |
| Education | Chi$^2$ | 7.495 | 0.112 | No significant differences |
| Income | Chi$^2$ | 23.199 | 0.000 | Respondents with missing values are overrepresented among the lower income groups and even more among respondents not reporting income |
| Month passed | T-test | 1.900 | 0.058 | No significant difference (albeit at the margin) |

## References

Apel, H; Aronica, G T; Kreibich, H & Thieken, A H.(2009), Flood risk analyses — how detailed do we need to be? Natural Hazards, v. 49, no. 1, p. 79–98.

Bartus, T. (2005), Estimation of marginal effects using margeff: Stata Journal, v. 5, no. 3, p. 309–329.

Blanc, J., Hall, J., Roche, N., Dawson, R., Cesses, Y., Burton, A. & Kilsby, C. (2012), Enhanced efficiency of pluvial flood risk estimation in urban areas using spatial–temporal rainfall simulations, J. Flood Risk Manage, v. 5, p. 143–152.

Bruwier, M., Erpicum, S., Pirotton, M., Archambeau, P. & Dewals, B.J. (2015), Assessing the operation rules of a reservoir system based on a detailed modelling chain, Nat. Hazards Earth Syst. Sci., v. 15, p. 365–379.

Collis, J., Grayson, A., Johal, S., 2010. Econometric Analysis of Alcohol Consumption in the UK. HM Revenue & Customs, London.

Elmer, F.; Thieken, A. H.; Pech, I. & Kreibich, H. (2010), Influence of flood frequency on residential building losses: Natural Hazards and Earth System Sciences, v. 10, p. 2145–2159.

European Environment Agency, 2012. Urban Adaptation to Climate Change in Europe. European Environment Agency, Copenhagen.

Jonkman, S. N.; Bockarjova, M.; Kok, M & Bernardini, P. (2008), Integrated hydrodynamic and economic modelling of flood damage in the Netherlands: Ecological Economics, v. 66, no. 1, p. 77–90.

Kellens, W.; Vanneuville, W.; Verfaillie, E.; Meire, E.; Deckers, P. & De Maeyer, P. (2013), Flood risk management in Flanders: past developments and future challenges, Water Resources Management, v. 27, no. 10, p. 3585–3606.

Kreibich, H.; Seifert, I.; Thieken, A. H.; Lindquist, E.; Wagner, K. & Merz, B. (2011), Recent changes in flood preparedness of private households and businesses in Germany, Regional Environmental Change, v. 11, no. 1, p. 59–71.

Kreibich, H.; Seifert, I., Merz, B. and Thieken, A. H. (2010), Development of FLEMOcs — a new model for the estimation of flood losses in the commercial sector, Hydrological Sciences Journal, v. 55, no. 8, p. 1302–1314.

Kreibich, H.; Piroth, K.; Seifert, I.; Maiwald, H.; Kunert, U.; Schwarz, J.; Merz, B. & Thieken, A. H. (2009), Is flow velocity a significant parameter in flood damage modelling?, Natural Hazards and Earth System Sciences, v. 9, p. 1679–1692.

Kreibich, H. & Thieken, A. H. (2008), Assessment of damage caused by high groundwater inundation, Water Resources Research, v. 44, no. 9.

Kreibich, H.; Thieken, A. H.; Petrow, T.; Müller, M & Merz, B. (2005), Flood loss reduction of private households due to building precautionary measures—lessons learned from the Elbe flood in August 2002, Natural Hazards and Earth System Science, v. 5, no. 1, p. 117–126.

Long, J.S., 1997. Regression Models for Categorical and Limited Dependent Variables. Sage.

McDonald, J. F. & Moffitt, R. A. (1980), The uses of Tobit analysis: Review of Economics and Statistics, v. 62, no. 2, p. 318–321.

Merz, B., Kreibich, H., Lall, U., 2013. Multi-variate flood damage assessment: a tree-based data-mining approach. Natural Hazards and Earth System Sciences, v. 13, 53–64.

Messner, F., Meyer, V., 2006. Flood damage, vulnerability and risk perception — challenges for flood damage research. In: Schanze, J., Zeman, E., Marsalek, J. (Eds.), Flood Risk Management: Hazards, Vulnerability and Mitigation Measures. Springer Netherlands, pp. 149–167 NATO Science Series.

Michel-Kerjan, E. O. & Kousky, C. (2010), Come rain or shine: evidence on flood insurance purchases in Florida, The Journal of Risk and Insurance, v. 77, no. 2, p. 369–397.

Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J., Green, C., 2005. The Benefits of Flood and Coastal Risk Management: A manual of Assessment Techniques (The Multi-coloured Manual). Middlesex university Flood Hazard Research Centre. Middlesex University Press, London.

Portaal Belgische overheid, 2012. Het Rampenfonds. http://www.belgium.be/nl/huisvesting/huisvestingsproblemen/natuurrampen/rampenfonds/.

Siegrist, M. & Gutscher, H. (2006), Flooding risks: a comparison of lay people's perceptions and expert's assessments in Switzerland, Risk Analysis, v. 26, no. 4, p. 971–979.

Sigelman, L. & Zeng, L. (1999), Analyzing censored and sample-selected data with Tobit and Heckit models, Political Analysis, v. 8, no. 2, p. 167–182.

Spekkers, M. H., Kok, M., Clemens, F. H. L. R. & J. A. E. ten Veldhuis (2014), Decision-tree analysis of factors influencing rainfall-related building structure and content damage, Natural Hazards and Earth System Science, v. 14, no. 9, p. 2531–2547.

Spekkers, M. H., Kok, M., Clemens, F. H. L. R. & J. A. E. ten Veldhuis (2013), A statistical analysis of insurance damage claims related to rainfall extremes, Hydrol. Earth Syst. Sci., v. 17, p. 913–922.

Spekkers, M.H., ten Veldhuis, J.A.E., Kok, M., Clemens, F.H.L.R., 2012a. Correlations between rainfall data and insurance damage data on pluvial flooding in The Netherlands. 10th International Conference on Hydroinformatics. Hamburg.

Spekkers, M. H.; Kok, M.; Clemens, F. H. L. R. & ten Veldhuis, J. A. E. (2012b), A statistical analysis of insurance damage claims related to rainfall extremes, Hydrology and Earth System Sciences Discussions, v. 9, no. 10, p. 11615–11640.

Spekkers, M. H.; ten Veldhuis, J. A. E.; Kok, M. & Clemens, F. H. L. R. (2011), Analysis of pluvial flood damage based on data from insurance companies in The Netherlands. On-line accessible at: http://repository.tudelft.nl/view/ir/uuid%3Ade414d3d-fe45-459a-bfd6-f46b91d5a7cd/.

Stulc, J., 2007. The 2002 Floods in the Czech Republic and their Impact on Built Heritage. In: Meier, H.R., Petzet, M., Will, T. (Eds.), Heritage at Risk — Cultural Heritage and Natural Disasters. Risk Preparedness and the Limits of Prevention. ICOMOS TUDPress, Dresden.

ten Veldhuis, J. A. E. (2011), How the choice of flood damage metrics influences urban flood risk assessment: Journal of Flood Risk Management, v. 4, no. 4, p. 281–287.

Thieken, A.H., Olschewski, A., Kreibich, H., Kobsch, S., Merz, B., 2008. Development and evaluation of FLEMOps — a new Flood Loss Estimation MOdel for the private sector. In: Proverbs, D., Brebbia, C.A., Penning-Roswell, E. (Eds.), Flood Recovery, Innovation and Response I. WIT Press, London, pp. 315–324.

Thieken, A. H., Petrow, T., Kreibiich, H. & Merz, B. (2006), Insurability and mitigation of flood lossses in private housholds in Germany, Risk Analysis, v. 26, no. 2, p. 383–395.

Thieken, A. H., Müller, M., Kreibich, H. & Merz, B. (2005), Flood damage and influencing factors, New Insights From the August 2002 Flood in Germany, v. 41, no. 42.

Tobin, J. (1958), Estimation of relationships for limited dependent variables, Econometrica, v. 26, no. 1, p. 24–36.

Van Ootegem, L, Verhofstadt, E., 2015. Well-being, life satisfaction and capabilities of flood disaster victims mimeo.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.

Zhou, Q.,Mikkelsen, P. S.; Halsnaes, K. & Arnbjerg-Nielsen, K. (2012), Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits, Journal of Hydrology, v. 414, p. 539–549.

Zhou, Q., Panduro, T. E., Thorsen, B. J. & Arnbjerg-Nielsen, K. (2013) Verification of flood damage modelling using insurance data, Water Sci. Technol., v. 68, p. 425–432.

**Luc Van Ootegem** is a professor in Economics at the University Gent (Faculty of Economics and Business Administration) where he teaches Welfare Economics, Environmental Economics and Social and Economic Policy. He has a Ph.D. in Economics from the KU Leuven (1995). From 2003 till 2008 he was responsible for the Master program in Applied Economics and Business Administration at the University College Gent. Since then, he is combining teaching and research on well-being (measurement and evaluation) and applied welfare economics. He is part-time project manager in the domain of research on well-being and sustainable development and environmental policy at the Research Institute for Work and Society (HIVA, KU Leuven) in the Research Team on "Environmental Policy and Sustainable Development".

**Elsy Verhofstadt** (PhD Economics) is an assistant professor at Ghent University, Department of Social Economics and affiliated to SHERPPA. She teaches Research Methods, Social and Economic Policy and Welfare Economics. Her research interests include amongst others the measurement and evaluation of well-being and the relation with sustainability. Her contributions to research have been published, amongst others, in Journal of Happiness Studies, Work & Stress and Social Indicators Research.

**Kristine Van Herck** is a senior researcher at the Research Institute for Work and Society (HIVA, KU Leuven) in the Research Team on "Environmental Policy and Sustainable Development". Her work focuses on sustainable development and agricultural and environmental policy. From 2007 until 2014, she worked as a researcher at LICOS Centre for Institutions and Economic Performance (KU Leuven), where she obtained in 2014 a PhD in Economics. From 2012 to 2014, she worked part time as a researcher at CEPS Centre for European Policy Studies in the research group "Food Security and Development Policy". In addition, she worked as an ad hoc consultant for several international organizations, such as the European Commission, FAO, EBRD and the World Bank.

**Tom Creten** worked as a researcher at the HIVA Research Institute for Work and Society, KU Leuven from 2013 to 2014. Being both an economist and a political scientist, he performed multidisciplinary research on the socio-economic consequences of floods and the related policy aspects. Apart from flood research, he also specialized in transition theory as a policy approach. Today he works for the Flemish government as a policy adviser at the Department of Economics, Science and Innovation.

Hydrology and
Earth System
Sciences

# Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium

**L. Foresti**[1], **M. Reyniers**[1], **A. Seed**[2], **and L. Delobbe**[1]

[1]Royal Meteorological Institute of Belgium, Brussels, Belgium
[2]Bureau of Meteorology, Melbourne, Australia

*Correspondence to:* L. Foresti (loris.foresti@gmail.com)

**Abstract.** The Short-Term Ensemble Prediction System (STEPS) is implemented in real-time at the Royal Meteorological Institute (RMI) of Belgium. The main idea behind STEPS is to quantify the forecast uncertainty by adding stochastic perturbations to the deterministic Lagrangian extrapolation of radar images. The stochastic perturbations are designed to account for the unpredictable precipitation growth and decay processes and to reproduce the dynamic scaling of precipitation fields, i.e., the observation that large-scale rainfall structures are more persistent and predictable than small-scale convective cells. This paper presents the development, adaptation and verification of the STEPS system for Belgium (STEPS-BE). STEPS-BE provides in real-time 20-member ensemble precipitation nowcasts at 1 km and 5 min resolutions up to 2 h lead time using a 4 C-band radar composite as input. In the context of the PLURISK project, STEPS forecasts were generated to be used as input in sewer system hydraulic models for nowcasting urban inundations in the cities of Ghent and Leuven. Comprehensive forecast verification was performed in order to detect systematic biases over the given urban areas and to analyze the reliability of probabilistic forecasts for a set of case studies in 2013 and 2014. The forecast biases over the cities of Leuven and Ghent were found to be small, which is encouraging for future integration of STEPS nowcasts into the hydraulic models. Probabilistic forecasts of exceeding $0.5\,\mathrm{mm\,h^{-1}}$ are reliable up to 60–90 min lead time, while the ones of exceeding $5.0\,\mathrm{mm\,h^{-1}}$ are only reliable up to 30 min. The STEPS ensembles are slightly under-dispersive and represent only 75–90 % of the forecast errors.

## 1 Introduction

The use of radar measurements for urban hydrological applications has substantially increased during the last years (e.g., Berne et al., 2004; Einfalt et al., 2004; Bruni et al., 2015). Given the fast response time of urban catchments and sewer systems, radar-based very short-term precipitation forecasting (nowcasting) has the potential to extend the lead time of hydrological and hydraulic flow predictions.

Nowcasting concerns the accurate description of the current weather situation together with very short-term forecasts obtained by extrapolating the real-time observations. Quantitative precipitation nowcasting (QPN) is traditionally done by estimating the apparent movement of radar precipitation fields using optical flow or variational echo tracking techniques and extrapolating the last observed precipitation field into the future (e.g., Germann and Zawadzki, 2002; Bowler et al., 2004a). During recent years there has been significant progress in NWP modeling with radar data assimilation techniques (see a review in Sun et al., 2014), which reduces the useful lead time of extrapolation-based nowcasts compared with NWP forecasts. The development of seamless forecasting systems that optimally blend the extrapolation nowcast with the output of NWP models makes the definition of the nowcasting time range even fuzzier (see, e.g., Pierce et al., 2010).

Due to the lack of predictability of rainfall growth and decay processes at small spatial scales (Radhakrishna et al., 2012), it is very important to provide together with a forecast an estimation of its uncertainty. The established method to represent the forecast uncertainty in Numerical Weather Prediction (NWP) is to generate an ensemble of forecasts by

perturbing the initial conditions of the model in the directions exhibiting the largest error growth, which amplify more the spread of the obtained ensemble. However, in the nowcasting range the computation of large NWP ensembles (50–100 members) that resolve features at the scales of 1 km and are updated every 5 min is still impossible to achieve. Consequently, the efforts in nowcasting research have recently focused on developing heuristic techniques for probabilistic precipitation nowcasting, which was the topic of the *Heuristic Probabilistic Forecasting Workshop* that was organized in Munich, Germany (Foresti et al., 2014).

Probabilistic QPN methods can be categorized into three main classes: analog, local Lagrangian and stochastic approaches. The analog-based approach derives the forecast probability density function (pdf) by retrieving a set of similar situations from an archive of precipitation events (Panziera et al., 2011; Foresti et al., 2015), the local Lagrangian approach derives the pdf by collecting the precipitation values in a neighborhood of a given grid point in Lagrangian frame of reference (Hohti et al., 2000; Germann and Zawadzki, 2004) and the stochastic approach exploits a random number generator to compute an ensemble of equally likely precipitation fields, for example by adding stochastic perturbations to a deterministic extrapolation nowcast (Pegram and Clothier, 2001a, b; Bowler et al., 2006; Metta et al., 2009; Berenguer et al., 2011; Seed et al., 2013; Atencia and Zawadzki, 2014; Dai et al., 2015). The stochastic approach is also extensively used to produce ensembles of precipitation fields that characterize the radar measurement uncertainty (e.g., Jordan et al., 2003; Germann et al., 2009) and for design storm studies (e.g., Willems, 2001a; Paschalis et al., 2013).

Uncertainty quantification is nowadays an integral part of both weather and hydrological forecasting (Pappenberger and Beven, 2006). Not surprisingly, an important part of hydro-meteorological research aims at understanding how to propagate the uncertainty of precipitation observations and forecasts into the hydrological models (e.g., Willems, 2001b; Cloke and Pappenberger, 2009; Collier, 2009; Zappa et al., 2010).

Several studies already analyzed the value of deterministic nowcasting systems for catchment hydrology (e.g., Berenguer et al., 2005) and for better control of urban drainage systems (e.g., Achleitner et al., 2009; Verworn et al., 2009; Thorndahl and Rasmussen, 2013). Since an important fraction of the uncertainty of hydrological predictions is due to the uncertainty of the input rainfall observations and forecasts, radar-based ensemble nowcasting systems are increasingly used as inputs for flood and sewer system modeling (e.g., Ehret et al., 2008; Silvestro and Rebora, 2012; Silvestro et al., 2013; Xuan et al., 2009, 2014). At longer forecast ranges, the NWP ensembles are also exploited for uncertainty propagation into hydrological models (see Roulin and Vannitsem, 2005; Thielen et al., 2009; Schellekens et al., 2011).

The Short-Term Ensemble Prediction System (STEPS) is a probabilistic nowcasting system developed at the Australian Bureau of Meteorology and the UK MetOffice (see the series of papers: Seed, 2003; Bowler et al., 2006; Seed et al., 2013). STEPS is operationally used at both weather services and provides short-term ensemble precipitation forecasts using both the extrapolation of radar images and the downscaled precipitation output of NWP models. The main idea behind STEPS is to represent the uncertainty due to the unpredictable precipitation growth and decay processes by adding stochastic perturbations to the deterministic extrapolation of radar images. The stochastic perturbations are designed to represent the scale-dependence of the predictability of precipitation and to reproduce the correct spatio-temporal correlation and growth of the forecast errors.

One of the first applications of STEPS in hydrology is presented in Pierce et al. (2005), who used the STEPS ensemble nowcasts to quantify the accuracy of flow predictions in a medium-sized catchment in the UK. The value of STEPS nowcasts for urban hydrology was extensively analyzed by Liguori and Rico-Ramirez (2012), Liguori et al. (2012), Liguori and Rico-Ramirez (2013) and Xuan et al. (2014). Liguori and Rico-Ramirez (2012) concluded that the performance of the radar-based extrapolation nowcast can be improved after 1 h lead time if blended with the output of a NWP model. They also found that, according to the Receiver Operating Characteristic (ROC) curve, the probabilistic nowcasts have more discrimination power than the deterministic ones. Liguori et al. (2012) integrated STEPS nowcasts as inputs into sewer system hydraulic models in an urban catchment in Yorkshire (UK). They concluded that the blending of radar and NWP forecasts has the potential to increase the lead time of flow predictions, but is strongly limited by the low accuracy of the NWP model in forecasting small-scale features. Liguori and Rico-Ramirez (2013) performed a detailed verification of the accuracy of flow predictions and concluded that the STEPS ensembles provide a similar performance as using a deterministic STEPS control forecast, but the ensembles lead to a slight underestimation of the flow predictions. Xuan et al. (2014) used ensemble STEPS nowcasts as inputs in a lumped hydrological model for a medium-sized catchment in the southwest of the UK. The hydrological model calibrated with rain gauges had lower RMSE than the one using radar data, but the ability of STEPS to account for the forecast uncertainty was useful in capturing some of the high flow peaks and extending the forecast lead time. However, the conclusions of the previous studies are strongly affected by the limited number of flood events analyzed. An extensive review of the usage of precipitation forecast systems for operational hydrological predictions in the UK from very short to long ranges (including STEPS) is provided in Lewis et al. (2015).

The goal of this paper is to present the development and verification of the STEPS system at the Royal Meteorological Institute of Belgium (RMI), referred to as STEPS-BE.

STEPS-BE provides in real-time 20-member ensemble precipitation nowcasts at 1 km and 5 min resolutions up to 2 h lead time on a 512 × 512 km domain using the Belgian 4 C-band radar composite as input. It was developed in the framework of the Belspo project PLURISK for better management of rainfall-induced risks in the urban environment. With respect to the original implementation of STEPS (Bowler et al., 2006), STEPS-BE includes two main improvements, which are designed to generate better STEPS nowcasts without NWP blending. The first one is related to the optical flow algorithm, which is extended with a kernel-based interpolation method to obtain smoother velocity fields. The second one concerns the generation of stochastic noise only within the advected radar composite. While the verification of STEPS nowcasts with NWP blending has already been extensive (Bowler et al., 2006; Seed et al., 2013), this paper analyzes the accuracy of STEPS ensemble nowcasts without NWP blending in the 0–2 h forecasting range.

Ensemble STEPS nowcasts are computed for a set of sewer overflow cases that affected the cities of Leuven and Ghent in 2013 and 2014. The accuracy of the ensemble mean forecast is verified using both continuous verification scores (multiplicative bias, RMSE) and categorical scores derived from the contingency table (probability of detection, false alarm ratio and Gilbert skill score). However, the most interesting part of this paper is the probabilistic and ensemble verification of STEPS nowcasts using both stratiform and convective rainfall events. Probabilistic nowcasts are verified using reliability diagrams and ROC curves. On the other hand, the dispersion of the nowcast ensembles is verified using rank histograms and by comparing the ensemble spread to the error of the ensemble mean.

The paper is structured as follows. Section 2 presents the radar data processing and case studies that are used to generate and verify the STEPS forecasts. Section 3 describes the STEPS nowcasting system and its extension and local implementation for Belgium (STEPS-BE). Section 4 illustrates the forecast verification results. Section 5 concludes the paper and discusses future perspectives.

## 2 Radar data and precipitation case studies

STEPS-BE integrates as input a composite image produced from the C-band radars of Wideumont (RMI, single-pol), Zaventem (Belgocontrol, single-pol), Jabbeke (RMI, dual-pol) and Avesnois (Meteo-France, dual-pol). The composite is produced on a 500 m resolution grid by combining single-radar pseudo Constant Altitude Plan Position Indicators (CAPPI) at a height of 1500 m a.s.l. The compositing algorithm takes the maximum reflectivity value from each radar at each grid point.

The radars have different hardware and scanning strategies, and are operated by different agencies (RMI, Belgocontrol and Meteo-France), which inevitably leads to differ-

ences in the data processing. The Wideumont and Zaventem radars eliminate the non-meteorological echoes using standard Doppler filtering. The Jabbeke radar includes an additional clutter filtering that uses a fuzzy logic algorithm based on the dual-polarization moments (essentially the co-polar correlation coefficient, the texture of the differential reflectivity and the texture of the specific differential phase shift). A static ground clutter map and a statistical filter are used by Meteo-France to remove the non-meteorological echoes of the Avensois radar. The French radar data processing chain is described in Tabary (2007) and in Figueras i Ventura and Tabary (2013).

Since the Zaventem radar is mainly used for aviation applications, its scanning strategy is optimized for the measurement of winds. Except for the lowest elevation scan, a dual PRF mode (1200/800 Hz) is used. The azimuths that are scanned with a high PRF (1200 Hz) only have a maximum range of 125 km and are more affected by the second trip echoes caused by convective cells located beyond the 125 km range.

All radars use the standard Marshall–Palmer relationship $Z = 200R^{1.6}$ to convert the measured reflectivity to rainfall rate. A composite image with more advanced radar-based quantitative precipitation estimation (QPE), that includes better ground clutter removal algorithms and also a correction for the bright band, was recently developed and the verification of the new product is ongoing.

STEPS forecasts were generated and verified for a set of sewer system overflow cases that affected the cities of Ghent and Leuven (see Table 1). The Ghent cases have a more stratiform character and occurred in late autumn and winter. On the other hand, the Leuven cases are more convective and occurred in summer months. A detailed climatology of convective storms in Belgium can be found in Goudenhoofdt and Delobbe (2009).

## 3 Short-Term Ensemble Prediction System (STEPS)

### 3.1 STEPS description

The Short-Term Ensemble Prediction System (STEPS) was jointly developed by the Australian Bureau of Meteorology (BOM) and the UK MetOffice (Bowler et al., 2006). STEPS forecasts are produced operationally at both weather services and are distributed to weather forecasters and a number of external users, in particular the hydrological services.

The key idea behind STEPS is to account for the unpredictable rainfall growth and decay processes by adding stochastic perturbations to the deterministic extrapolation of radar images (Seed, 2003). In order to be effective, the stochastic perturbations need to reproduce important statistical properties of both the precipitation fields and the forecast errors:

**Table 1.** List of precipitation events that caused sewer system floods in Ghent and Leuven.

| City | Date | Event start (UTC) | Event end (UTC) | Duration | Predominant precipitation | Main wind direction |
|------|------|-------------------|-----------------|----------|---------------------------|---------------------|
| Ghent | 10 Nov 2013 | 13:50 | 22:00 | 8:10 h | Stratiform | WNW → NNW |
| Ghent | 3 Jan 2014 | 03:00 | 14:00 | 11 h | Stratiform | SW → WSW |
| Leuven | 9–10 June 2014 | 06:30, 9th | 15:30, 10th | 33 h | Convective | SW |
| Leuven | 19–20 July 2014 | 22:00, 19th | 06:30, 20th | 8:30 h | Convective | SSW |

1. spatial scaling of precipitation fields,

2. dynamic scaling of precipitation fields,

3. spatial correlation of the forecast errors,

4. temporal correlation of the forecast errors.

The *spatial scaling* considers the precipitation field as arising from multiplicative cascade processes (Schertzer and Lovejoy, 1987; Seed, 2003). The presence of spatial scaling can be demonstrated by computing the 2-D Fourier power spectrum (PS) of a precipitation field. A 1-D PS can be obtained by radially averaging the 2-D PS. The precipitation field is said to be *scaling* if the 1-D PS draws a straight line on the log-log plot of the power against the spatial frequency (power law), which can be parametrized by one or two spectral exponents (see, e.g., Seed et al., 2013; Foresti and Seed, 2014). Within the multiplicative framework, a rainfall field is not represented as a collection of convective cells of a characteristic size but rather as a hierarchy of precipitation structures embedded in each other over a continuum of scales. STEPS considers the spatial scaling by decomposing the radar rainfall field into a multiplicative cascade using a fast Fourier transform (FFT) to isolate a set of eight spatial frequencies (Seed, 2003; Bowler et al., 2006; Seed et al., 2013). The top cascade levels (0, 1 and 2) represent the low spatial frequencies (large precipitation structures), while the bottom cascade levels (5, 6, 7) represent the high spatial frequencies (small precipitation structures). Another important behavior of rainfall fields is known as *dynamic scaling*, which is the empirical observation that the rate of temporal development of rainfall structures is a power law function of their spatial scale (Venugopal et al., 1999; Foresti and Seed, 2014). This means that large precipitation features are more persistent and hence predictable compared with small precipitation cells, which is closely related to the concept of scale-dependence of the predictability of precipitation (Germann and Zawadzki, 2002; Turner et al., 2004).

The stochastic perturbations should be able to reflect the properties of the forecast errors. Generating spatially and temporally correlated forecast errors is mandatory for hydrological applications, in particular when the correlation length of the errors is comparable or superior to the size and response time of the catchment. *Spatially correlated stochastic noise* can be constructed by applying a power law filter to a white noise field (Schertzer and Lovejoy, 1987). In practice it consists of three steps: computing the FFT of a white noise field, multiplying the obtained components in frequency domain by a given filter and applying the inverse FFT to return back to the spatial domain. The 1-D or 2-D power spectra of the rainfall field can be used as a filter to obtain noise fields that have the same scaling and spatial correlation of the rainfall field. The 1-D PS of the precipitation fields often appears to be a power law of the spatial frequency and explains why the procedure is also called power law filtering of white noise. In order to represent the anisotropies of the precipitation field, the 2-D PS can also be used as a filter. In the absence of a target precipitation field from which to derive the PS, the filter can be parametrized by using a climatological power law (see Seed et al., 2013). Finally, the *temporal correlations* are imposed by auto-regressive (AR) filtering. A hierarchy of AR processes defines the temporal evolution of the cascade levels. With the exception of forecast lead times beyond 2–3 h (Atencia and Zawadzki, 2014), an AR process of order 1 or 2 is already a good approximation to describe the temporal decorrelation of the forecast errors.

The practical implementation of STEPS to reproduce these important properties consists of the following steps (see Bowler et al., 2006; Foresti and Seed, 2014).

1. Estimation of the velocity field using optical flow on the last two radar rainfall images (Bowler et al., 2004a)

2. Decomposition of both rainfall fields into a multiplicative cascade using an FFT to isolate a set of eight spatial frequencies

3. Estimation of the rate of temporal evolution of rainfall features at each level of the cascade (Lagrangian auto-correlation)

4. Generation of a cascade of spatially correlated stochastic noise using as a filter the 1-D or 2-D power spectra of the last observed radar rainfall field. A Gaussian filter is used to isolate a given spatial frequency (see Foresti and Seed, 2014).

5. Stochastic perturbation of the rainfall cascade using the noise cascade (level by level)

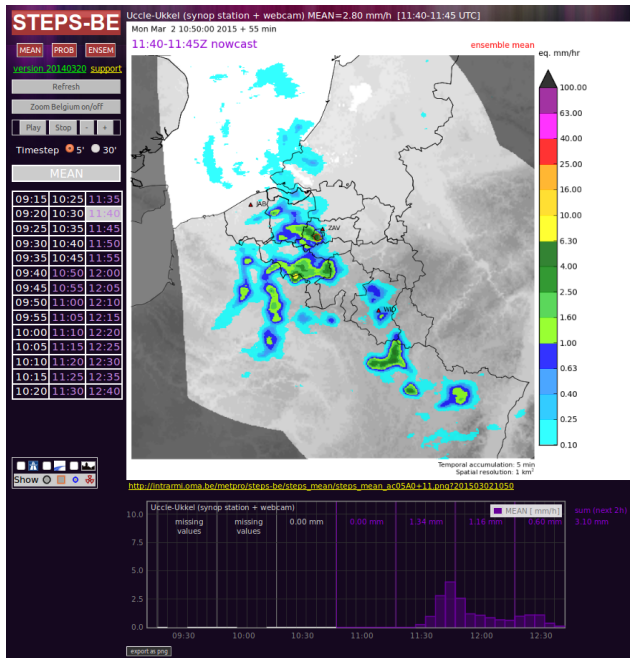6. Extrapolation of the cascade levels using a semi-Lagrangian advection scheme

**Figure 1.** Web platform of STEPS-BE showing the ensemble mean forecast.

7. Application of the AR(1) or AR(2) model for the temporal update of the cascade levels at each forecast lead time using the Lagrangian auto-correlations estimated in step (3)

8. Recomposition of the cascade into a rainfall field

9. Probability matching of the forecast rainfall field with the original observed field (Ebert, 2001)

10. Computation of the forecast rainfall accumulations from the instant forecast rainfall rates. This procedure is known as advection correction and consists of advecting the instant rainfall rate forward over the 5 min period by discretizing the advection into smaller time steps.

### 3.2 STEPS implementation at RMI (STEPS-BE)

Bowler et al. (2006) introduced a general framework for blending a radar-based extrapolation nowcast with one or more outputs of downscaled NWP models (see also Pierce et al., 2010; Seed et al., 2013). Because of being designed for urban applications, the maximum lead time of STEPS-BE is restricted to 2 h. The operational NWP model of RMI (ALARO) runs only four times daily using a grid spacing of 4 km. Considering the model spin-up time and the absence of radar data assimilation, it is very unlikely that ALARO provides useful skill for blending its output with a radar-based extrapolation nowcast within the considered nowcasting range. It must also be remembered that the effective res-

olution of NWP models is much larger than the grid spacing. For instance, Grasso (2000) estimates the effective resolution to be at least 4 times the grid spacing, while Skamarock (2004) estimates it to be up to 7 times the grid spacing. ALARO would then only be able to resolve features that are greater than 20 km. For all these reasons, STEPS-BE only involves an extrapolation nowcast without NWP blending.

The STEPS-BE forecast domain is smaller than the extent of the 4 C-band radars composite (see Fig. 1). The radar field was upscaled from the original resolution of 500 m to 1 km and a sub-region of $512 \times 512$ grid points centered over Belgium was extracted. The forecast domain was extended by 32 pixels on each side to reduce the edge effects due to the FFT. This leads to an eight-level multiplicative cascade representing the following spatial scales (rounded to the nearest integer): $576$–$256$–$114$, $256$–$114$–$51$, $114$–$51$–$23$, $51$–$23$–$11$, $23$–$11$–$4$, $11$–$4$–$2$, $4$–$2$–$1$ and $2$–$1$ km. Italic characters mark the scales on which the Gaussian filter is centered in the frequency domain (see Foresti and Seed, 2014, for a more detailed explanation and visualization of the Gaussian FFT filter). The Gaussian filters of the largest and smallest spatial scales are truncated in order to preserve the power of the field. The top cascade level represents scales between the 576 and 256 km wavelengths and is not a perfect Gaussian filter. One can notice that the spatial scales are not exact multiples of 2. In fact, a multiplication factor of 2.246 was employed to match the enlarged STEPS-BE domain size.

STEPS-BE includes a couple of improvements compared with the original implementation of the BOM:

1. kernel interpolation of optical flow vectors;

2. generation of stochastic noise only within the advected radar mask.

The optical flow algorithm of Bowler et al. (2004a) estimates the velocity field by dividing the radar domain into a series of blocks within which the optical flow equation is solved. The minimization of the field divergence is only performed at the level of the block, which leaves sharp discontinuities in the velocity field between the blocks. In order to overcome this issue, a Gaussian kernel smoothing was applied to interpolate the velocity vectors located at the center of the blocks onto the fine radar grid. The bandwidth of the Gaussian kernel was chosen to be $\sigma = 24\,\text{km} = 0.4k$, where $k = 60$ grid points is the block size. This setting has the advantage of obtaining velocity fields that are less affected by the differential motion of small rainfall features and the presence of ground clutter. A too precise velocity field would provide increased predictability at very short lead times but worse forecasts at longer lead times due to excessive convergence and divergence of precipitation features during the advection. Smooth velocity fields could also be obtained by using a smaller block size and by compensating with a larger bandwidth of the smoothing kernel.

In STEPS-BE the 1-D power spectrum of the last observed rainfall field is used as a filter to generate the spatially correlated stochastic perturbations. The PS is parameterized using two spectral slopes to account for a scaling break that is often observed at the wavelength of 40 km (see Seed et al., 2013; Foresti and Seed, 2014). To simplify the computations, an auto-regressive model of order 1 (AR(1)) was employed for imposing the temporal correlations and to model the growth of forecast errors.

The original STEPS implementation (Bowler et al., 2006) was designed to blend the radar extrapolation nowcasts with the output of NWP models. However, the domain covered by the radars is smaller than the rectangular domain of the NWP model and small amounts of stochastic noise are generated by default also outside the radar composite. This setting was not adapted for radar-based nowcasts without NWP blending and needed some adaptation. In fact, when advecting the radar mask over several time steps, large areas with small amounts of stochastic rain appear outside the validity domain of the forecast and perturb the probability matching. In STEPS-BE the stochastic perturbations are only generated within the advected radar domain and set to zero elsewhere.

STEPS-BE can also account for the uncertainty in the estimation of the velocity field. The STEPS version that is implemented in the UK (Bowler et al., 2006) includes a detailed procedure to generate velocity perturbations that reproduce various statistical properties of the differences between the forecast velocity and the actual future diagnosed velocity (see details in Bowler et al., 2004b). In the BOM and RMI implementations a simpler procedure is applied. The diagnosed velocity field is multiplied by a single factor $C$ that is drawn from the following distribution:

$$C = 10^{1.5N/10}, \tag{1}$$

where $N$ is a normally distributed random variable with zero mean and unit variance. In other words, the velocity field is accelerated or decelerated by a single random factor without affecting the direction of the vectors. In fact, the uncertainty on the diagnosed speed was observed to be higher than that of the direction of movement (Bowler et al., 2006).

The BOM and RMI versions of STEPS also include a stochastic model for the radar measurement error, a broken-line model to account for the unknown future evolution of the mean areal rainfall and the possibility to use time-lagged ensembles. However, a nowcasting model with too many components is harder to calibrate and complicates the interpretation of the forecast fields. Because of these reasons, STEPS-BE only exploits the basic stochastic model for the velocity field and for the evolution of rainfall fields (due to growth and decay processes).

The core of STEPS and the STEPS-BE extensions are implemented in C/C$^{++}$ and the production of figures in python. Bash scripts were written to call multiple STEPS instances and compute the ensemble members in parallel over several processors. Once all the ensemble members are computed, a separate script collects the corresponding netCDF files and calculates the forecast probabilities. Most of the computational cost of STEPS consists of filtering the white noise field with FFT, advecting and updating the radar cascade with the AR model. The re-calculation of optical flow fields on each processor takes less than 10 % of the total computational time.

The python matplotlib library is used to read the netCDF files, export the PNG figures and the time series of observed and forecast rainfall at the location of major cities and weather stations. A single STEPS nowcast generates more than 600 figures, which takes a significant fraction of the total computational time. In order to optimize the timing, a bash script monitors continuously the directory with incoming radar composites and triggers STEPS-BE once a field with a new time stamp is found. All these implementation details ensure that the user/forecaster can have access to an ensemble STEPS nowcast in less than 5 min after receiving the radar composite image.

The visualization system of STEPS-BE is very similar to the one of INCA-BE, the local Belgian implementation of the Integrated Nowcasting through Comprehensive Analysis system (INCA, Haiden et al., 2011) developed at the Austrian weather office (ZAMG). Figure 1 illustrates the web interface with an example of an ensemble mean nowcast. The user can highlight the major cities, weather stations and click to visualize the time series of observed and forecast precipitation/probability, which appears at the bottom of the web page. The navigation through the observations and forecast lead times is facilitated by the scroll wheel of the mouse. On the other hand, by clicking on the image it is possible to easily scroll through the various ensemble members or probability levels for a given lead time. Scrolling the ensemble members at different lead times is very instructive and can make the user aware of the forecast uncertainty. In fact, at a lead time of 5 min the ensemble members agree very well on the intensity and location of precipitation. This means that the ensemble spread is small and the probabilistic forecast is sharp; i.e., most of the forecast probabilities are close to 1 or 0 (see an explanation in Appendix A). On the other hand, at 1 or 2 h lead time the ensemble members disagree on the location and intensity of rainfall, which enhances the ensemble spread and decreases the sharpness of the probabilistic forecast. The web page includes extensive documentation to guide the user and a set of case studies to help understanding the strengths and limitations of STEPS. The visualization system was implemented with great attention to take full advantage of the multi-dimensional information content of probabilistic and ensemble forecasts.
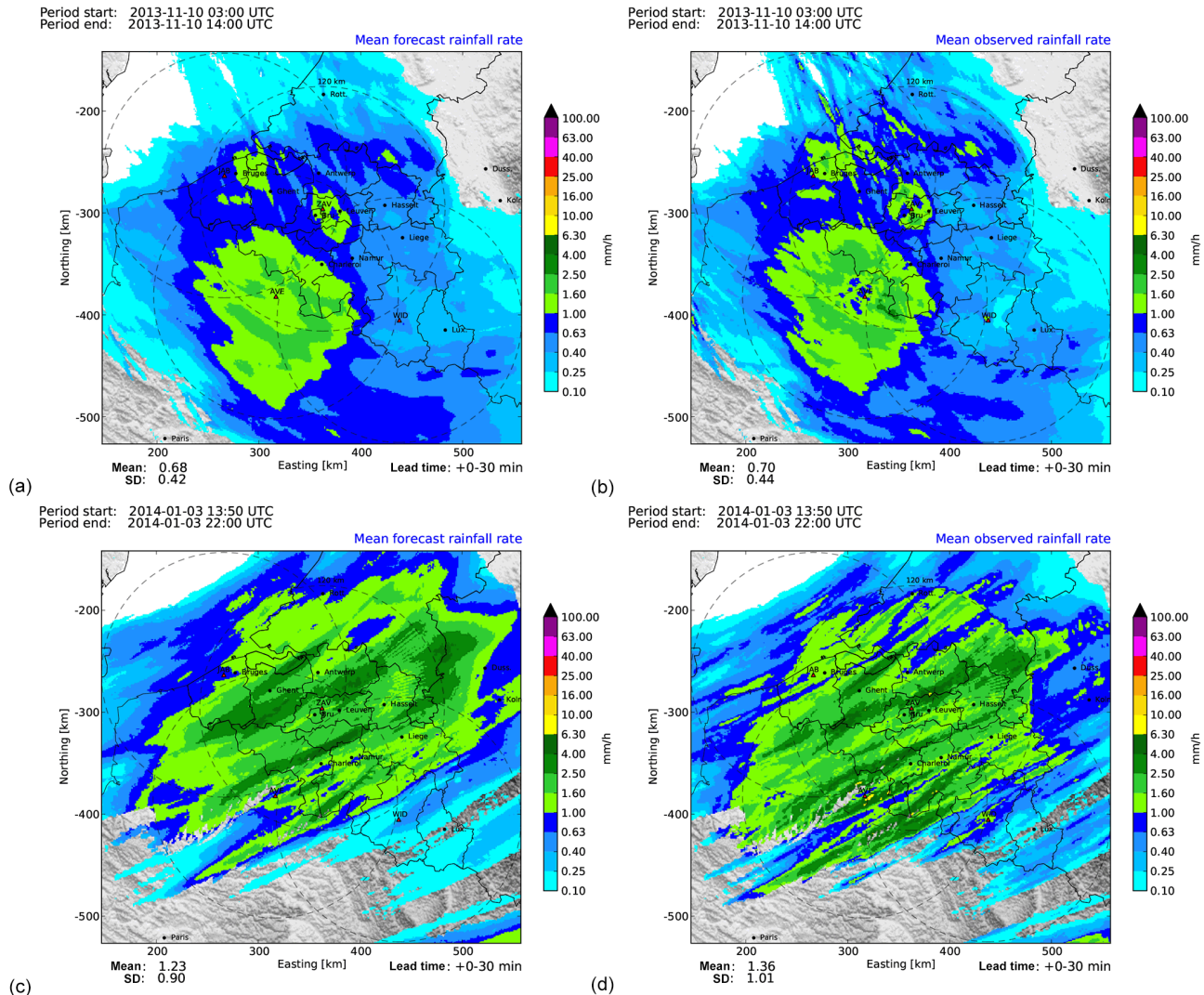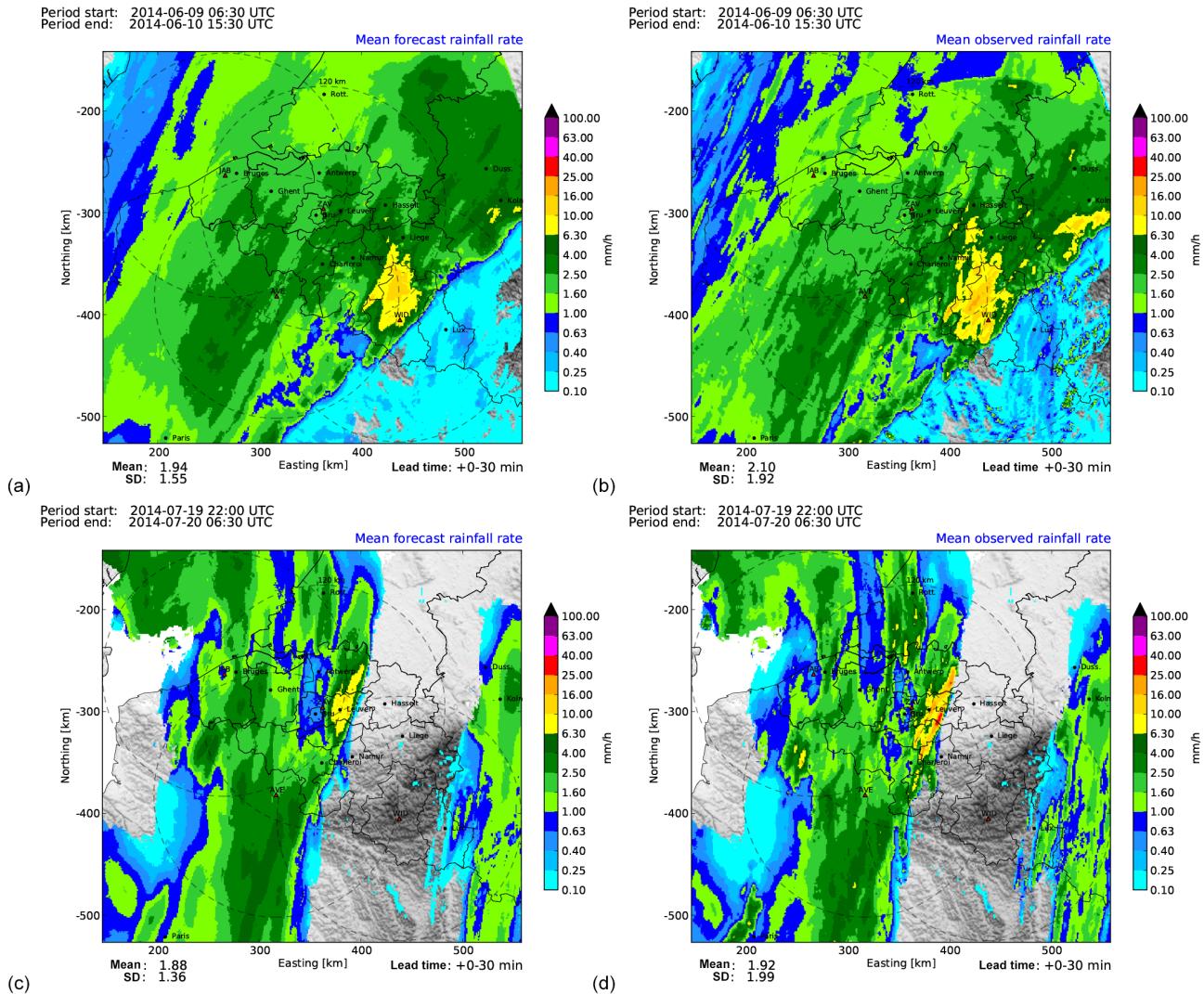
**Figure 2.** Average forecast and observed rainfall accumulations for the Ghent cases. **(a)** Forecast and **(b)** observed 0–30 min rainfall accumulations on 10 November 2013. **(c)** Forecast and **(d)** observed 0–30 min rainfall accumulations on 3 January 2014. The mean and standard deviation of the field within the 120 km range of the radars are shown on the bottom left corner. Field values are shown only if there are at least 10 samples for the computation of the mean. The red triangles denote the location of the Wideumont (WID, coordinates 438 km east/−405 km north), Zaventem (ZAV, 363/−296), Jabbeke (JAB, 266/−263) and Avesnois (AVE, 317/−382) radars. The 120 km range from the radar is displayed as a dashed circle. The mountain range of the Ardennes covers the three most southern districts of Belgium and Luxembourg (LUX).

## 4   Forecast verification

### 4.1   Verification set-up

This section presents the verification of STEPS-BE forecasts using a set of case studies (see Sect. 2). The accumulated radar observations were employed as reference for the verification. The rainfall rates are accumulated over the last 5 min by reversing the field vectors based on the observations and then performing the advection correction. The 30 min ensemble mean forecast was verified against the observed 30 min radar accumulations using both continuous and categorical

verification scores. The deterministic verification procedure follows the one presented in Foresti and Seed (2015), which was designed to analyze the spatial distribution of the forecast errors. More details about the forecast verification setup and scores are given in Appendix A.

The continuous scores include the multiplicative bias and the root mean squared error (RMSE), while the categorical scores include the probability of detection (POD), false alarm ratio (FAR) and Gilbert skill score (GSS) derived from the contingency table for rainfall thresholds of 0.5 and 5.0 mm h$^{-1}$. The rainfall thresholds are given in equivalent intensity independently of the forecast rainfall accumulation.

**Figure 3.** Average observed and forecast rainfall accumulations for the Leuven cases. **(a)** Forecast and **(b)** observed 0–30 min rainfall accumulations on 9–10 June 2014. **(c)** Forecast and **(d)** observed 0–30 min rainfall accumulations on 19–20 July 2014.

Thus, a threshold of $5.0\,\mathrm{mm\,h^{-1}}$ on a 30 min accumulation corresponds to 2.5 mm of rain. The multiplicative bias and the RMSE were evaluated only at the locations where the forecast or the verifying observations exceeded $0.1\,\mathrm{mm\,h^{-1}}$, which can be referred to as a *weakly conditional verification*. The probabilistic forecast of exceeding 0.1, 0.5 and $5.0\,\mathrm{mm\,h^{-1}}$ was verified using the reliability diagrams and ROC curves. Finally, the dispersion of the ensemble was analyzed by comparing the ensemble spread to the RMSE of the ensemble mean and by using rank histograms. The probabilistic and ensemble verification does not consider the spatial distribution of the errors and pools the data together in both space and time to derive the statistics.

## 4.2 Deterministic verification

Figures 2 and 3 show the average forecast and observed rainfall rates corresponding to the 0–30 min ensemble mean accumulation nowcast for the Ghent and Leuven cases, respectively. In other words, they represent the average forecast and observed rainfall rates over the duration of the precipitation event (for the 0–30 min lead time). The average was computed using the weak conditional principle explained above.

The average forecast and observed accumulations generally agree very well for the 0–30 min lead time forecast. The Ghent case on 10 November 2013 (Fig. 2a and b) is the only one with northwesterly flows and is characterized by the lowest average rainfall rates. The Avesnois radar demonstrates very well the range dependence of the average rainfall rates, which gradually decrease with increasing distance from the
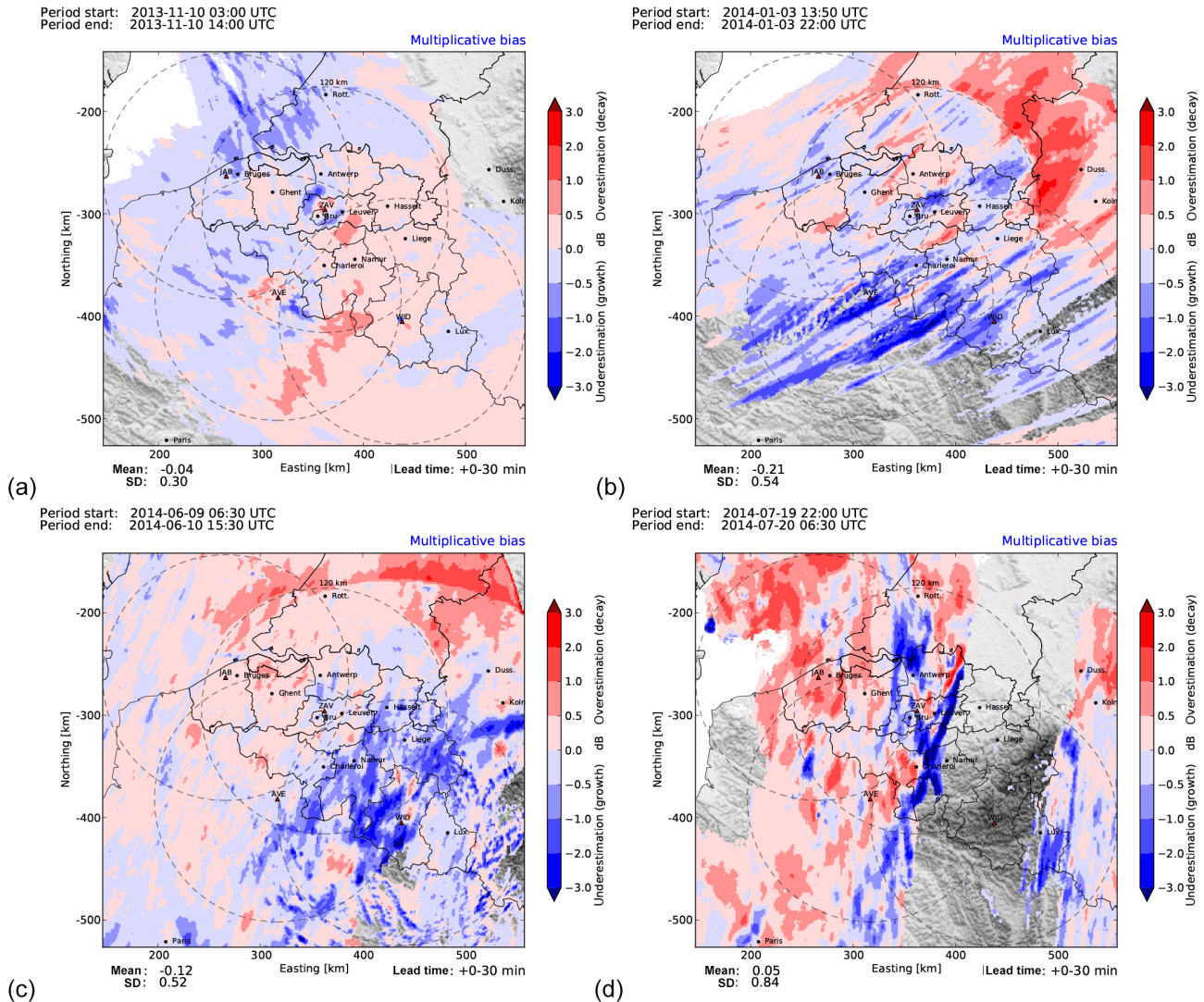
**Figure 4.** Average 0–30 min multiplicative forecast biases for the Ghent cases on **(a)** 10 November 2013 and on **(b)** 3 January 2014 and the Leuven cases on **(c)** 9–10 June 2014 and on **(d)** 19–20 July 2014. The interpretation of under- and over-estimations by STEPS as systematic rainfall growth and decay or simply as radar measurement biases is subject to interpretation as explained in the text.

radar. On the contrary, the smaller ring of high rainfall rates around the Zaventem radar is mostly due to the bright band (Fig. 2b).

The bright band effect influences the radar observations and hence the nowcasts based on their extrapolation. At longer lead times the larger rainfall estimates due to the bright band are extrapolated far from the location of the radar. The stochastic perturbations of STEPS can help to gradually dissolve the circular patterns introduced by the bright band effect. However, the bright band affects more the observations used for the verification, in particular when the rainfall is advected from upstream over the radar region. In such a case, the local larger rainfall estimates lead to a verification bias and the forecasts are wrongly supposed of rainfall underestimation. In spite of these issues, bright band effects might

not be so important for urban hydrological applications. In fact, except for one stratiform case presented in this paper, pluvial floods mainly happen in summer with convective precipitation events, during which the bright band is absent or negligible.

The Ghent case on 3 January 2014 has higher rainfall rates and the elongated structures of precipitation areas demonstrate well the southwesterly flow regime (Fig. 2c and d). For this case the measurements of the Zaventem radar are also affected by second trip echoes, which appear as a set of radially oriented rainfall structures northwest of the radar. These alternating patterns are explained by the dual PRF mode of Zaventem (see Sect. 2).

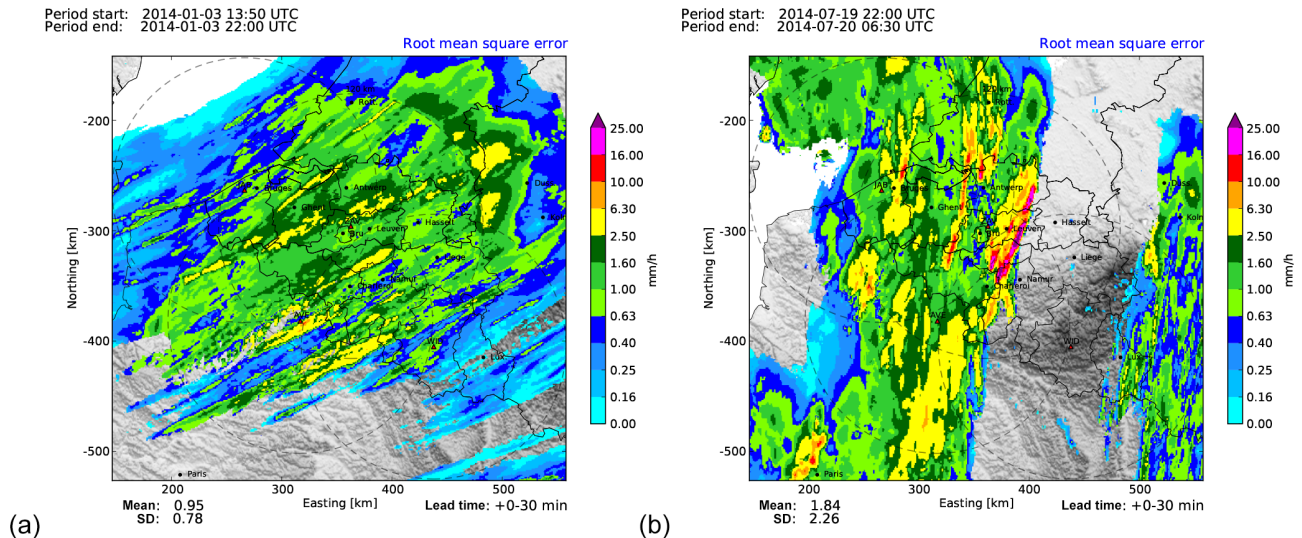The Leuven cases on 9 June and 19–20 July 2014 have an important convective activity (Fig. 3a–d). The maximum av-

**Figure 5.** Average 0–30 min forecast RMSE for **(a)** the Ghent winter case on 3 January 2014 and **(b)** the Leuven summer case on 19–20 July 2014.

erage rainfall rates are located over the Ardennes mountain range and the city of Leuven, respectively. Since urban flash floods can be triggered by a single convective cell, the average rainfall rate over the duration of the event may not be as high in the considered city (e.g., Fig. 3b).

Figure 4 illustrates the multiplicative bias of the 0–30 min nowcast averaged over each of the four events. A detailed interpretation of such forecast biases using Australian radar data and their connection to orographic features is given in Foresti and Seed (2015), which point out that an important fraction of the forecast errors is caused by the biases of the verifying radar observations rather than systematic rainfall growth and decay processes due to orography. In Fig. 4a it is easy to notice the effect of bright band, which causes a series of systematic forecast biases around the Zaventem radar and perpendicularly oriented with respect to the prevailing flow direction (NW). Systematic rainfall underestimation occurs along the Belgian coast of the North Sea. One factor which contributes to this underestimation is the absence of visibility of the radar at longer ranges. The incoming precipitation is suddenly detected by the radar and therefore strongly underestimated by STEPS. The only situation where the range dependence of the rainfall estimation does not affect the forecast verification occurs when the velocity field is perfectly rotational and centered on the radar (assuming no beam blockage). All the other cases have to deal with the fact that the rainfall nowcast also extrapolates the biases of the radar observations! Contrary to expectation, on the upwind side of the Ardennes there is overestimation, which may depict a region of systematic rainfall decay. The bias over the city of Ghent is fortunately small and is included in the range from 0 to +0.5 dB (light overestimation, rainfall decay). Having small systematic biases over the cities of interest is very im-

portant for future integration of STEPS nowcasts as input in hydraulic models.

In Fig. 4b the systematic underestimation is also located upstream with respect to the prevailing winds (SW). The strong overestimations in Germany and the Netherlands are mostly due to the underestimation of rainfall by the verifying radar observations rather than caused by systematic rainfall decay. This is particularly visible after a range of 125 km from the Zaventem radar, which demonstrates again that discontinuities and biases in the radar observations lead to biases in the extrapolation forecast. Also in this case the bias over the city of Ghent is small but in the range from −0.5 to 0 dB (light underestimation, rainfall growth). Radar composite discontinuities are also visible in Fig. 4c but this time located at a range of 240 km north of the Wideumont radar when entering the area covered by the Jabbeke radar. This forecast bias is mainly explained by the negative calibration bias of the Jabbeke radar, which is known to slightly underestimate the rainfall rates with respect to the Wideumont radar. Strong underestimation occurs over the Ardennes due to the systematic initiation and growth of convection that cannot be predicted by STEPS (Fig. 4c). Fortunately the city of Leuven is located in a region with small biases in the range from −0.5 to +0.5 dB. Figure 4d is quite interesting since strong underestimations are located in front of the rain band (from Charleroi to Leuven and beyond) and overestimations at the rear of the rain band (west of the Jabbeke radar). The underestimations are due to systematic rainfall initiation in front of the rain band, while the overestimations are probably caused by a too slow extrapolation of rainfall, which tends to drag at the rear of the rain band. The two bands of underestimations south of Leuven are caused by two different thunderstorms. The first one passed over the city of Leuven and had
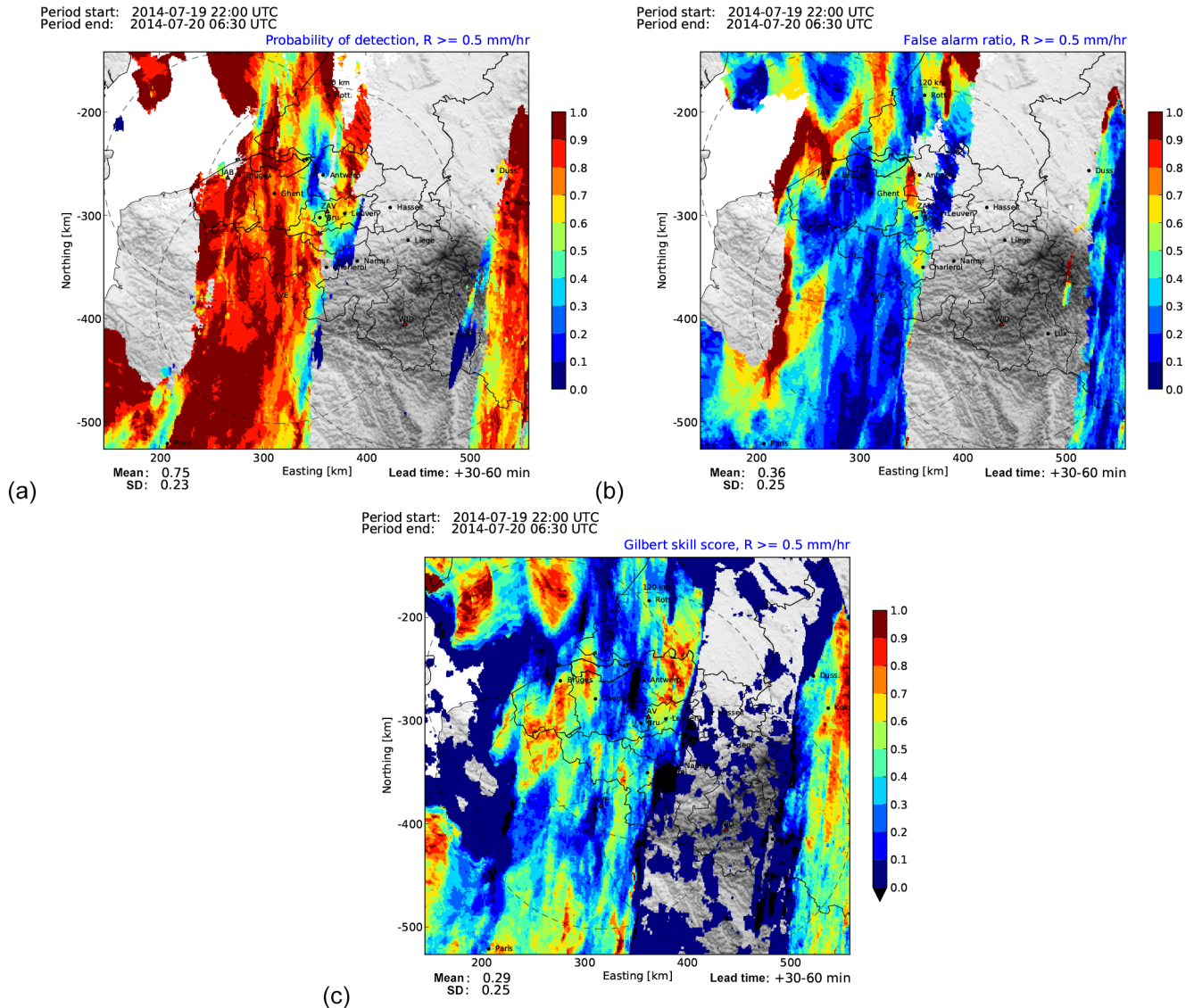
(a)



(b)



(c)

**Figure 6. (a)** POD, **(b)** FAR and **(c)** GSS of the 30–60 min ensemble mean forecast of exceeding $0.5\,\mathrm{mm\,h^{-1}}$ for the Leuven case on 19–20 July 2014.
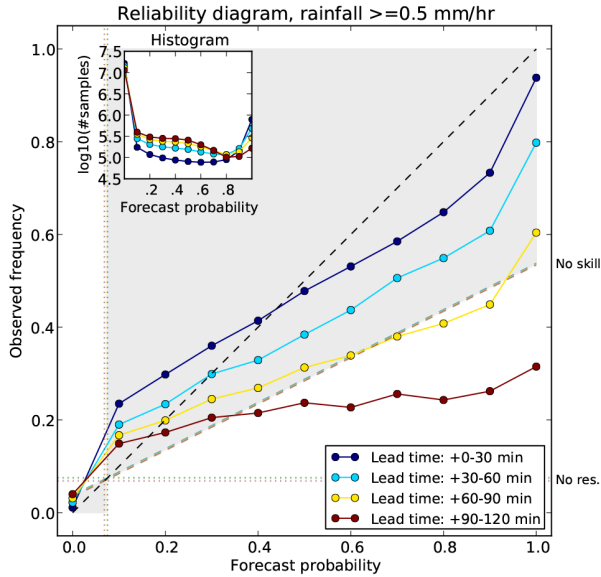
a stronger westerly component with respect to the prevailing southerly flow. The second thunderstorm was weaker and had a stronger easterly component. When isolated convection does not follow the prevailing movement of the rainfall field, strong biases can appear in the nowcast during the first lead times.

Figure 5 shows the spatial distribution of the RMSE for the stratiform event on 3 January 2014 in Ghent and the convective event on 19–20 July 2014 in Leuven. If compared with Figs. 2d and 3d it is clear that the RMSE is strongly correlated with the regions having the highest mean rainfall accumulations (proportional effect). Thus, it is not surprising that the RMSE of the convective case (Fig. 5b) displays values exceeding $10\,\mathrm{mm\,h^{-1}}$ over the city of Leuven. The winter

case only shows RMSE values below $2\,\mathrm{mm\,h^{-1}}$ over the city of Ghent.

Figure 6 illustrates an example of categorical verification of the 30–60 min ensemble mean forecast for the Leuven case on 19–20 July 2014 relative to the rainfall threshold of $0.5\,\mathrm{mm\,h^{-1}}$. The probability of detection is high everywhere (mean of 0.75) except in the neighborhood of Antwerp and south of Leuven, where the initiation of thunderstorms could not be predicted by STEPS (Fig. 6a). The false alarm ratio is quite low (mean of 0.36) and the regions with high values are mainly located at the rear of the front where the rainfall is advected too slowly compared with the actual movement of the front (Fig. 6b). A high Gilbert skill score generally coincides with the regions with the highest rainfall accumulations and
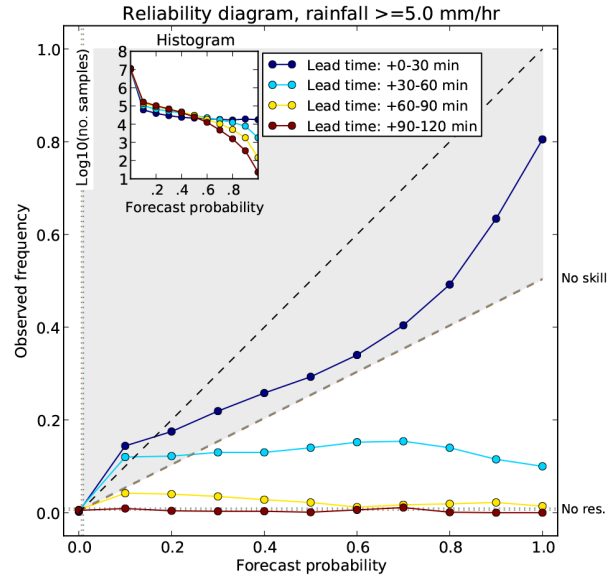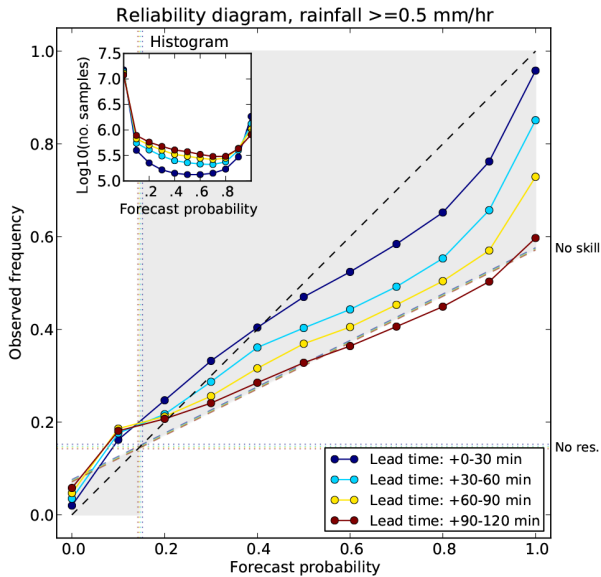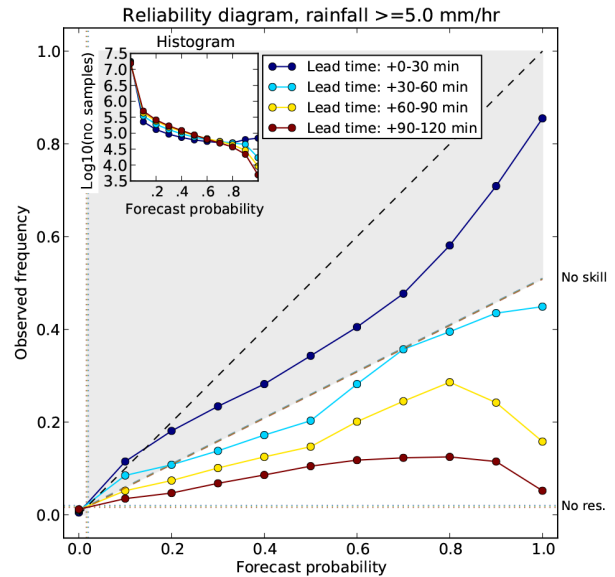
**Figure 7.** Reliability diagrams for the Ghent case on 3 January 2014 relative to the probabilistic forecast of exceeding **(a)** 0.5 mm h$^{-1}$ and **(b)** 5.0 mm h$^{-1}$. **(c, d)** Same as **(a, b)** but for the Leuven case on 19–20 July 2014.

becomes lower at the edges of the rain areas (Fig. 6c). This finding can be explained conceptually if one thinks about the verification of the future path of a single convective cell. The regions with the highest uncertainty are located along the edges of the predicted thunderstorm path and the highest skill is obtained in the center of the predicted path.

## 4.3 Probabilistic verification

Figure 7 shows the reliability diagrams relative to the probabilistic forecast of exceeding the 0.5 and 5.0 mm h$^{-1}$ rainfall thresholds for the Ghent case on 3 January 2014 (Fig. 7a and b) and the Leuven case on 19–20 July 2014 (Fig. 7c and d). The reference probabilistic forecast is taken as the climatological frequency of exceeding a given rainfall threshold during that precipitation event (horizontal dashed line). Unexpectedly, the forecasts of the stratiform case in Ghent
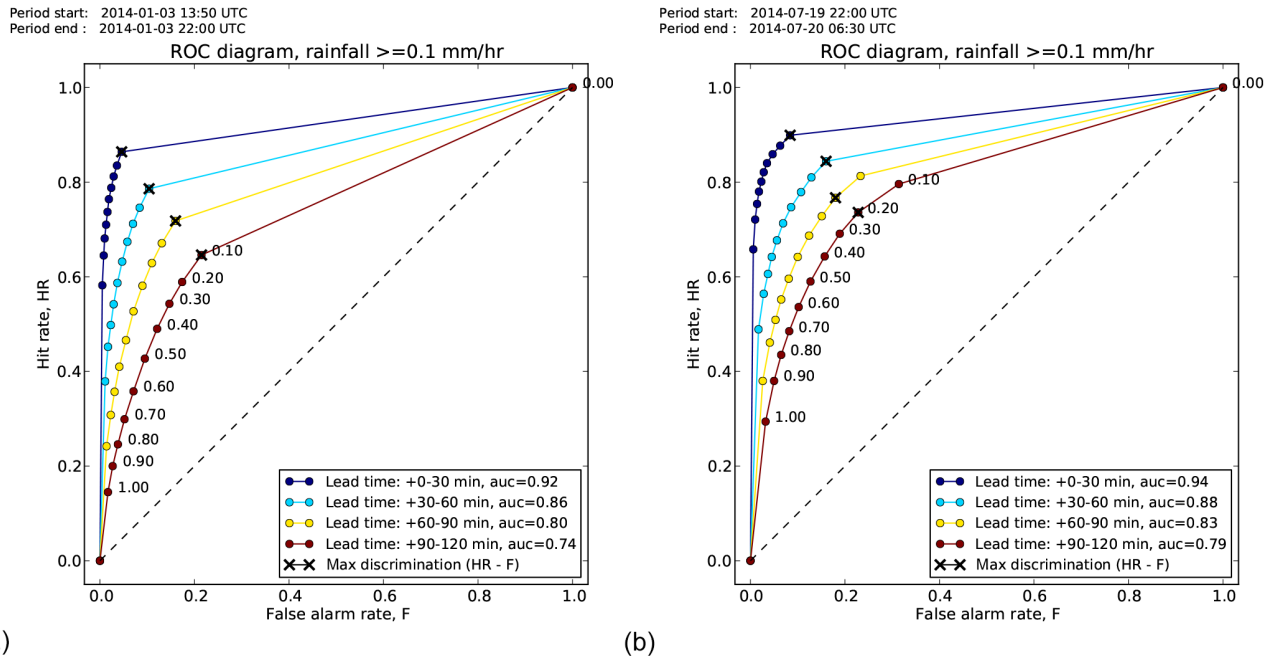
**Figure 8.** ROC curves relative to the probabilistic forecast of exceeding $0.1\,\mathrm{mm\,h^{-1}}$ for **(a)** the Ghent case on 3 January 2014 and **(b)** the Leuven case on 19–20 July 2014.

are less reliable than the ones of the convective case in Leuven for both rainfall thresholds. Probabilistic forecasts of exceeding $0.5\,\mathrm{mm\,h^{-1}}$ for the Ghent case have a good reliability and positive Brier skill score (BSS) up to 60 min lead time (Fig. 7a). The higher rainfall threshold of $5.0\,\mathrm{mm\,h^{-1}}$ is harder to predict and there is skill only up to 30 min lead time (Fig. 7b). The convective case in Leuven is more predictable and the probabilistic forecast of exceeding $0.5\,\mathrm{mm\,h^{-1}}$ exhibits skill up to 90 min lead time (Fig. 7c). It is interesting to note that forecast probabilities that are close to the climatological frequency (intersection of lines around the probability 0.15) can easily fall outside the skillful region (Fig. 7c). In fact, a small systematic forecast bias is likely to be worse than the event climatology at those frequencies. The rainfall threshold of $5.0\,\mathrm{mm\,h^{-1}}$ shows again a limit of predictability of 30 min (Fig. 7d). Despite having a negative BSS, the following lead times (Fig. 7d) have higher resolution than the stratiform case in Ghent (Fig. 7b).

Figure 8 illustrates the ROC curves relative to the probabilistic forecast of exceeding $0.1\,\mathrm{mm\,h^{-1}}$ for the Ghent case on 3 January 2014 (Fig. 8a) and the Leuven case on 19–20 July 2014 (Fig. 8b). All the ROC curves are very far from the diagonal line of no skill. The probability level that is marked with a cross is the one that maximizes the difference between the hit rate (HR) and the false alarm rate ($F$) (not to be confused with the false alarm ratio, which is conditioned on the forecasts). This point is located within the probabilities 0.1 and 0.2, which means that an optimal forecast of the probability of rain is achieved when only 10–20 % of the

ensemble members exceed the $0.1\,\mathrm{mm\,h^{-1}}$ threshold. A forecaster who is not scared of making false alarms would choose a lower probability level to increase the number of hits. On the contrary, an unconfident forecaster who would like to minimize the false alarms would choose a higher probability level, which has however the consequence of reducing the number of hits. As expected, the area under the ROC curves (AUC) decreases for increasing lead times. The discrimination skill for the convective event in Leuven is slightly higher than the one of the stratiform event in Ghent, which confirms the findings on the reliability diagrams (Fig. 7). This does not mean that small-scale features are easier to forecast than larger-scale features, which is known to be false (see Foresti and Seed, 2014). It means that the predictability of well-defined and organized convective systems is higher than the one of more moderate convection with shorter lifetime, at least for the cases analyzed in this paper.

### 4.4 Ensemble verification

Figure 9 compares the error of the ensemble mean (RMSE) and the ensemble spread for the Ghent case on 3 January 2014 and the Leuven case on 19–20 July 2014 (see interpretation of ensemble spread in Appendix A). In both cases the RMSE increases up to a lead time of 50–60 min and then starts a slow decrease, which can be counter-intuitive. However, it must be remembered that the ensemble mean forecast becomes smoother for increasing lead times, which reduces the double penalty error due to forecasting a thunderstorm at the wrong location. The ensemble spread also increases up to
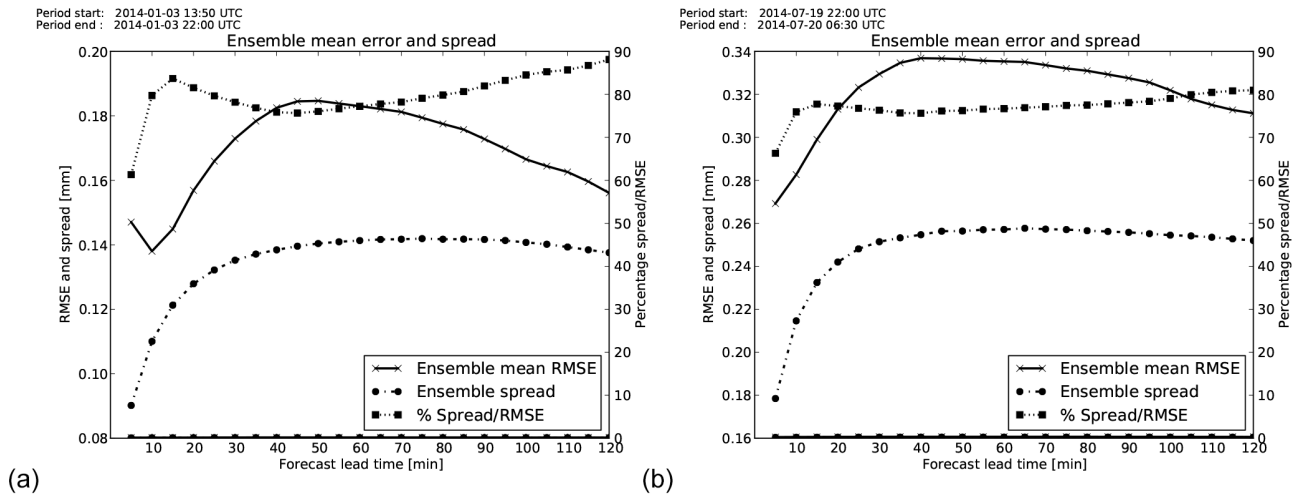
**Figure 9.** Comparison of ensemble spread and RMSE of the ensemble mean forecast at 5 min resolution for **(a)** the Ghent case on 3 January 2014 and **(b)** the Leuven case on 19–20 July 2014.
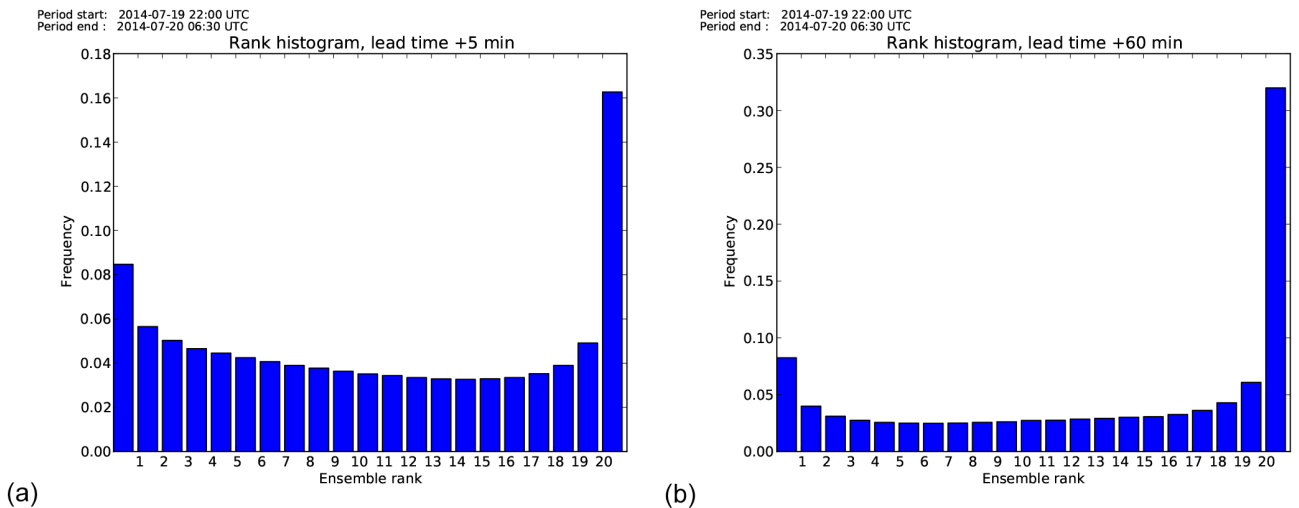


**Figure 10.** Rank histograms for the Leuven case on 19–20 July 2014 for a lead time of **(a)** 5 min and **(b)** 60 min.

50–60 min lead time and then slowly stabilizes. For both the Ghent and Leuven cases the ensemble spread is lower than the error of the ensemble mean at all lead times, which suggests that the ensemble forecasts are under-dispersive. The degree of under-dispersion is highest at a lead time of 5 min, with the spread values being equal to 60 % of the forecast error for the winter event in Ghent (Fig. 9a) and 70 % for the summer event in Leuven (Fig. 9b). Except for the 5 min lead time, the ensemble spread represents 75–90 % of the forecast error for the Ghent case (Fig. 9a) and 75–80 % for the Leuven case (Fig. 9b), which is a good result. It is not yet clear why the RMSE at a lead time of 5 min is higher than the one at 10 min for the winter case in Ghent (Fig. 9a).

The underestimation of the ensemble dispersion at the first lead time can be attributed to both the underestimation of the ensemble spread and the overestimation of the ensem-

ble mean RMSE, but with different degrees according to the different causes. High RMSEs at the start of the nowcast can be due to using a very smooth velocity field for the advection (see Sect. 3.2), which does not exploit sufficiently the very short-term predictability of small-scale precipitation features, but is optimized for predictions at longer lead times. Another explanation for this underestimation of ensemble dispersion could be due to the space–time variability of the Z–R relationship. Spatial and temporal changes in the drop size distribution (DSD) can lead to changes in the estimated rainfall rate that is used for the verification. Therefore, there could be a mismatch between the "fixed" DSD of the forecasts and the variable DSD underlying the verifying observations. Another possible source of mismatch could be due to the advection correction with optical flow when computing the rainfall accumulations. The forecast accumu-

lations are computed by advecting the previous rainfall field forward. On the other hand, the observed accumulations are computed by reversing the optical flow vectors and advecting the rainfall field backwards (see Sect. 4.1). This choice increases the differences when comparing the +0–5 min forecast accumulations (advection of the 0 min image forward) with the +0–5 min observed accumulations a posteriori (advection of the +5 min image backwards). The ideal approach would be to derive the accumulation by advecting both the previous image forward and the last image backwards. An optimal accumulation could be computed by a weighted average of the two advected images by discretizing the 5 min interval. However, such an approach is not very pragmatic and would require additional computational time in order to obtain a marginal improvement in the forecasts.

Figure 10 illustrates the rank histograms for the Leuven case on 19–20 July 2014 for lead times of 5 and 60 min. The U-shape of the rank histograms demonstrates again a certain degree of ensemble under-dispersion. In particular, all the ensemble members for the 5 min lead time are inferior to the observations in $\sim 16\%$ of the cases (Fig. 10a), while for the 60 min lead time it happens in more than $\sim 30\%$ of the cases (Fig. 10b). On the other hand, the fraction of observations falling below the value of the lowest ensemble member is only 8 % for both lead times. Despite the fact that STEPS is designed to reproduce the space–time variability of rainfall, it underestimates a certain fraction of the observed rainfall extremes. This underestimation grows with increasing lead time and depicts an increasing smoothness of the STEPS ensembles, which is probably due to the advection of the radar rainfall cascade (see Sect. 3, step 6). In fact, the small-scale rainfall features represented by the bottom cascade levels suffer more from numerical diffusion during the Lagrangian extrapolation, which is observed as a gradual loss of variability in the forecast ensembles.

## 4.5 Verification summary of the events

Table 2 provides a comparison of the verification scores for each event. The average standard deviation of the multiplicative biases of the 30 min lead time forecast is in the range 0.3–0.8. Except for the event on 19–20 July 2014 the biases remain well below 1 dB for all lead times, which is a positive result. Of course, these are average values, and locally they can even exceed 3 dB (see Fig. 4).

On the other hand, the RMSE values mark more the distinction between the two winter cases in Ghent and the two summer cases in Leuven. For the winter cases the RMSE values increase from 0.38–0.95 at a lead time of 30 min to 0.78–1.48 at 120 min, while for the summer cases from 1.84–2.45 to 2.52–3.38 mm h$^{-1}$. Thus, the RMSE of a 30 min lead time nowcast of the two convective cases is higher than the RMSE of a 120 min nowcast of the two stratiform cases, as might be expected. It is interesting to mention that linear verification scores such as the RMSE strongly depend on the variance

of the data. Consequently, it would be difficult to compare the error of the STEPS ensemble mean nowcast with the one of a deterministic nowcast, for example computed by INCA-BE. In fact, the ensemble averaging process filters out the unpredictable precipitation features and is rewarded in terms of RMSE. Similar results were already observed in Foresti et al. (2015), who also pointed out the difficulty of comparing ensemble prediction systems having a different number of ensemble members.

The probability of detection relative to the 0.5 mm h$^{-1}$ threshold decreases from 78–86 to 33–58 %, while the false alarm ratio increases from 10–17 to 46–65 %. The Gilbert skill score starts with values of 0.58–0.64 and 0.29–0.40 at the 30 and 60 min lead times, respectively, and decays to values of 0.08–0.20 at 120 min. Wang et al. (2009) reported a critical success index value of 0.45 for STEPS nowcasts of 0–60 accumulations relative to the 1 mm h$^{-1}$ threshold. Considering that the GSS is the CSI corrected by random chance, this value is comparable with the ones of the 30–60 min accumulations obtained in this paper. The GSS values relative to the threshold of 5.0 mm h$^{-1}$ are much lower. They oscillate between 0.15 and 0.44 for the first lead time and become very low and close to 0 afterwards. Thus, the predictability of rainfall structures exceeding 5.0 mm h$^{-1}$ rarely exceeds 30 min according to the GSS.

The areas under the ROC curve values characterizing the potential discrimination power of the probabilistic forecast of exceeding 0.5 mm h$^{-1}$ start at 0.92–0.95 at 30 min lead time and decrease to 0.69–0.79 at 120 min. For the probabilistic forecast of exceeding 5.0 mm h$^{-1}$ they start at 0.88–0.90 and decrease to 0.62 for the convective cases and to 0.50 for the stratiform cases (no discrimination).

From all these results we can conclude that there is not much predictability beyond 2 h lead time by extrapolating the 4 C-band composite radar images in Belgium. Therefore, a maximum lead time of 2 h in STEPS-BE is a good choice. Extending this lead time requires blending the radar-extrapolation nowcast with the output of NWP models to increase the predictability of precipitation.

## 5 Conclusions

The Short-Term Ensemble Prediction System (STEPS) is a probabilistic nowcasting system based on the extrapolation of radar images developed at the Australian Bureau of Meteorology in collaboration with the UK MetOffice. The principle behind STEPS is to produce an ensemble forecast by perturbing a deterministic extrapolation nowcast with stochastic noise. The perturbations are designed to reproduce the spatial and temporal correlations of the forecast errors and the scale dependence of the predictability of precipitation.

This paper presented the local implementation, adaptation and verification of STEPS at the Royal Meteorological Institute of Belgium, referred to as STEPS-BE. STEPS-BE pro-

**Table 2.** Summary of the forecast verification scores of the next four 30 min accumulation forecasts for the precipitation events in Ghent and Leuven. The lead time shown is the end of the 30 min accumulation period (e.g., 60 min is relative to the 30–60 min accumulation). The bias values correspond to the standard deviation of the multiplicative bias, which is more interesting than its mean (often close to 0).

| Event | Bias 30 min | Bias 60 min | Bias 90 min | Bias 120 min | RMSE 30 min | RMSE 60 min | RMSE 90 min | RMSE 120 min |
|---|---|---|---|---|---|---|---|---|
| | (dB) | | | | (mm h$^{-1}$) | | | |
| 10/11/2013 | 0.30 | 0.49 | 0.61 | 0.70 | 0.38 | 0.59 | 0.71 | 0.78 |
| 03/01/2014 | 0.54 | 0.74 | 0.82 | 0.89 | 0.95 | 1.39 | 1.53 | 1.48 |
| 9–10/06/2014 | 0.52 | 0.63 | 0.66 | 0.69 | 2.45 | 3.26 | 3.40 | 3.38 |
| 19–20/07/2014 | 0.84 | 1.18 | 1.30 | 1.35 | 1.84 | 2.36 | 2.49 | 2.52 |
| Event | POD 30 min | POD 60 min | POD 90 min | POD 120 min | FAR 30 min | FAR 60 min | FAR 90 min | FAR 120 min |
| | Forecast $>= 0.5$ mm h$^{-1}$ | | | | Forecast $>= 0.5$ mm h$^{-1}$ | | | |
| 10/11/2013 | 0.83 | 0.71 | 0.62 | 0.54 | 0.17 | 0.30 | 0.38 | 0.46 |
| 03/01/2014 | 0.80 | 0.63 | 0.49 | 0.33 | 0.10 | 0.25 | 0.45 | 0.65 |
| 9–10/06/2014 | 0.78 | 0.65 | 0.55 | 0.46 | 0.15 | 0.32 | 0.44 | 0.54 |
| 19–20/07/2014 | 0.86 | 0.75 | 0.66 | 0.58 | 0.17 | 0.36 | 0.50 | 0.61 |
| Event | GSS 30 min | GSS 60 min | GSS 90 min | GSS 120 min | GSS 30 min | GSS 60 min | GSS 90 min | GSS 120 min |
| | Forecast $>= 0.5$ mm h$^{-1}$ | | | | Forecast $>= 5.0$ mm h$^{-1}$ | | | |
| 10/11/2013 | 0.58 | 0.38 | 0.27 | 0.20 | 0.15 | 0.02 | 0.0 | 0.0 |
| 03/01/2014 | 0.64 | 0.40 | 0.20 | 0.08 | 0.28 | 0.06 | 0.0 | 0.0 |
| 9–10/06/2014 | 0.59 | 0.38 | 0.26 | 0.17 | 0.44 | 0.20 | 0.09 | 0.04 |
| 19–20/07/2014 | 0.58 | 0.29 | 0.14 | 0.07 | 0.27 | 0.09 | 0.04 | 0.02 |
| Event | AUC 30 min | AUC 60 min | AUC 90 min | AUC 120 min | AUC 30 min | AUC 60 min | AUC 90min | AUC 120 min |
| | Forecast $>= 0.5$ mm h$^{-1}$ | | | | Forecast $>= 5.0$ mm h$^{-1}$ | | | |
| 10/11/2013 | 0.95 | 0.89 | 0.84 | 0.79 | 0.88 | 0.67 | 0.56 | 0.50 |
| 03/01/2014 | 0.92 | 0.85 | 0.78 | 0.69 | 0.90 | 0.72 | 0.57 | 0.50 |
| 9–10/06/2014 | 0.93 | 0.86 | 0.81 | 0.76 | 0.89 | 0.77 | 0.68 | 0.62 |
| 19–20/07/2014 | 0.94 | 0.87 | 0.82 | 0.77 | 0.88 | 0.75 | 0.68 | 0.62 |

duces in real-time 20-member ensemble nowcasts at 1 km and 5 min resolutions up to 2 h lead time using the four C-band radar composite of Belgium. Compared with the original implementation, STEPS-BE includes a kernel-based interpolation of optical flow vectors to obtain smoother velocity fields and an improvement to generate stochastic noise only within the advected radar composite to respect the validity domain of the nowcasts.

The performance of STEPS-BE was verified using the radar observations as reference on four case studies that caused sewer system floods in the cities of Ghent and Leuven during the years 2013 and 2014. The ensemble mean forecast of the next four 30 min accumulations was verified using the multiplicative bias, the RMSE as well as some categorical scores derived from the contingency table: the probability of detection, false alarm ratio and Gilbert skill score (equitable skill score). The spatial distribution of multiplicative biases revealed regions of systematic over- and under-estimation by STEPS. The underestimations are often associated with the locations of convective initiation and thunderstorm growth, which cannot be predicted by STEPS. On the other hand, the regions of overestimation are mostly due to the underestimation of rainfall by the verifying observations rather than systematic rainfall decay (see Foresti and Seed, 2015, for a more detailed discussion). In order to disentangle the forecast and observation biases, detailed knowledge about the spatial distribution of the radar measurement errors for a given weather situation is needed. The multiplicative biases over the cities of Leuven and Ghent are very low (from −0.5 to +0.5 dB), which is a good starting point to integrate STEPS nowcasts as inputs into sewer system hydraulic models. The categorical forecast verification helped discovering the places with low probability of detection due to convective initiation at the front of the rain band and high false alarm ratio at the

rear of the rain band, likely due to a too slow rainfall extrapolation by STEPS. Reliability diagrams demonstrated that probabilistic forecasts of exceeding $0.5\,\mathrm{mm\,h^{-1}}$ have skill up to 60–90 min lead time. On the contrary, convective features exceeding $5.0\,\mathrm{mm\,h^{-1}}$ are only predictable up to 30 min. In terms of reliability and discrimination ability, it was also observed that the forecasts of convective events have more skill than the ones of stratiform events. The STEPS ensembles are characterized by a certain degree of underestimation of the forecast uncertainty, with values of the ensemble spread close to 75–90 % of the forecast error.

The current contribution focused on the verification of STEPS-BE nowcasts using only four precipitation cases of different character. The deterministic and categorical verifications require many more cases to analyze the climatological distribution of the forecast errors, e.g., as done in Foresti and Seed (2015). On the other hand, the probabilistic and ensemble verification pools the data in both space and time and converges much faster to stable statistics.

From a research perspective, STEPS-BE could also be extended by including a stochastic model to account for the residual radar measurement errors, in particular to obtain more accurate estimations of the forecast uncertainty at short range. The STEPS framework also allows blending of the extrapolation nowcast with the output of NWP models, which is a necessary step to increase the predictability of precipitation for lead times beyond 2 h.

**Table A1.** Correspondence between the decibel scale and the power ratio.

| dB | −6 | −3 | −1 | −0.5 | 0 | +0.5 | +1 | +3 | +6 |
|---|---|---|---|---|---|---|---|---|---|
| Power ratio $(F/O)$ | 0.251 | 0.501 | 0.794 | 0.891 | 1 | 1.122 | 1.259 | 1.995 | 3 .981 |

## Appendix A: Forecast verification scores

Forecast verification is an important aspect of a forecasting system. A forecast without an estimation of its accuracy is not very informative. For an in-depth description of forecast verification science and corresponding scores we refer to Jolliffe and Stephenson (2011) and the verification website maintained at the Bureau of Meteorology (http://www.cawcr.gov.au/projects/verification/).

The STEPS *ensemble mean forecast* was verified using the following scores.

– Multiplicative bias:

$$\text{bias} = \frac{1}{N} \sum_{i=1}^{N} 10\log_{10}\left(\frac{F_i + b}{O_i + b}\right), \tag{A1}$$

where $F_i$ is the forecast rainfall at a given grid point, $O_i$ is the observed rainfall at a given grid point, $b = 2\,\text{mm}\,\text{h}^{-1}$ is an offset to eliminate the division by zero and to reduce the contribution of the forecast errors at low rainfall intensities, and $N$ is the number of samples. For the specific case of the verification of the spatial distribution of forecast biases, the summation is performed over time. Thus, $N$ corresponds to the number of forecasts where either the forecast or the observed rainfall are greater than $0.1\,\text{mm}\,\text{h}^{-1}$ at a given grid point (denoted as weak conditional verification). The bias is given in decibels (dB) in order to obtain a more symmetric distribution of the multiplicative errors centered at 0, which is not possible with the simple power ratio $F/O$. Table A1 summarizes the correspondence between the decibel scale and the power ratio. For example, a bias of $+3$ dB occurs when the forecast rainfall $F$ is twice as much as the observed rainfall $O$.

– Root mean square error:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (F_i - O_i)^2} \tag{A2}$$

– Contingency table of a dichotomous (yes/no) forecast, see Table A2, where the "hits" is the number of times that both the observation and the forecast exceed a given rainfall threshold (at a given grid point), the "false alarms" is the number of times that the forecast exceeds the threshold but the observation does not, the "misses" is the number of times that the forecast does not exceed

**Table A2.** Contingency table of a categorical forecast.

| | | Observed | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Forecast | Yes | Hits | False alarms | Forecast yes |
| | No | Misses | Correct negatives | Forecast no |
| | Total | Observed yes | Observed no | Total |

the threshold but the observation does and the "correct negatives" is the number of times that both the observation and the forecast do not exceed the threshold.

– Different scores can be derived from the contingency table to characterize a particular feature or skill of the forecasting system:

  – Probability of detection (hit rate):

  $$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} = \frac{\text{hits}}{\text{observed yes}}, \tag{A3}$$

  The "POD" characterizes the fraction of observed events that were correctly forecast and is also known as the hit rate (HR).

  – False alarm ratio:

  $$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} = \frac{\text{false alarms}}{\text{forecast yes}}, \tag{A4}$$

  The "FAR" characterizes the fraction of forecast events that were wrongly forecast.

  – False alarm rate:

  $$F = \frac{\text{false alarms}}{\text{false alarms} + \text{correct negatives}}$$
  $$= \frac{\text{false alarms}}{\text{observed no}}. \tag{A5}$$

  The false alarm rate $F$ is conditioned on the observations, while the false alarm ratio FAR on the forecasts.

  – Gilbert skill score (equitable threat score):

  $$\text{GSS} = \tag{A6}$$
  $$\frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}},$$

  where

  $$\text{hits}_{\text{random}} = \frac{(\text{hits} + \text{misses})(\text{hits} + \text{false alarms})}{\text{total}}$$
  $$= \frac{(\text{observed yes})(\text{forecast yes})}{\text{total}} \tag{A7}$$

  is the number of hits obtained by random chance, which is calculated by multiplying the marginal

sums of the observed and forecast events (such as computing the theoretical frequencies for the Chi-squared test). The GSS characterizes the detection skill of the forecasting system with respect to random chance. In practice it corresponds to the critical success index (CSI) adjusted for the hits obtained by random chance.

The accuracy of probabilistic forecasts can be verified in various ways. In this paper we employ the reliability diagram and the Receiver Operating Characteristic (ROC) curve. The reliability diagram compares the forecast probability with the observed frequency. Reliability characterizes the agreement between the forecast probability and observed frequency. For a reliable forecasting system the two values should be the same, which happens for example when we observe rain 80 % of the time when it is forecast with 80 % probability (in average, diagonal line of Fig. 8). Unreliable forecasts exhibit departures from this optimum (bias). Resolution characterizes the ability of the forecasts to categorize the observed frequencies into distinct classes. The complete lack of resolution occurs when the forecast probabilities are completely unable to distinguish the observed frequencies, which generally corresponds to the climatological frequency of exceeding a given precipitation threshold (horizontal dashed line in Fig. 8). The Brier skill score (BSS) characterizes the relative accuracy of the probabilistic forecast compared to a reference system (see Jolliffe and Stephenson, 2011). Although the climatology or sample climatology of the event is often used as a reference, the BSS can also be computed against other reference forecasts, e.g., another probabilistic forecasting method or even a deterministic forecasting method treated as a probabilistic binary forecast. However, in such cases it is not possible to draw a unique horizontal line representing complete lack of skill in Fig. 8. The region where the probabilistic forecast has a positive BSS, i.e., it is better than the climatological frequency, is grayed out. In fact, the points located below the no skill line are closer to the climatological frequency and produce a negative BSS. Reliability diagrams usually contain the histogram of the forecast probabilities to analyze the sharpness of the forecasts (small inset in Fig. 8). Sharpness characterizes the ability to forecast probabilities that are different from the reference forecast. Sharp forecasting systems are "confident" about their predictions and give many probabilities around one and zero.

The ROC curve is used to analyze the discrimination power of a probabilistic forecast of exceeding a given rainfall threshold. It is constructed by plotting the hit rates and false alarm rates evaluated at increasing probability thresholds to make the binary decision whether it will rain or not. The ROC curve of a random probabilistic forecast system lies on the diagonal where the hit rate equals the false alarm rate (no skill): the forecast probabilities do not have discrimination power. When the false alarm rate is higher than the hit rate the forecast is worse than that obtained by random

chance (below the diagonal). A skilled forecasting system is observed when the hit rates are higher than the false alarm rates, which draws a characteristic curve. The area under the ROC curve (AUC) measures the discrimination power of the probabilistic forecasts, with a maximum value of 1 (100 % of hits and 0 % of false alarms) and a minimum value of 0.5 for a random forecasting system. Values below 0.5 denote a forecasting system that performs worse than random chance. The AUC is computed by integrating over all the trapezoids that can be drawn below the ROC curve. The AUC is not sensitive to the forecast bias and the reliability of the forecast could still be improved through calibration. For this reason the AUC is only a measure of potential skill.

The ensemble forecasts are verified to detect whether there is over- or under-dispersion. It is common practice to compare the "skill" (error) of the ensemble mean with the ensemble spread (Whitaker and Loughe, 1998; Foresti et al., 2015):

$$\text{spread} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\frac{1}{M-1} \sum_{m=1}^{M} \left( F_{im} - \overline{F_i} \right)^2}, \qquad \text{(A8)}$$

where $M$ is the number of ensemble members (ensemble size), $F_{im}$ is the forecast of a given ensemble member and $\overline{F_i}$ is the ensemble mean forecast (at a given grid point). Since we are not analyzing the spatial or temporal distribution of the ensemble spread, $N$ corresponds to the total number of samples in space and time, which is the number of forecasts within a rainfall event multiplied by the number of grid points within a radar field. The weak conditional verification is also applied to the computation of the spread. The ensemble spread characterizes the variability of the ensemble members about the ensemble mean (standard deviation). For a good ensemble prediction system, the ensemble spread should be equal to the average variability of the observations about the ensemble mean, as measured by the RMSE of the ensemble mean (Eq. 2). If the spread is larger than the RMSE, the ensemble is overestimating the forecast uncertainty (over-dispersion), otherwise it is underestimating it (under-dispersion). It is interesting to mention that the ensemble mean RMSE and ensemble spread could also be computed starting from the logarithm of rainfall rates to account for the skewed distribution of precipitation (not used in this paper).

Another way to analyze the spread of ensemble forecasts is based on rank histograms (also known as a Talagrand diagram). First, the precipitation values of the ensemble members are ranked in increasing order. Then, the rank of the observation is evaluated by checking in which of the $M + 1$ bins it falls. By repeating the operation for a large number of cases and forecasts it is possible to construct a histogram. A good ensemble prediction system displays a flat histogram; i.e., the observations are indistinguishable from the forecasts and each ensemble member is an equi-probable realization of the future state of the atmosphere. A bell-shaped histogram with a peak in the middle is observed in the case of ensemble

over-dispersion. On the contrary, a U-shape histogram with peaks at the edges is observed in the case of ensemble under-dispersion, which is more common (in particular for NWP ensembles). In this case the values of the observations often fall below or above the lowest or highest value of the ranked ensemble, which is not dispersive enough to capture the extremes.

Edited by: R. Uijlenhoet

# References

Achleitner, S., Fach, S., Einfalt, T., and Rauch, W.: Nowcasting of rainfall and of combined sewage flow in urban drainage systems, Water Sci. Technol., 59, 1145–51, 2009.

Atencia, A. and Zawadzki, I.: A comparison of two techniques for generating nowcasting ensembles – Part I: Lagrangian ensemble technique, Mon. Weather Rev., 142, 4036–4052, 2014.

Berenguer, M., Corral, C., Sánchez-Diezma, R., and Sempere-Torres, D.: Hydrological validation of a radar-based nowcasting technique, J. Hydrometeorol., 6, 532–549, 2005.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G. S.: SBMcast – an ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, J. Hydrol., 404, 226–240, 2011.

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

Bowler, N. E. H., Pierce, C. E., and Seed, A. W.: Development of a precipitation nowcasting algorithm based upon optical flow techniques, J. Hydrol., 288, 74–91, 2004a.

Bowler, N. E. H., Pierce, C. E., and Seed, A. W.: STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Forecast Research Technical Report No. 433, MetOffice, Exeter, UK, 2004b.

Bowler, N. E. H., Pierce, C. E., and Seed, A. W.: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Q. J. Roy. Meteor. Soc., 132, 2127–2155, 2006.

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., and ten Veldhuis, J. A. E.: On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution, Hydrol. Earth Syst. Sci., 19, 691–709, doi:10.5194/hess-19-691-2015, 2015.

Cloke, H. L. and Pappenberger, F.: Ensemble flood forecasting: a review, J. Hydrol., 375, 613–626, 2009.

Collier, C. G.: On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models, Meteorol. Appl., 16, 35–40, 2009.

Dai, Q., Rico-Ramirez, M. A., Han, D., Islam, T., and Liguori, S.: Probabilistic radar rainfall nowcasts using empirical and theoretical uncertainty models, Hydrol. Process., 29, 66–79, 2015.

Ebert, E. E.: Ability of a poor man's ensemble to predict the probability and distribution of precipitation, Mon. Weather Rev., 129, 2461–2480, 2001.

Ehret, U., Götzinger, J., Bárdossy, A., and Pegram, G. G. S.: Radar-based flood forecasting in small catchments, exemplified by the Goldersbach catchment, Germany, International Journal of River Basin Management, 6, 323–329, 2008.

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

Figueras i Ventura, J. and Tabary, P.: The new French operational polarimetric radar rainfall rate product, J. Appl. Meteorol. Clim., 52, 1817–1835, 2013.

Foresti, L. and Seed, A.: The effect of flow and orography on the spatial distribution of the very short-term predictability of rainfall from composite radar images, Hydrol. Earth Syst. Sci., 18, 4671–4686, doi:10.5194/hess-18-4671-2014, 2014.

Foresti, L. and Seed, A.: On the spatial distribution of rainfall nowcasting errors due to orographic forcing, Meteorol. Appl., 22, 60–74, 2015.

Foresti, L., Seed, A., and Zawadzki, I.: Report of the Heuristic Probabilistic Forecasting Workshop, Munich, Germany, 13 pp., 30–31 August 2014, available at: https://sites.google.com/site/lorisforesti/projects/nowcasting/ScientificReport_HeuristicProbForecastingWorkshop_Munich_2014_121214.pdf, last access: 3 July 2015, 2014.

Foresti, L., Panziera, L., Mandapaka, P. V., Germann, U., and Seed, A.: Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall, Meteorol. Appl., 22, 141–155, 2015.

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

Germann, U. and Zawadzki, I.: Scale-dependence of the predictability of precipitation from continental radar images – Part II: Probability forecasts, J. Appl. Meteorol., 43, 74–89, 2004.

Germann, U., Berenguer, M., Sempere-Torres, D., and Zappa, M.: REAL – ensemble radar precipitation estimation for hydrology in a mountainous region, Q. J. Roy. Meteor. Soc., 135, 445–456, 2009.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.

Grasso, L. D.: The differentiation between grid spacing and resolution and their application to numerical modeling, B. Am. Meteorol. Soc., 81, 579–580, 2000.

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern Alpine region, Weather Forecast., 26, 166–183, 2011.

Hohti, H., Koistinen, J., Nurmi, P., Saltikoff, E., and Holmlund, K.: Precipitation nowcasting using radar-derived atmospheric motion vectors, in: Proc. of the 1st European Conf. on Radar in Meteorology and Hydrology (ERAD), Bologna, Italy, 4–8 September 2000, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25, 1323–1327, 2000.

Jolliffe, I. T. and Stephenson, D. B.: Forecast Verification: a Practitioner's Guide in Atmospheric Science, 2nd edn., John Wiley and Sons, Chichester, 2011.

Jordan, P., Seed, A. W., and Weinnman, P. E.: A stochastic model of radar measurement errors in rainfall accumulations at catchment scale, J. Hydrometeorol., 4, 841–855, 2003.

Lewis, H., Mittermaier, M., Mylne, K., Norman, K., Scaife, A., Neal, R., Pierce, C., Harrison, D., Jewell, S., Kendon, M., Saunders, R., Brunet, G., Golding, B., Kitchen, M., Davies, P., and Pilling, C.: From months to minutes – exploring the value of high-resolution rainfall observation and prediction during the UK winter storms of 2013/2014, Meteor. Appl., 22, 90–104, 2015.

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

Liguori, S. and Rico-Ramirez, M. A.: A practical approach to the assessment of probabilistic flow predictions, Hydrol. Process., 27, 18–32, 2013.

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

Metta, S., Rebora, N., Ferraris, L., von Hardernberg, J., and Provenzale, A.: PHAST: a phase-diffusion model for stochastic nowcasting, J. Hydrometeorol., 10, 1285–1297, 2009.

Panziera, L., Germann, U., Gabella, M., and Mandapaka, P. V.: NORA – nowcasting of orographic rainfall by means of analogues, Q. J. Roy. Meteor. Soc., 137, 2106–2123, 2011.

Pappenberger, F. and Beven, K. J.: Ignorance is bliss: or seven reasons not to use uncertainty analysis, Water Resour. Res., 42, W05302, doi:10.1029/2005WR004820, 2006.

Paschalis, A., Molnar, P., Fatichi, S., and Burlando, P.: A stochastic model for high-resolution space–time precipitation simulation, Water Resour. Res., 49, 8400–8417, 2013.

Pegram, G. G. S. and Clothier, A. N.: High resolution space–time modelling of rainfall: the "String of Beads" model, J. Hydrol., 241, 26–41, 2001a.

Pegram, G. G. S. and Clothier, A. N.: Downscaling rainfields in space and time, using the String of Beads model in time series mode, Hydrol. Earth Syst. Sci., 5, 175–186, doi:10.5194/hess-5-175-2001, 2001b.

Pierce, C., Bowler, N., Seed, A., Jones, A., Jones, D., and Moore, R.: Use of a stochastic precipitation nowcast scheme for fluvial flood forecasting and warning, Atmos. Sci. Lett., 6, 78–83, 2005.

Pierce, C., Hirsch, T., and Bennett, A. C.: Formulation and evaluation of a post-processing algorithm for generating seamless, high resolution ensemble precipitation forecasts, Forecasting R&D Technical Report 550, MetOffice, Exeter, UK, 2010.

Radhakrishna, B., Zawadzki, I., and Fabry, F.: Predictability of precipitation from continental radar images. Part V: growth and decay, J. Atmos. Sci., 69, 3336–3349, 2012.

Roulin, E. and Vannitsem, S.: Skill of medium-range hydrological ensemble predictions, J. Hydrometeorol., 6, 729–744, 2005.

Schellekens, J., Weerts, A. H., Moore, R. J., Pierce, C. E., and Hildon, S.: The use of MOGREPS ensemble rainfall forecasts in operational flood forecasting systems across England and Wales, Adv. Geosci., 29, 77–84, doi:10.5194/adgeo-29-77-2011, 2011.

Schertzer, D. and Lovejoy, S.: Physical modelling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res., 92, 9696–9714, 1987.

Seed, A.: A dynamic and spatial scaling approach to advection forecasting, J. Appl. Meteorol., 42, 381–388, 2003.

Seed, A. W., Pierce, C. E., and Norman, K.: Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme, Water Resour. Res., 49, 6624–6641, 2013.

Silvestro, F. and Rebora, N.: Operational verification of a framework for the probabilistic nowcasting of river discharge in small and medium size basins, Nat. Hazards Earth Syst. Sci., 12, 763–776, doi:10.5194/nhess-12-763-2012, 2012.

Silvestro, F., Rebora, N., and Cummings, G.: An attempt to deal with flash floods using a probabilistic hydrological nowcasting chain: a case study, Nat. Hazards Earth Syst. Sci. Discuss., 1, 7497–7515, doi:10.5194/nhessd-1-7497-2013, 2013.

Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J., Joe, P., Barker, D. M., Li, P.-W., Golding, B., Xu, M., and Pinto, J.: Use of NWP for nowcasting convective precipitation: recent progress and challenges, B. Am. Meteor. Soc., 95, 409–426, 2014.

Tabary, P.: The new French operational radar rainfall product – Part I: Methodology, Weather Forecast., 22, 393–408, 2007.

Thielen, J., Bartholmes, J., Ramos, M.-H., and de Roo, A.: The European Flood Alert System – Part 1: Concept and development, Hydrol. Earth Syst. Sci., 13, 125–140, doi:10.5194/hess-13-125-2009, 2009.

Thorndahl, S. and Rasmussen, M. R.: Short-term forecasting of urban storm water runoff in real-time using extrapolated radar rainfall data, J. Hydroinform., 15, 897–912, 2013.

Turner, B. J., Zawadzki, I., and Germann, U.: Predictability of precipitation from continental radar images. Part III: Operational nowcasting implementation (MAPLE), J. Appl. Meteorol., 43, 231–248, 2004.

Venugopal, V., Foufoula-Georgiou, E., and Sapozhnikov, V.: Evidence of dynamic scaling in space–time rainfall, J. Geophys. Res., 104, 31599–31610, 1999.

Verworn, H. R., Rico-Ramirez, M. A., Krämer, S., Cluckie, I., and Reichel, F.: Radar-based flood forecasting for river catchments, Water Manage., 162, 159–168, 2009.

Wang, J. Keenan, T., Joe, P., Wilson, J., Lai, E. S. T., Liang, F., Wang, Y., Ebert, E. E., Ye, Q., Bally, J., Seed, A., Chen, M., Xue, J., and Conway, B.: Overview of the Beijing 2008 Olympics Project. Part I: Forecast Demonstration Project, WMO World Weather Research Programme, Report, 133 pp., 2009.

Whitaker, J. S. and Loughe, A. F.: The relationship between ensemble spread and ensemble mean skill, Mon. Weather Rev., 26, 3292–3302, 1998

Willems, P.: A spatial rainfall generator for small spatial scales, J. Hydrol., 252, 126–144, 2001a.

Willems, P.: Stochastic description of the rainfall input errors in lumped hydrological models, Stoch. Env. Res. Risk A., 15, 132–152, 2001b.

Skamarock, W. C.: Evaluating mesoscale NWP models using kinetic energy spectra, Mon. Weather Rev., 132, 3019–3032, 2004.

Xuan, Y., Cluckie, I. D., and Wang, Y.: Uncertainty analysis of hydrological ensemble forecasts in a distributed model utilising short-range rainfall prediction, Hydrol. Earth Syst. Sci., 13, 293–303, doi:10.5194/hess-13-293-2009, 2009.

Xuan, Y., Zhu, D., Triballi, P., and Cluckie, I.: Forecast uncertainty of a lumped hydrological model coupled with the STEPS radar rainfall nowcasts, in: Int. Symp. Weather Radar and Hydrol., Washington DC, US, 7–10 April 2014, 9 pp., 2014.

Zappa, M., Beven, K., Bruen, M., Cofino, A., Kok, K., Martin, E., Nurmi, P., Orfila, B., Roulin, E., Seed, A., Schroter, K., Szturc, J., Vehvilainen, B., Germann, U., and Rossa, A.: Propagation of uncertainty from observing systems and NWP into hydrologi-

cal models: COST-731 Working Group 2, Atmos. Sci. Lett., 11, 83–91, 2010.

# Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements

EDOUARD GOUDENHOOFDT AND LAURENT DELOBBE

*Royal Meteorological Institute of Belgium, Brussels, Belgium*

## ABSTRACT

Volumetric measurements from a C-band weather radar in Belgium are reprocessed over the years 2005–14 to improve the quantitative precipitation estimation (QPE). The data quality is controlled using static clutter and beam blockage maps and clutter identification based on vertical gradients, horizontal texture, and satellite observations. A new QPE is obtained using stratiform–convective classification, a 40-min averaged vertical profile of reflectivity (VPR), a brightband identification, and a specific transformation to rain rates for each precipitation regime. The rain rates are interpolated on a 500-m Cartesian grid, linearly accumulated, and combined with hourly rain gauge measurements using mean field bias or kriging with external drift (KED). The algorithms have been fine-tuned on 13 cases with various meteorological situations. A detailed validation against independent daily rain gauge measurements reveals the importance of VPR correction. A 10-yr verification shows a significant improvement of the new QPE, especially at short and long range, with roughly 50% increase in coverage. Adding the KED allows average improvements of 38%, 35%, and 80% for the mean absolute difference, the multiplicative error spread, and the fraction of good estimates, respectively. The benefit is higher in widespread situations and increases when considering higher rainfall amounts. The mitigation of radar artifacts is clearly visible on 10-yr statistics, including mean annual totals, probabilities to exceed 10 mm, and maxima for hourly and daily accumulation. The correlation of mean totals with rain gauges increases from 0.54 to 0.66 with the new QPE and to 0.8 adding KED.

## 1. Introduction

Surface precipitation impacts human activities at a wide range of space and time scales. Therefore, a detailed knowledge of its characteristics is needed for various applications. Predicting surface precipitation is one of the biggest challenges for numerical weather models, and thus, the verification of forecast accuracy for this variable is important. Precipitation is also the main driving force in agriculture and hydrology. The interest of precipitation observations at high spatial and temporal resolution is particularly high in urban hydrology. For many applications, long time series of precipitation are needed. In hydrology, the estimation of areal rainfall extreme statistics is required for risk assessment. It is also useful for the verification of regional climate models. In a changing climate the evolution

of the precipitation characteristics must be carefully monitored.

Various efforts to construct reference precipitation datasets have been reported in the literature. The simplest approach is the collection of gauge measurements that are interpolated on a gridded dataset. In Europe, the reference 50-yr daily E-OBS (Haylock et al. 2008) is used in many studies. The main weakness remains the low spatial representativeness of rain gauges that can miss small convective cells. Model reanalyses offer estimations over the globe but are prone to large biases (de Leeuw et al. 2015). Their ever-increasing resolutions are still not able to resolve convection properly. Satellite-based precipitation estimation is available at the global scale from infrared imagers or passive microwave radiometers, but their resolution and accuracy are still limited (Tang et al. 2014). More reliable estimation is available from spaceborne precipitation radar, but only in the tropics (Yang and Nesbitt 2014).

The potential of single-polarization ground-based weather radar to provide a reference precipitation dataset at very high temporal and spatial resolution with

*Corresponding author address*: Edouard Goudenhoofdt, Observations Department, Royal Meteorological Institute of Belgium, Avenue Circulaire 3, 1180 Brussels, Belgium.
E-mail: edouard.goudenhoofdt@meteo.be

good accuracy is large. Dual-polarization technology offers a greater potential, but long time series are not available yet. Using merged radar–gauge products, Overeem et al. (2009) derived annual maximum rainfall depths for durations from 15 min to 24 h and area sizes from 6 to $1.7 \times 10^3$ km$^2$ for the Netherlands. A reanalysis based on the French radar network demonstrated good results for streamflow simulation in some situations (Lobligeois et al. 2014). However, rainfall estimation from weather radar suffers from many sources of error and uncertainties that have been extensively discussed in the literature. The reader is referred to Uijlenhoet and Berne (2008) and Villarini and Krajewski (2010) for a detailed review. This has prevented the wide usage of radar observations in operational applications (e.g., hydrology). Recently, the most important issues have been tackled for operational quantitative precipitation estimation (QPE) algorithms (Tabary 2007; Germann et al. 2006). Since most weather services only archive 2D operational radar products (Nelson et al. 2010; Fairman et al. 2015), the quality of their dataset depends on the operational algorithm that was used to derive 2D products from the polar scans. Because of inherent updates of QPE algorithms, time periods with consistent data are limited. Improving the 2D radar product quality without 3D information is a difficult task (Wagner et al. 2012; Tabary 2007). For example, it is not straightforward to correct for overestimation due to the melting layer in an interpolated 2D product. Archiving of volume radar data was not common a decade ago because of telecommunication and storage costs and the lack of interest in precipitation reanalysis. Therefore, the number of long-term precipitation datasets making use of the full volumetric radar information is limited. Thorndahl et al. (2014) made a reanalysis of a relatively complete set of 10-yr volume data in Denmark. It has been produced using careful processing but without a correction for the vertical profile of reflectivity (VPR). Krajewski et al. (2011) developed a flexible framework for surface rainfall dataset generation from NEXRAD level 2 (volumetric) measurements that are available since 2002. The produced datasets have been used in several hydrological studies (Smith et al. 2012).

At the Royal Meteorological Institute of Belgium (RMIB), the volumetric data from a single-polarization C-band Doppler radar are archived since 2002. Those data have already been used to derive convective storm-track statistics in Goudenhoofdt and Delobbe (2013). The operational QPE algorithm used since 2004 is an interpolation of reflectivity data from different elevation angles at a given height combined with a Marshall–Palmer relationship. This algorithm, which has been applied on the archived volume data using a height of

800 m above the radar level, is defined as QPE1. In this study, a careful processing (defined as QPE2) of the volumetric data is made, including beam blockage correction, application of a static clutter map, dynamical clutter identification, correction for the height of the measurement (i.e., VPR correction), and specific reflectivity–rain rate ($Z$–$R$) relationships. A flowchart summarizing the processing steps and data flow of QPE2 is presented in Fig. 1. The algorithms have been developed based on a literature review and with simplicity and robustness in mind for future operational use in real time. A dense hourly rain gauge network with quality control available since 2005 is used in combination with the radar-based QPE using two methods. The new QPE algorithms are compared to QPE1 in order to evaluate the improvements. It is first validated on a selection of cases and then on the period 2005–15 using an independent rain gauge network. To our knowledge, such long-term verification of radar-based precipitation estimates has not yet been done. Several interesting statistics are derived from the 10-yr dataset and compared to rain gauge statistics.

## 2. Weather radar measurements

### a. Radar reflectivity measurement

Since 2001, RMIB has been operating a C-band (5.62 GHz) Doppler radar located in Wideumont (49.9°N, 5.5°E), southeast Belgium, at 592 m above sea level (Fig. 2). With measurements up to 240 km, the radar covers Belgium, Luxembourg, and parts of France, the Netherlands, and Germany. It is a single-polarization radar with Doppler capability used to filter ground echoes. During the study period, the radar performed a dedicated reflectivity scan at five elevation angles (0.3°, 0.6°, 1.8°, 3.3°, and 6.0°) every 5 min with a pulse repetition frequency of 600 Hz and pulse duration of 0.836 μs The radar volume data have a resolution of 1° in azimuth (an average of 33 pulses) and 250 m in range (an average of two successive range bins). Because of the height of the radar tower and its position near the top of the Ardennes ridge, beam blockage effects are relatively limited. More information regarding the radar characteristics and scanning strategy can be found in Delobbe and Holleman (2006). The volumetric data of the Wideumont radar have been archived at RMIB since 2002. No significant changes in the radar calibration have been encountered, except in April 2013 when software and hardware updates occurred. During the update the radar constants have been reset to default, resulting in a decrease of 1.3 dB. The Doppler filtering is active since April 2004 and is combined with a clutter-to-signal ratio (CSR) of 10 dB. The CSR has been reset to
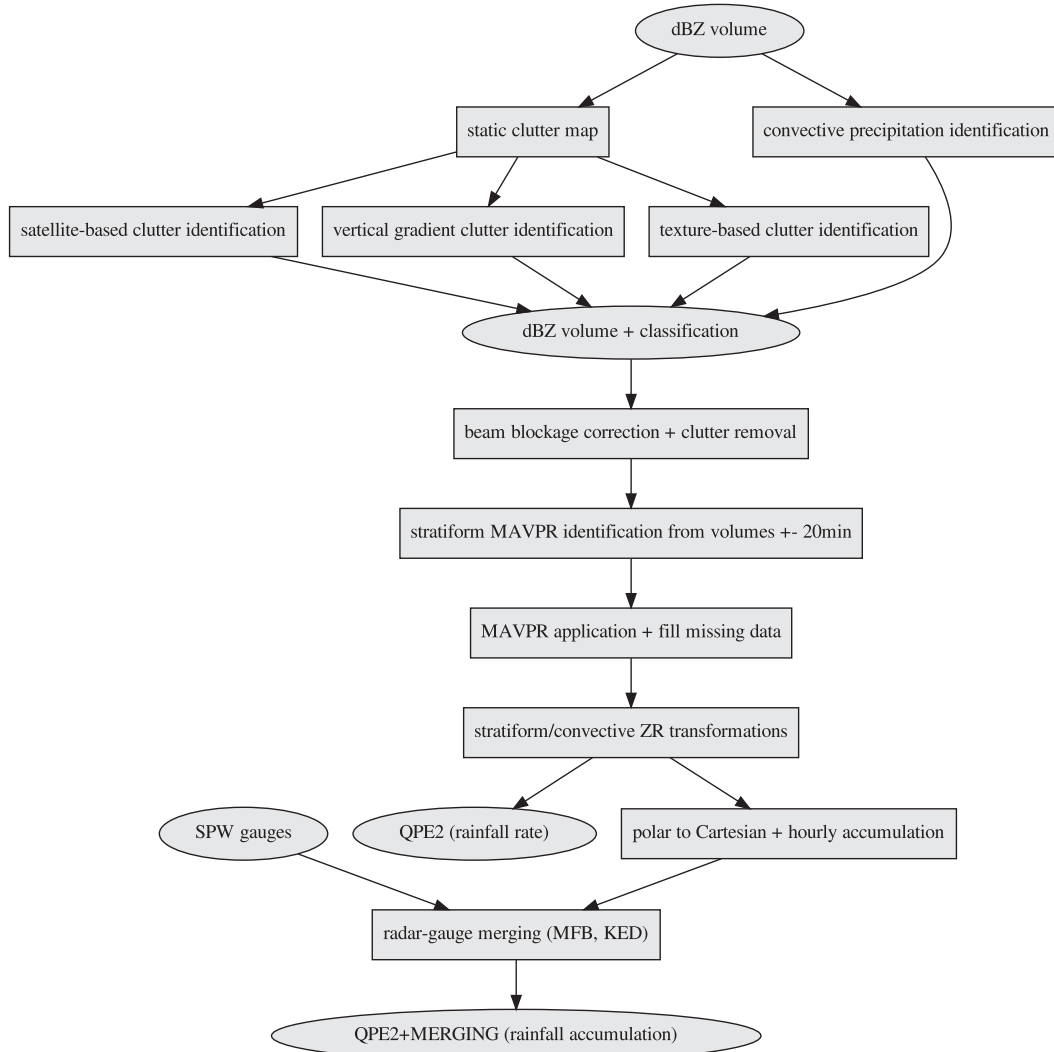
FIG. 1. Flowchart of QPE2: ellipses represent data and rectangles represent processing steps.

the default value of 15 dB during the 2013 update, resulting in fewer values set to zero. No further postprocessing corrections have been performed on the volumetric data. The radar data availability is about 96% for the 2005–15 period. Since the data are coded with a resolution of 0.5 dBZ, a random uniform field (−0.25 dBZ, +0.25 dBZ) is added to simulate actual reflectivity values. This allows for generating a smooth averaged empirical profile of reflectivity when the median is used as the averaging statistic.

## b. Beam blockage correction

To compute beam blockage, we consider the trajectory of the beam in normal propagation conditions (i.e., using a 4/3 factor for the earth radius). The Shuttle Radar Topography Mission (SRTM) data (Farr et al. 2007) are used as a digital elevation model (DEM) and have a resolution of 3 arc s (90 m). The beam blockage has been estimated using a formula proposed by Bech et al. (2007) that assumes a constant ground elevation across the beam. To simplify the computation, the DEM is interpolated on a 300-m grid, which is the resolution of the beamwidth at 10 km from the radar, where the first blockage occurs. If the blockage occurs close to the radar, a high accuracy of the height of the radar is required. It was found that using the Belgian reference for the radar height (592 m) results in a significant underestimation of the closest blockage. A correction of about 15 m was applied to this height to match the World Geodetic System used by SRTM. Because of the relatively low elevation angle (0.3°), blockages up to 18% were found for several narrow sectors (Fig. 3). Uncertainties in elevation angles, beam propagation, DEM measurements, and blockage computation method can
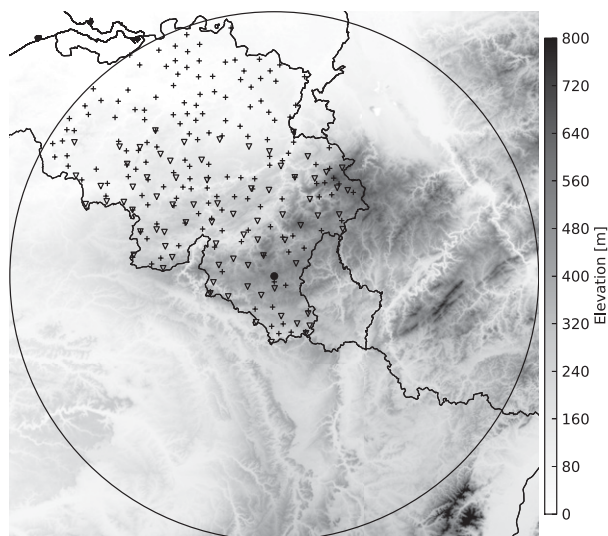
FIG. 2. Elevation map centered at the Wideumont radar (black dot) up to 180 km range (circle) with SPW (triangle) and CLIM (plus signs) rain gauge networks (gauges removed because of clustering are not shown). Country borders with France, Luxembourg, Germany, and the Netherlands are also displayed.

lead to under- or overestimation of the actual beam blockage. The data are corrected for beam blockage using a one-way correction. A smoothing of 1° is performed to take into account sample averaging in azimuth.

## c. Clutter identification

Weather radar measurements can be contaminated by clutter (i.e., nonmeteorological signals). In the atmosphere, airplanes, birds, and insects are a dynamical source of clutter. Nevertheless, the main source of clutter comes from the ground, which is hit by the side lobes of the beam (or even its main lobe). Ground clutter occurs mainly at elevated places (e.g., hills) surrounding the radar. During abnormal propagation (AP) conditions, the lowest part of the beam can bend progressively toward the ground and clutter can be found at any distance from the radar. The identification and mitigation of clutter remains a challenging problem in radar meteorology because it can be difficult to distinguish from precipitation. Conventional radars use Doppler filtering to remove near-zero velocity pulse echoes. A statistical filter can also be used based on pulse-to-pulse fluctuations. However, those filters are not perfect, and too aggressive settings can partly eliminate precipitation. Since pulse data are usually not sent to the data center, postprocessing of the radar moments (based on several pulses) is usually performed for remaining clutter. A summary of clutter mitigation techniques can be found in Hubbert et al. (2009), who recommend the fuzzy logic approach for its simplicity

and practicality. This technique uses a combination of features and probability functions to make a clutter identification decision. For example, Berenguer et al. (2006) combine three statistics from 3D data: shallow vertical extent, high spatial variability, and low radial velocities. In this study we combine satellite cloud-free echoes, strong vertical gradient, and high spatial variability in a deterministic fashion. This corresponds to step probability functions in the fuzzy logic framework with a probability of one associated to clutter. Eventually, a measurement is considered as clutter if at least one feature is equal to one.

### 1) PERMANENT CLUTTER

Radar measurements are contaminated by static ground clutter even after Doppler filtering. Topographic features such as hills are permanent while others (e.g., trees, towers, buildings) might evolve in time because of human activities. Static clutter can be identified using detection probability maps over a given period. In an operational context, this map should be updated regularly (i.e., each month) to take into account new sources of static clutter (e.g., wind farms). The bins whose measurements exceed 7 dBZ more than 50% of the time are identified as static clutter. The threshold is chosen to match the threshold used for "no precipitation" in subsequent algorithms. This probability threshold is chosen to be significantly higher than the expected monthly maximum rainfall occurrence in Belgium. To compute this map, using unfiltered data is simple and robust. Using filtered (e.g., Doppler) data is more difficult but allows keeping meteorological information for the slightly contaminated bins. For the Wideumont radar, no choice was possible since only filtered data have been archived. More problematic is the fact that bins exceeding a defined CSR have been set to "undetected" without a flag and are therefore wrongly interpreted as zeros. To solve this problem, radar data without CSR thresholding have been generated during one day without rain in 2013 and used to compute the static clutter map (Fig. 3). In the lowest elevation, about 10% of static clutter is found in the range 0–70 km. The static clutter map is applied before using the dynamical algorithms to improve their performance.

### 2) SATELLITE-BASED IDENTIFICATION

Since 2005 a cloud type classification (Derrien and Le Gléau 2005) with a resolution of 5 km is available each 15 min from the Nowcasting Satellite Application Facility (NWC SAF) products based on Meteosat Second Generation (MSG) satellite measurements. Radar echoes in areas classified as "sea" or "ground" by the satellite are identified as clutter. Areas classified as snow or
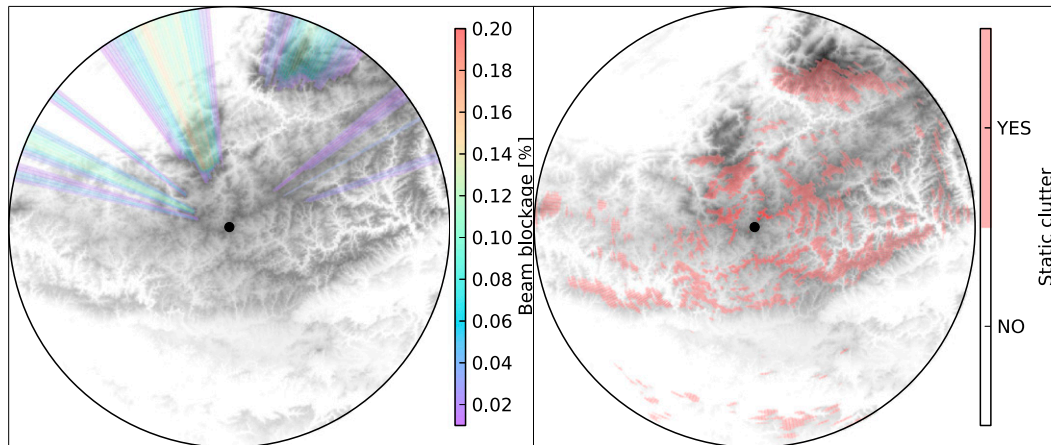
FIG. 3. Static quality information for the lowest elevation (0.3°) scan up to 50 km: (left) beam blockage and (right) clutter.

nonprecipitating clouds could also be identified as clutter. Since areas with actual rain can be misclassified in these classes, they are not used here. It is interesting to note that a parallax correction would have been required for nonprecipitating cloud areas. The time lag between radar and satellite observations (up to 8 min) is taken into account. An upper bound on the steering wind, estimated from the ground wind, is used to reduce the potential cloud-free areas. The clutter identification based on satellite is very robust but limited to well-defined cloud-free areas. It is particularly useful when clutter is difficult to distinguish from rainfall (e.g., in strong anomalous propagation conditions or at very large distance from the radar). Unfortunately, the availability of derived satellite products at RMIB was limited to 90% during the study period. Using a more complete dataset could help to remove even more clutter.

### 3) VERTICAL GRADIENT

Algorithms based on the vertical profile of reflectivity have been proposed in several papers (Steiner and Smith 2002; Berenguer et al. 2006). In case of precipitation, clutter can be identified by unrealistic vertical reflectivity gradients between two measurements at different elevations. A measurement at a given elevation is considered as clutter if the gradient between its value and the corresponding (horizontally) interpolated value on a higher (lower) elevation exceeds in magnitude $-20\,\mathrm{dB}Z\,\mathrm{km}^{-1}$ ($+10\,\mathrm{dB}Z\,\mathrm{km}^{-1}$). Because of variations from signal fluctuations, a minimum absolute difference of $5\,\mathrm{dB}Z$ between two corresponding values at different elevations is required for clutter identification. For this particular algorithm, all measurements below $10\,\mathrm{dB}Z$ are set to $10\,\mathrm{dB}Z$ since gradients are not

meaningful for low reflectivity values. Wrong clutter identification may occur in the bright band (BB) where high vertical gradient can be found. Therefore, the BB is identified using the radar data or constructed from the estimated freezing level. The freezing level is estimated based on the extrapolation (1°C per 150 m) of the mean surface temperature measured by nearby automatic weather stations. Extra (e.g., model) information on the freezing level is not used because of technical limitations. In the BB, a safety margin is applied with higher thresholds ($-60$ and $40\,\mathrm{dB}Z\,\mathrm{km}^{-1}$, respectively) for clutter identification. Since measurements from different elevations are not simultaneous (there is a lag of 30 s between successive elevations), a smoothing distance is computed based on the estimated steering wind and applied to the data. An upper bound for the steering wind is computed as three times the wind measurements at 10 m averaged over 10 stations. This algorithm is useful for clutter inside precipitation regions that cannot be easily identified by looking at the reflectivity texture.

### 4) TEXTURE

We use a simple and fast texture-based technique originally developed by Gabella and Notarpietro (2002) for Cartesian data. It consists of a two-part identification algorithm using 1) echo continuity and 2) minimum echo area. The algorithm has been improved and made applicable to polar data. For the first part, each bin is compared with a neighborhood that is selected based on a maximum distance. Since finding a circular neighborhood is computationally intensive, it is approximated by a square window. If the proportion of neighbors with similar values is too small, the bin is considered as clutter. In this study, two bins are similar if their difference is less than 7 dBZ. The proportion of similar bins

in the predefined neighborhood has to exceed 30% for the bin to be identified as nonclutter. The maximum distance defining the square window is the most important parameter of the algorithm. It should be chosen close to the smallest precipitation scale. The default value of 1 km used in this study should be interpreted as a lower bound for convective storm size. This corresponds to a window of 27 bins for a range of 100 km and 9 bins for a range of 200 km, where only the radial information is used. Using the proportion of similar values instead of counting their total number allows dealing with bins that have no values (e.g., after the application of the static clutter map). For the second part (echo area), too thin contiguous echo areas (with values above 0 dB$Z$) are considered as clutter. An echo area is too thin if more than 90% of its bins belong to its border. This second part is not applied when the bin azimuthal resolution is larger than the precipitation scale (within the range of 70 km). This texture-based technique can identify many different kinds of clutter. It is particularly useful for AP clutter in the lowest elevation and at distances where vertical information is limited.

## 3. Surface rainfall estimation

### a. Radar-based estimation

The estimation of rainfall rate at the ground can be derived from radar reflectivity measurements. Because of the earth curvature and the positive elevation angles, these measurements are obtained at heights increasing with range. It is well known that the VPR depends on the variations of hydrometeors' phase and size distribution. The most striking effect is the BB, which is an increase of reflectivity caused by melting snow in a layer under the freezing level. Another important effect is the increasing underestimation with range due to the lower reflectivity of snow and partial overshooting. Correction for the VPR has been an active topic in radar meteorology for decades. Parameterized models can be fitted to observations using inverse methods (Vignal et al. 2003), but the most straightforward approach is to derive averaged empirical profiles. The profile is best estimated using data at close range to the radar to limit the beam broadening effect. Germann and Joss (2002) compute the unconditional mean rain rate up to 70 km while Bellon et al. (2007) compute a reflectivity average over 20-km range intervals. Some authors stress the importance of distinguishing different types of precipitation (Kirstetter et al. 2010). For some cases, a local VPR correction should be preferred (Kirstetter et al. 2010; Zhang and Qi 2010; Hazenberg et al. 2013), but this has a computational cost. A hybrid approach

combining empirical and modeled profiles has been recently proposed by Koistinen and Pohjola (2014). At the RMIB, a VPR correction has been proposed (Goudenhoofdt and Delobbe 2012) based on a global average of normalized local profiles from data in the 0–70-km range. This paper presents a new version of the algorithm with a focus on robustness.

### 1) CONVECTIVE RAINFALL IDENTIFICATION

For convective precipitation a uniform VPR is assumed and no correction is applied. Since their VPRs are different, we first distinguish between stratiform and convective precipitation. Following the Steiner algorithm discussed in Biggerstaff and Listemaa (2000), a reflectivity value is considered as convective if it exceeds 40 dB$Z$ or if it is significantly higher than its local background (defined by a search radius of 11 km). In a last step, convective areas are expanded by range-dependent smoothing. Values below 10 dB$Z$ are set to 10 dB$Z$ for proper background computation (only for this specific task). The original algorithm has been adapted to be applied on polar data with possible missing data. Another limitation was the false detection of the bright band as convective precipitation. To consider a point as convective, we require that the value from the lowest elevation angle and the corresponding interpolated value at 3000 m above the radar level are both identified as convective. No convective identification is made before the highest elevation angle reaches 1500 m, which corresponds to a range of 15 km.

### 2) AVERAGE PROFILE ESTIMATION

Assuming that the VPR does not vary in space, the mean apparent VPR (MAVPR) is computed for each height interval as the median of reflectivity values considered as stratiform (i.e., not convective and exceeding 10 dB$Z$). Only data at close range are used to limit the beam broadening effect. A maximum distance of 45 km, which corresponds already to an 800-m beamwidth, seems reasonable. A minimum distance of 5 km is chosen to avoid computing the lowest part of the VPR on a limited area. A minimum vertical resolution of 50 m is chosen to match the vertical resolution of the beam at 5 km from the radar. Even if the probability of residual ground clutter is low, we prefer not to use the lowest elevation in the computation of the MAVPR to make our reanalysis robust. To support this choice, one notes that the vertical coverage of the lowest elevation is limited and that more useful data can be obtained from higher elevations. Using an unconditional average of reflectivity can be problematic in case of localized precipitation areas. The robustness of the profile depends on the amount of available measurements. As in Bellon

et al. (2007), we impose a minimal number (i.e., 1000) of measurements for each vertical interval. This is unfortunately not enough since it does not take into account the increasing sampling area per bin with range. Therefore, we use an additional criterion based on the total sampled area (which is set to 1000 km$^2$) for each vertical interval. To be representative, the MAVPR should also cover a sufficient vertical extent. The MAVPR is considered as valid if it is available between 500 m and 2 km above the radar. If the mean 7-dB$Z$ echo top in the selected area is lower than 2 km, it is taken as the upper limit required for a valid VPR. To increase the availability of a valid MAVPR, we use successive radar scans as in Hazenberg et al. (2013). The underlying assumption of VPR stationarity seems reasonable for an interval of 20 min. This corresponds to nine radar volumes using both previous and next data. One notes that in real-time conditions, data from the future are not available and then only five radar volumes can be used. The MAVPR (denoted $\mathbf{z_m}$) is checked for abnormal curvature by computing the second-order discrete difference dd2$_i$ for each vertical interval $i$:

$$\mathbf{dd}(i) = \mathbf{z_m}(i+1) - \mathbf{z_m}(i) \quad \text{and} \tag{1}$$

$$\mathbf{dd2}(i) = \mathbf{dd}(i+1) - \mathbf{dd}(i). \tag{2}$$

The MAVPR is not valid if the maximum dd2$_i$ (expected at the brightband peak) exceeds 10 dB$Z$ or if the number of dd2$_i$ higher than 1 dB$Z$ exceeds 5. If the MAVPR is not valid, the vertical resolution is increased by 50 m and the thresholds for dd2$_i$ adapted. This procedure is applied to a resolution not higher than 200 m to ensure that the bright band is properly resolved. If no valid MAVPR can be obtained, the closest one (up to 20 min) is used. In the worst case, a climatological profile with a constant slope ($-2$ dB$Z$ km$^{-1}$) above the freezing level is used. The climatological profile is uniform below the freezing level.

### 3) BRIGHTBAND IDENTIFICATION

The identification of the BB is important for the extrapolation of the MAVPR. Because of beam broadening with increasing range, the BB depth increases and the BB peak decreases. The peak of the BB is identified by the maximum of the MAVPR. It is checked that the peak is not a spike, that a strong gradient exists below, that a very strong negative gradient exists above, and that the peak value is 3 dB$Z$ larger than the values at $\pm$500 m (at least one value is needed). Starting from the peak, the bottom of the BB is identified by the first gradient lower than 3 dB$Z$ km$^{-1}$ and at least 300 m below. Similarly, the top of the BB is identified by the first gradient exceeding $-5$ dB$Z$ km$^{-1}$ and at least 300 m above. If no lower or upper bound for the BB can be found (i.e., when no profile value is available), the MAVPR is extrapolated linearly to $\pm$500 m by $-5$ and $-8$ dB$Z$, respectively. If no BB can be identified, the freezing level is estimated as in the clutter identification. If the top of the MAVPR is above the freezing level, its value is extrapolated aloft using a constant gradient of $-2$ dB$Z$ km$^{-1}$. The MAVPR is finally extrapolated (uniformly) to the ground using the value at the lowest height.

### 4) MAVPR APPLICATION

The ratio between the median profile $Z_m(h)$ and the apparent profile at a given range $Z_a(r, h)$ is applied to the measured equivalent reflectivity factor $Z_e(h)$ to obtain the surface reflectivity $Z_e(h = 0)$:

$$Z_e(h = 0) = Z_e(h) \frac{Z_m(h = 0)}{Z_a(r, h)}, \tag{3}$$

$$z = 10 \log_{10}(Z), \quad \text{and} \tag{4}$$

$$z_e(h = 0) = z_e(h) + z_m(h = 0) - z_a(r, h), \tag{5}$$

where $h$ is the height above the radar and $r$ is the distance to the radar (note that quantities $Z_e$, $Z_m$, and $Z_a$ are in units of mm$^6$ m$^{-3}$, and the corresponding $z_e$, $z_m$, and $z_a$ are expressed on a decibel scale with units in dB$Z$). The quantity $z_m(h)$ is estimated at any height from the discrete profile $\mathbf{z_m}(i)$ (MAVPR) using linear interpolation (decibel scale). The quantity $Z_a(r, h)$ is obtained by the convolution of the median profile $Z_m(h)$ with the two-way normalized power gain of the beam at distance $r$. Since $Z_m(h)$ is based on data at a given range interval, the convolution is started in the middle of this interval. A more accurate but costly computation of the convolution (Kirstetter et al. 2010) is not expected to significantly improve the results. According to Joss and Lee (1995), the correction factor should be limited to avoid instability. Therefore, after a sensitivity test, a maximum correction factor of 10 dB$Z$ has been chosen. The height of the radar is taken as the ground reference for the extrapolation of the VPR. The case when the radar is in the bright band is problematic since finding a good $Z$–$R$ relationship for melting snow can be complicated. We solve this problem by using a theoretical VPR profile in order to obtain a reflectivity value corresponding to rain. In practice we use the extrapolation procedure described above to find the corresponding value at the artificial lower limit of the bright band. The estimation of reflectivity at the ground reference is applied for each elevation. Using a weighted (e.g., based on distance) average of several elevations to mitigate residual errors and reflectivity measurement uncertainty is common. However, we think that after quality control the estimation from the lowest elevation is much better

than from higher elevation. Therefore, estimations from higher elevations are only used to fill gaps caused by the clutter removal. For this purpose, the data from higher elevations are bilinearly interpolated in a range to match the lowest elevation. Remaining gaps (i.e., when no precipitation information is found in the higher elevation) in the grid are interpolated using a nearest-neighbor technique up to a 2-km maximum distance.

### 5) Z–R RELATIONSHIP

The Marshall–Palmer relationship ($Z = 200R^{1.6}$) has been used operationally since the beginning of the radar measurements. For the reanalysis a more refined relationship from the Radar-Online-Aneichung (RADOLAN) product (Wagner et al. 2012) of the German weather service has been considered:

$$Z = 200R^{1.6} \quad \text{if} \quad \text{dB}Z \leq 44 \quad \text{and} \tag{6}$$

$$Z = 77R^{1.9} \quad \text{if} \quad \text{dB}Z > 44. \tag{7}$$

For high reflectivity, the rain rate obtained with the latter is a bit lower. To deal with hail, a maximum reflectivity of $55\,\text{dB}Z$ is used, which corresponds to $88\,\text{mm}\,\text{h}^{-1}$. To avoid remaining artifacts (e.g., insects), reflectivity values below $7\,\text{dB}Z$ (i.e., $0.1\,\text{mm}\,\text{h}^{-1}$) are considered as no precipitation.

### b. Radar–gauge merging

In a final stage, radar estimates can be combined with rain gauge measurements. A large variety of radar–gauge merging methods have been tested in Goudenhoofdt and Delobbe (2009) at the daily scale. The results showed that a simple mean field bias correction (MFB) reduces the mean absolute error significantly and that kriging with external drift (KED) performs best. Cokriging (Krajewski 1987) is an attractive alternative method, but its computational cost is too high to be considered in this study. Sideris et al. (2013) proposed to add temporal information to KED via cokriging. Because of the added complexity, this method is also not used here. The benefit of using hourly over daily adjustment has been shown by Thorndahl et al. (2014) and Berndt et al. (2014). In this study both MFB and KED methods are used on hourly Cartesian data. For a given gauge, the corresponding radar-based estimate is the one of the pixel where the gauge is located. Averaging over several pixels does not improve the correlation in most cases.

### 1) RAIN GAUGE MEASUREMENTS

The hydrological service of the Walloon region (SPW) operates a dense (one gauge per $135\,\text{km}^2$) and integrated

network of 90 telemetric rain gauges (Fig. 2). Most of them are tipping-bucket systems providing hourly rainfall accumulation. The collected data are used for hydrological modeling and directly sent to RMIB. The rain gauges are controlled on site every 3 months and in a specialized workshop every year. Every day, a quality control of the data is performed by RMIB using a comparison with neighboring stations. Radar data are also used in this quality control for the elimination of outliers. The classification of gauge data for the period 2005–15 are 94.5% validated, 2.7% corrected, 2.3% classified as dry snow, and 0.5% unclassified. Only validated data are used for combination with radar data. However, the quality control is not perfect since errors in classification have been found in dry snow and convective situations by comparison with the climatological network. To make the merging methods more robust, the clustering of the rain gauge network is reduced by removing a few gauges.

### 2) RADAR RAINFALL RATE ACCUMULATION

Before accumulation, rainfall rate values from the polar grid are interpolated on a Cartesian grid of 500-m resolution. The interpolation is done using a uniform 500-m square window filter on the polar data followed by a nearest-neighbor interpolation. This method is more accurate than an interpolation of points because it takes into account the actual areas of the polar bins and the Cartesian pixels. Rainfall accumulation over a given period is obtained by linear interpolation of the radar rain rates. No correction of temporal sampling errors (e.g., using optical flow) is performed because their occurrence is relatively low with a time step of 5 min. Indeed, small convective storms moving at a high speed are relatively rare (Goudenhoofdt and Delobbe 2013). Using Cartesian instead of polar data allows for a better correlation between gauge and radar estimation by reducing time and space sampling differences, especially at close range.

### 3) MEAN FIELD BIAS CORRECTION

The assumption here is that the radar estimates are affected by a uniform multiplicative error. This error can be due to a bad electronic calibration or an erroneous coefficient $a$ in the $Z = aR^b$ relationship. For each time interval, the adjustment factor is computed as the median of the ratios of gauge and radar values. The adjustment is valid if there are at least 11 pairs with both values exceeding 0.2 mm. Because of a potential decrease (at least for QPE1) of radar rainfall estimates caused by partial overshooting farther than a certain distance, only gauges up to 100 km are used.
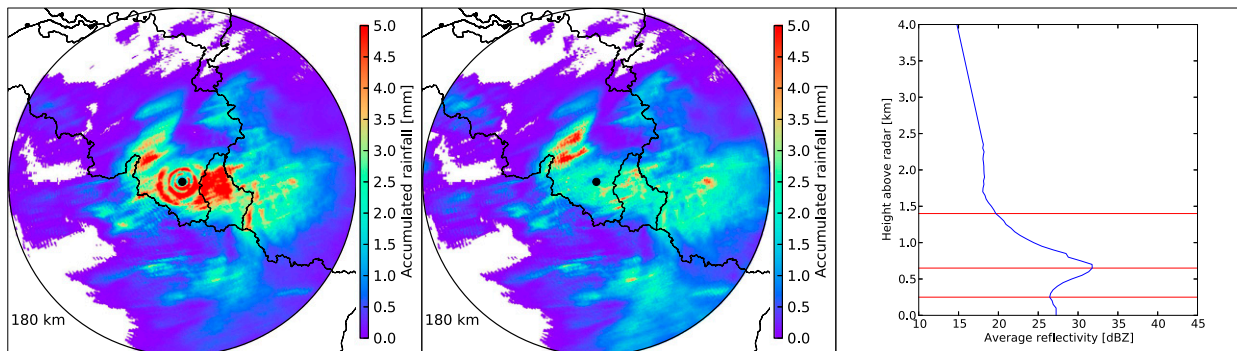
FIG. 4. Stratiform case on 12 Apr 2013: hourly accumulation from 0000 to 0100 UTC based on (left) QPE1 or (center) QPE2. (right) Median averaged vertical profile of reflectivity based on nine polar volumes; the brightband peak and limits are identified by lines.

#### 4) KRIGING WITH EXTERNAL DRIFT

This is a geostatistical method that combines the rain gauge values linearly and uses the radar as auxiliary information. It follows the same scheme as the ordinary kriging, except that the expected value of the estimated precipitation field is now considered as a linear function of the radar field (the drift). The weights are obtained by minimizing the estimation error that depends on the spatial correlation between radar and gauge values and their residuals. As proposed by Erdin et al. (2012), a square-root transformation is applied to precipitation values to approach Gaussianity, which is an underlying assumption in kriging. Determining a suitable variogram for the residuals between the estimated value and the drift is challenging. The underlying assumption is that the residuals are correlated in space. A robust method is proposed by Schiemann et al. (2011), while Delrieu et al. (2014) suggest deriving event-based variograms. For the sake of simplicity, we use an exponential climatological variogram with a nugget of 0.1 and a range of 10 km. It is not easy to determine the decorrelation distance of the residuals. We prefer not to make assumptions on the correlation between the residuals beyond 10 km, since QPE2 might have removed the large-scale errors. The stability of the method depends on the validity of rain gauge measurements and also the collocated radar estimates. Therefore, no gauge locations are used beyond 160 km because of potentially severe underestimation by the radar. Additionally, a minimum correlation (Pearson coefficient) of 0.5 is required between the two estimates. The performance of this method decreases significantly with the distance to the rain gauge network.

### 4. Evaluation

#### a. Rain gauge measurements

RMIB maintains a climatological network (CLIM) including 270 stations with daily measurements of precipitation accumulation between 0800 and 2000 local time (LT). From Fig. 2, one notes that the coverage of CLIM is larger than the SPW network and that only a few locations belong to both networks. Most of these stations are manual and the data are generally available with some delay. The data are manually inspected on a monthly basis by well-trained operators. By arraying observers' records geographically in time sequence, pattern analysis performed by the trained staff can reveal inconsistencies or anomalies. Pattern analysis will be used here to evaluate and compare the performance of the QPE methods. As for the merging network, a declustering is applied to reduce the influence of areas with high rain gauge density on the results. The hourly accumulation obtained by the merging methods are summed to match the 24-h accumulation of CLIM.

#### b. Examples

In Fig. 4, a typical winter case with stratiform precipitation is shown. A bright band appears in the MAVPR between 600 and 1600 m. If we compare QPE2 (new) with QPE1, we see that the rings due to the bright band have been significantly reduced. The underestimation at long range is also reduced.

In summer (Fig. 5), the distinction between convective and stratiform precipitation plays an important role. Indeed, for the identified convective precipitation on each volume, no profile correction is performed. Furthermore, because of the limited area covered by stratiform precipitation, a climatological profile is used. There is a slight effect of decreasing the rainfall rate by using a specific $Z$–$R$ relationship for convective precipitation in QPE2, but the main effect in this case is a significant decrease of the rainfall rates with QPE2 due to the maximum (hail) threshold (55 dB$Z$).

#### c. Case validation

For the first six months of 2013, 13 days with different rainfall situations have been chosen and sorted in
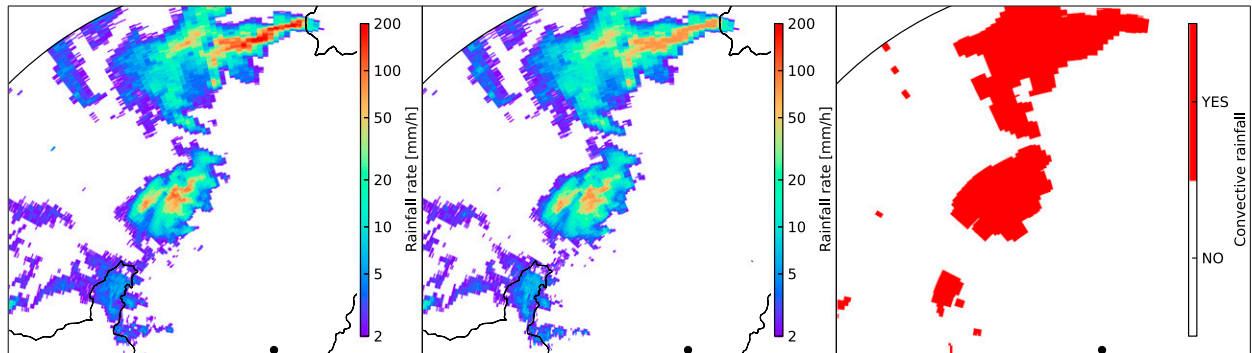
FIG. 5. Convective case on 19 Jun 2013: rainfall rate estimation at 1905 UTC for (left) QPE1 and (center) QPE2 and (right) corresponding identified convective area.

chronological order. Stratiform precipitation (cases 1 ,4, 6, 9, and 10), a mix of rain and snow precipitation (cases 2, 3, 5, 7, and 8), and convective precipitation (cases 11, 12, and 13) were observed. The various algorithms of QPE2 (and especially the VPR correction) have been fine-tuned on those cases. A categorical verification (Wilks 1995) that hourly values exceed 0.1 mm against the SPW network is reported in Table 1. The scores are averaged over the 13 cases, plus a few cases with little or no rain. It shows a decrease of the false alarm ratio (FAR) and an increase of the probability of detection (POD) with QPE2. This highlights the effect of the mitigation of clutter and correction of underestimation at long range, respectively. The high values of FAR are due to some cases with gauges covered by dry snow but not flagged by the quality control. A lower FAR and higher POD both increase the critical success index (CSI), which measures the global performance. The quantitative performance of the algorithms is analyzed using three different scores. The mean absolute difference (MAD) measures the average additive error between the radar estimate $R$ and the gauge measurement $G$:

$$\text{MAD} = \text{mean}|R_i - G_i|. \qquad (8)$$

This score is strongly influenced by heavy rainfall and outliers. It is therefore a good indicator of the robustness of the method. The scatter score (SCS; Germann et al. 2006) measures the multiplicative error spread:

$$\text{SCS} = r_b - r_a, \qquad (9)$$

$$r_i = \frac{R_i}{G_i} \quad \text{where} \quad r_i \geq r_{i-1}, \quad \text{and} \qquad (10)$$

$$c_i = \sum_1^i G_i \Big/ 1 \sum^n G_i \quad \text{where} \quad c_a \approx 0.16 \quad \text{and} \quad c_b \approx 0.84, \qquad (11)$$

where $n$ is the number of radar–gauge pairs and $r_i$ are the ratios between radar and gauge estimates sorted in increasing order. The ratios are expressed on a decibel scale and set to $-20\,\text{dB}$ if the radar estimation is zero. The indices $a$ and $b$ correspond to the normalized cumulative sum of gauge values (defined as $c_i$) closest to 0.16 and 0.84. The fraction of radar estimates with less than 20% error $F_{20}$ measures the number of good estimates:

$$F_{20} = \frac{\sum_0^n \delta_i}{n}, \qquad (12)$$

$$\delta_i = 1 \quad \text{if} \quad 0.8 \leq r_i \leq 1.25, \qquad (13)$$

$$\delta_i = 0 \quad \text{otherwise}, \qquad (14)$$

where $n$ and $r_i$ are defined above. This score allows a tolerance for the sampling difference between the radar and the gauge. It is worth pointing out that the SCS and $F_{20}$ are both immune to outliers. In Fig. 6, one can see the relative performance of the quantitative precipitation estimates (QPE1 and QPE2) without merging [original radar data (ORI)] or with a merging method (MFB, KED) against the 24-h CLIM up to 180 km. The scores are computed for each case on a set of radar–gauge pairs and then averaged. We do not compute the scores using all radar–gauge pairs at once because it would favor cases with widespread

TABLE 1. POD, FAR, and CSI for hourly accumulation exceeding 0.1 mm against the SPW network. The scores are computed for all hours of the cases and then averaged.

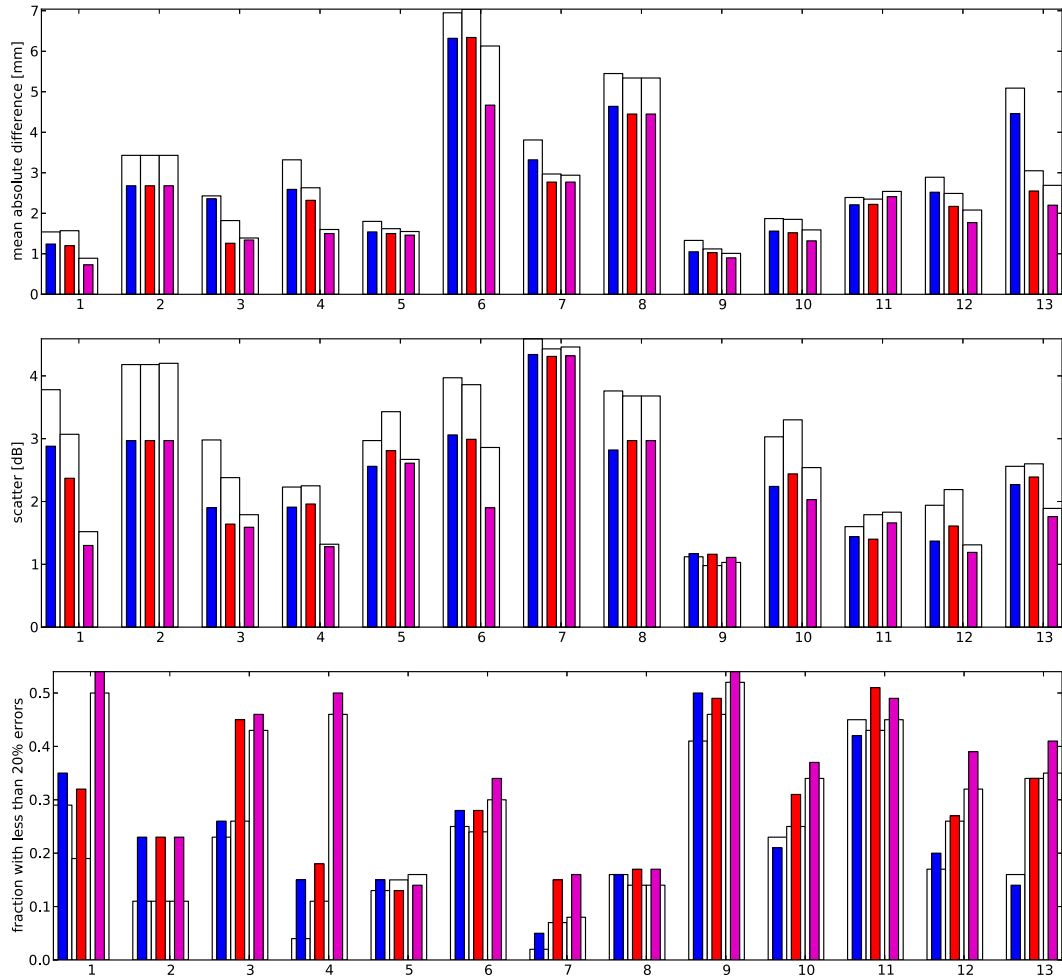| | POD | FAR | CSI |
|---|---|---|---|
| QPE1 | 0.601 | 0.550 | 0.322 |
| QPE2 | 0.632 | 0.512 | 0.363 |

FIG. 6. Validation of daily accumulation against CLIM for the 13 cases: (top) MAD, (middle) SCS, and (bottom) $F_{20}$. For each numbered case the bars correspond (from left to right) to ORI (blue), MFB (red), and KED (magenta) merging methods. The white and colored bars correspond to QPE1 and QPE2, respectively.

precipitation. As in the merging process, the corresponding radar-based estimate is chosen at the pixel where the gauge is located. To properly compare the methods, we use a fixed set of gauge values exceeding 1 mm. One first notes that the MAD is lower for QPE2 than QPE1 for all cases and merging methods. This is also true for SCS and $F_{20}$ in most cases. The benefit of QPE2 is higher for the stratiform cases than for convective or snow/rain mix cases. The benefit of QPE2 is reduced after merging for most cases, especially with KED. It could be explained by the fact that the merging is able to mitigate part of the error corrected by QPE2. The MFB method has little effect on SCS since only the bias is corrected. The $F_{20}$ is clearly improved using QPE2 and the merging methods for most cases and especially for stratiform cases.

Table 2 reports the averaged statistics over all cases for QPE1 and QPE2 and four additional estimates to study the effect of the processing:

- PPI1: the uncorrected lowest elevation angle (PPI);
- PPI2: the lowest PPI where identified clutter have been replaced by values from higher elevations;
- PPI3: PPI2 with beam blockage correction; and
- QPE2$_{MP}$: QPE2 using the Marshall–Palmer $Z$–$R$ relationship.

As expected, PPI1 performs the worst for all scores and merging methods. For all merging methods, PPI2 performs slightly better than QPE1 for the MAD score but slightly worse for SCS and $F_{20}$. A slight benefit of the beam blockage (PPI3) is obtained after merging for SCS and $F_{20}$. The significant improvement of QPE2

TABLE 2. Verification of daily accumulation against CLIM for gauges up to 180 km and values above 1 mm. The results of QPE1, QPE2, and four intermediate steps in combination of the merging methods are shown. The scores are averaged over 13 cases with more than five valid pairs.

|  | QPE1 | PPI1 | PPI2 | PPI3 | QPE2$_{MP}$ | QPE2 |
|---|---|---|---|---|---|---|
| MAD (mm) | | | | | | |
| ORI | 3.272 | 3.325 | 3.143 | 3.171 | 2.821 | 2.808 |
| MFB | 2.872 | 3.040 | 2.857 | 2.864 | 2.478 | 2.465 |
| KED | 2.546 | 2.793 | 2.485 | 2.488 | 2.176 | 2.170 |
| SCS (dB) | | | | | | |
| ORI | 2.980 | 3.196 | 3.173 | 3.180 | 2.409 | 2.391 |
| MFB | 2.938 | 3.113 | 3.137 | 3.109 | 2.441 | 2.426 |
| KED | 2.406 | 2.712 | 2.488 | 2.497 | 2.056 | 2.056 |
| $F_{20}$ | | | | | | |
| ORI | 0.199 | 0.205 | 0.209 | 0.208 | 0.244 | 0.242 |
| MFB | 0.234 | 0.214 | 0.220 | 0.225 | 0.293 | 0.293 |
| KED | 0.323 | 0.280 | 0.311 | 0.317 | 0.364 | 0.364 |

TABLE 3. POD, FAR, and CSI for daily accumulation exceeding 1 mm against CLIM. The scores are computed for all days and then averaged over the 10 years.

|  | POD | | FAR | | CSI | |
|---|---|---|---|---|---|---|
|  | QPE1 | QPE2 | QPE1 | QPE2 | QPE1 | QPE2 |
| ORI | 0.663 | 0.704 | 0.365 | 0.323 | 0.466 | 0.517 |
| MFB | 0.671 | 0.710 | 0.364 | 0.322 | 0.475 | 0.525 |
| KED | 0.713 | 0.739 | 0.362 | 0.318 | 0.510 | 0.554 |

compared to PPI3 demonstrates the benefit of the VPR correction. The replacement of the Marshall–Palmer $Z$–$R$ relationship with an adapted $Z$–$R$ relationship has only a small positive impact on the averaged results. On average, QPE2 without merging (ORI) performs better than QPE1 + MFB, especially for SCS. QPE1 with KED gives a similar performance as QPE2 with MFB.

### d. 10-yr verification

The precipitation estimates have been verified using CLIM for the 2005–15 period. The results of the categorical verification (values exceeding 1 mm) are shown in Table 3. There is a positive impact of QPE2 with respect to QPE1 on POD because of the correction of underestimation at long range. The reduction of FAR suggests that clutter mitigation has played a role. The merging methods allow for improving the POD, but using KED reduces slightly the benefit of QPE2.

Figure 7 shows the scores computed at each rain gauge for values exceeding 1 mm. It is displayed as a function of the distance between the radar and the gauge. It is important to note that a perfect match with the gauge measurements is impossible since precipitation is estimated by the radar on a much bigger area (500 m × 500 m). It must be noted that a gauge with a very high MAD for QPE1 (due to ground clutter) has been removed to obtain a smooth average curve. The main result is that with QPE2 the MAD, SCS, and $F_{20}$ are improved for 98%, 92%, and 86% of the gauges compared to QPE1, respectively. The general behavior for QPE1 and QPE2 is a decrease of performance further than a certain distance because of partial and complete overshooting, respectively. The benefit of QPE2 with respect to QPE1 is particularly apparent at short and

long range. It results from the correction of the brightband effect and the correction of partial overshooting, respectively. The former result is more pronounced for SCS and $F_{20}$. It can be explained by the fact that the brightband error is highly variable. Another observation is that QPE2 has a relatively linear decrease of performance along the range for SCS and $F_{20}$. This could be explained by the increasing probability of attenuation by intense rainfall and hail, which is not corrected. There is little effect at long range for SCS when comparing QPE1 and QPE2 because the VPR correction (i.e., linear profile above the bright band) does not decrease the spread of the error. The effective range (i.e., where the MAD stays relatively constant) increases approximatively from 100 to 125 km, which corresponds to a 50% increase in areal coverage.

In Fig. 8, one can see the mean absolute error for the merging methods. The MFB allows us to improve both QPE1 and QPE2 significantly, except at long range. The relative benefit of QPE2 is slightly smaller in the midrange. The KED method allows us to further improve the performance of QPE1 and QPE2 but reduces the additional benefit of QPE2 at long range. The benefit of QPE2 at short range is enhanced after merging and is mainly due to the VPR correction. The results (not shown) are similar for the other scores. The average improvements from QPE1 to QPE2 + KED for gauges up to 150 km are 38%, 35%, and 80% for MAD, SCS, and $F_{20}$, respectively.

The scores averaged over all days with at least five gauge values exceeding a given threshold can be found in Table 4. Since the same weight is given for each day, the results are slightly different than the gauge-averaged statistics, which favor widespread situations. The range has been limited to 150 km, where the performance starts to drop significantly. Looking at gauge values exceeding 1 mm, QPE2 + KED reduces the MAD from 2.44 to 1.66 mm and the SCS from 2.26 to 1.95 dB while $F_{20}$ is increased from 0.22 to 0.37. The added value of the new processing increases with higher thresholds. At 10 mm the MAD decreases from 6.17 to 3.92 mm, the SCS decreases from 1.40 to 1.14 dB, and $F_{20}$ increases from 0.28 to 0.53. It should be noted that part of the
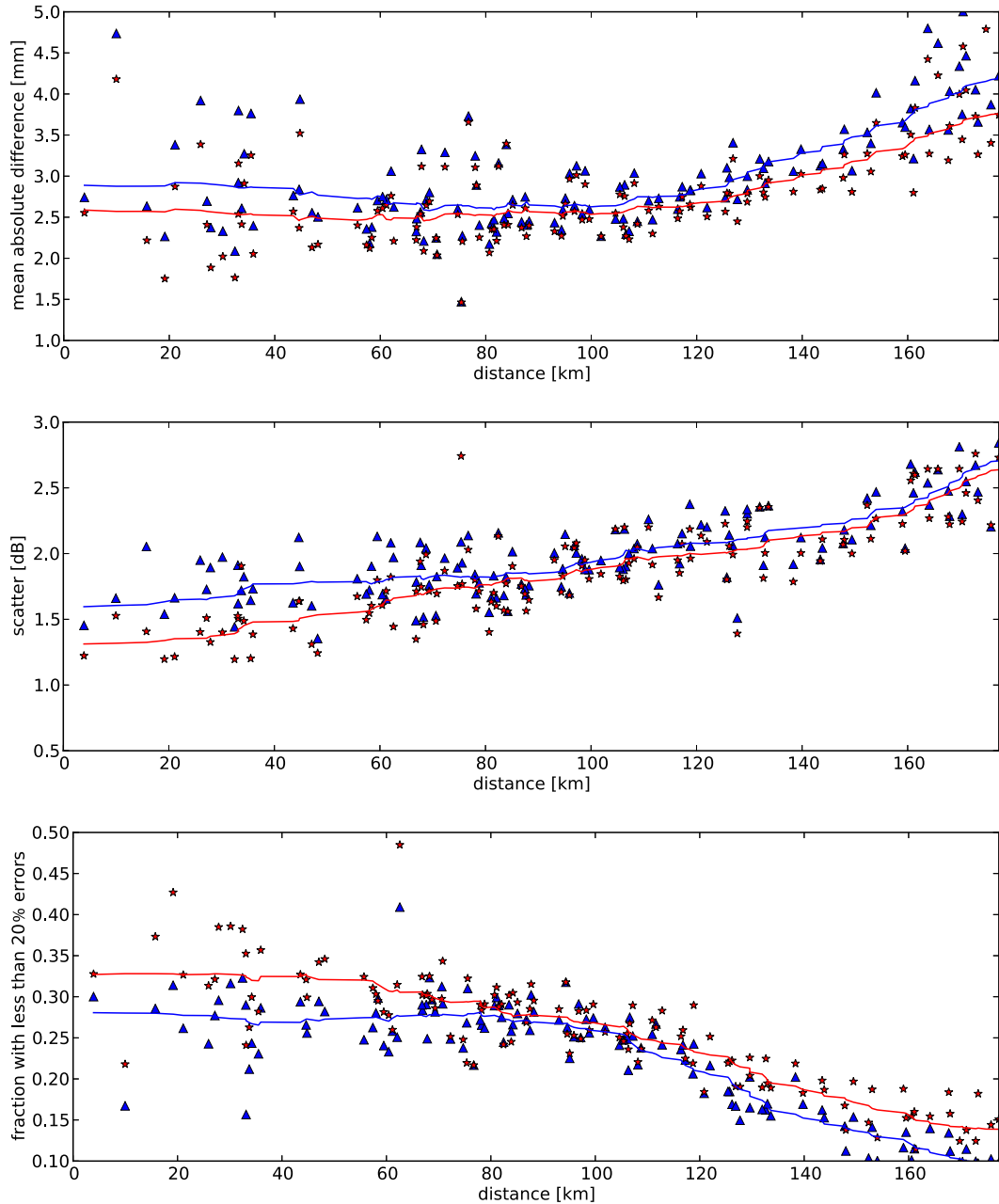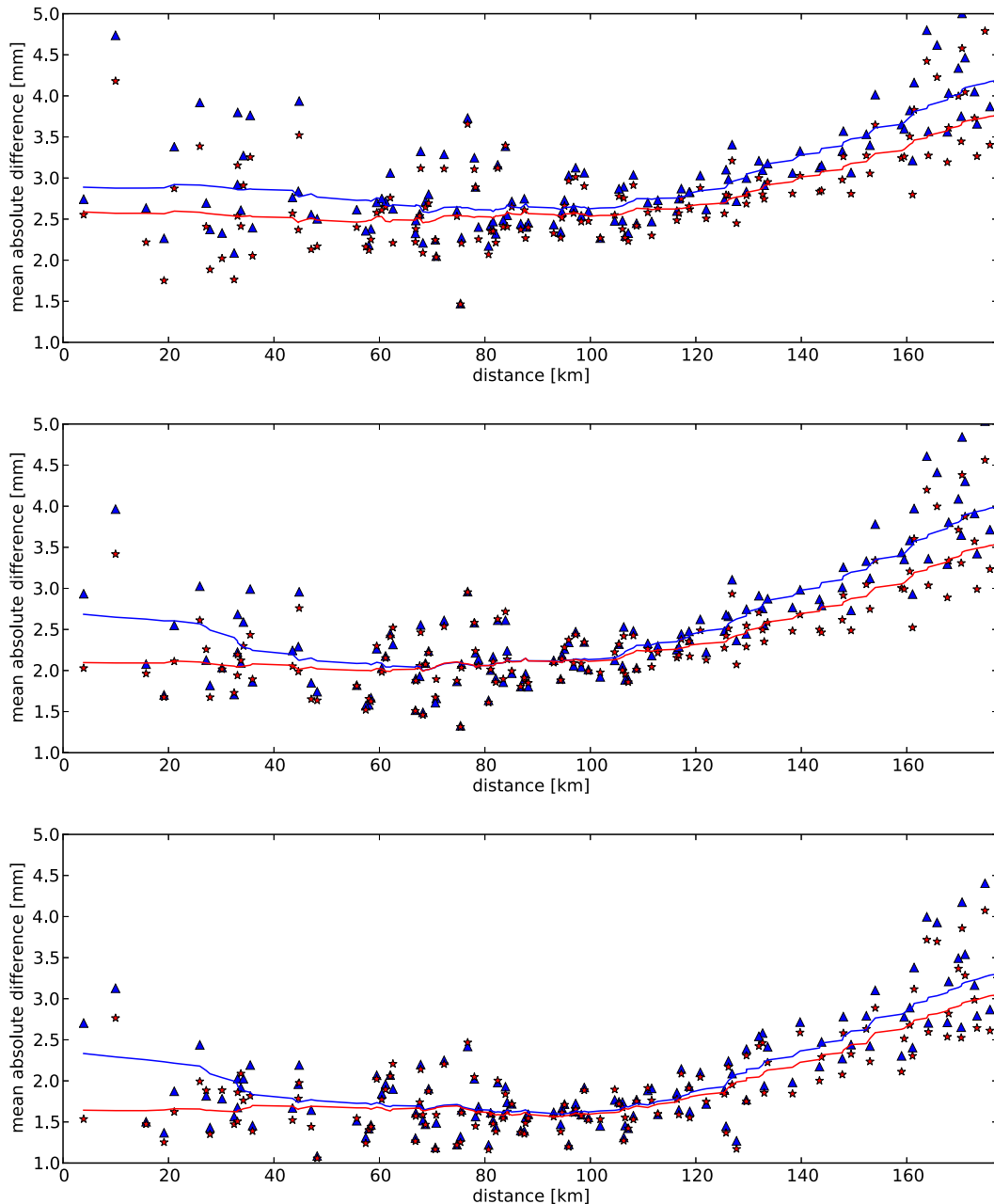
FIG. 7. The 2005–15 verification of daily accumulation against CLIM for QPE1 (blue triangles) and QPE2 (red stars) without merging. Only radar–gauge pairs with gauge value exceeding 1 mm are selected. Shown are (top) MAD, (middle) SCS, and (bottom) $F_{20}$.

performance of the KED method could be due to smoothing effects.

### e. Statistics

The global and local benefit of QPE2 and the merging methods can be seen by looking at Fig. 9, which shows the mean annual total for the 10-yr period. The full radar coverage is shown mainly to highlight the mitigation of

radar artifacts. It is important to note that the estimates outside Belgium have not been verified, since it is not the purpose of the paper. We can, however, expect good results given the relatively similar meteorological conditions within the radar coverage. The results of QPE2 + KED have little value outside Belgium but are presented for the sake of consistency with the other figures. It can be seen that the impact of clutter has been

FIG. 8. The 2005–15 verification of daily accumulation (MAD) against CLIM for QPE1 (blue triangles) and QPE2 (red stars). Only radar–gauge pairs with gauge values exceeding 1 mm are selected. Shown are the merging methods (top) ORI, (middle) MFB, and (bottom) KED.

significantly reduced. In particular, the small lines due to airplanes (mainly southwest of the radar from 200-km range) and the interference line (south of the radar) have been removed. However, there are still some limited areas contaminated by ground clutter that are slightly amplified by the merging methods. It is important to note that only a few gauges are located in areas with a higher probability of clutter. The underestimation

due to beam blockage is reduced but still visible. This can be at least partially explained by variations in propagation conditions. The effect of the bright band (i.e., concentric circles) on QPE1 is not very pronounced since it depends on the height of the freezing level and is therefore spread over the range. The smoothed circles are properly removed by the usage of the lowest elevation in QPE2. The underestimation at long

TABLE 4. Verification of daily accumulation against CLIM (up to 150 km). The scores are computed for all days with more than five valid pairs and then averaged over the 10 years. Valid pairs are those with a gauge value exceeding a threshold (mm) given in the table header.

| | 0 mm | | 1 mm | | 5 mm | | 10 mm | |
|---|---|---|---|---|---|---|---|---|
| | QPE1 | QPE2 | QPE1 | QPE2 | QPE1 | QPE2 | QPE1 | QPE2 |
| MAD (mm) | | | | | | | | |
| ORI | 1.263 | 1.170 | 2.437 | 2.274 | 4.100 | 3.731 | 6.170 | 5.574 |
| MFB | 1.104 | 1.038 | 2.107 | 2.002 | 3.513 | 3.252 | 5.157 | 4.751 |
| KED | 0.920 | 0.875 | 1.736 | 1.660 | 2.886 | 2.690 | 4.257 | 3.924 |
| SCS (dB) | | | | | | | | |
| ORI | 2.963 | 2.770 | 2.255 | 2.211 | 1.622 | 1.538 | 1.397 | 1.312 |
| MFB | 2.977 | 2.782 | 2.271 | 2.223 | 1.635 | 1.550 | 1.401 | 1.333 |
| KED | 2.922 | 2.661 | 2.011 | 1.953 | 1.397 | 1.331 | 1.205 | 1.138 |
| $F_{20}$ | | | | | | | | |
| ORI | 0.371 | 0.441 | 0.216 | 0.243 | 0.266 | 0.305 | 0.278 | 0.317 |
| MFB | 0.392 | 0.460 | 0.264 | 0.288 | 0.339 | 0.374 | 0.369 | 0.404 |
| KED | 0.405 | 0.481 | 0.351 | 0.370 | 0.451 | 0.479 | 0.496 | 0.526 |

ranges is reduced using QPE2 and even more using QPE2 + KED.

A comparison with mean annual totals of the SPW and CLIM (up to 180 km) is shown in Fig. 10. Compared to the SPW network, the correlation (Pearson coefficient) increases from 0.70 (QPE1) to 0.85 (QPE2). There is underestimation of radar estimates, even with QPE2, for values exceeding 1100 mm, which are located over the hills nearby the radar. This can be at least partially explained by the higher height of valid measurements in
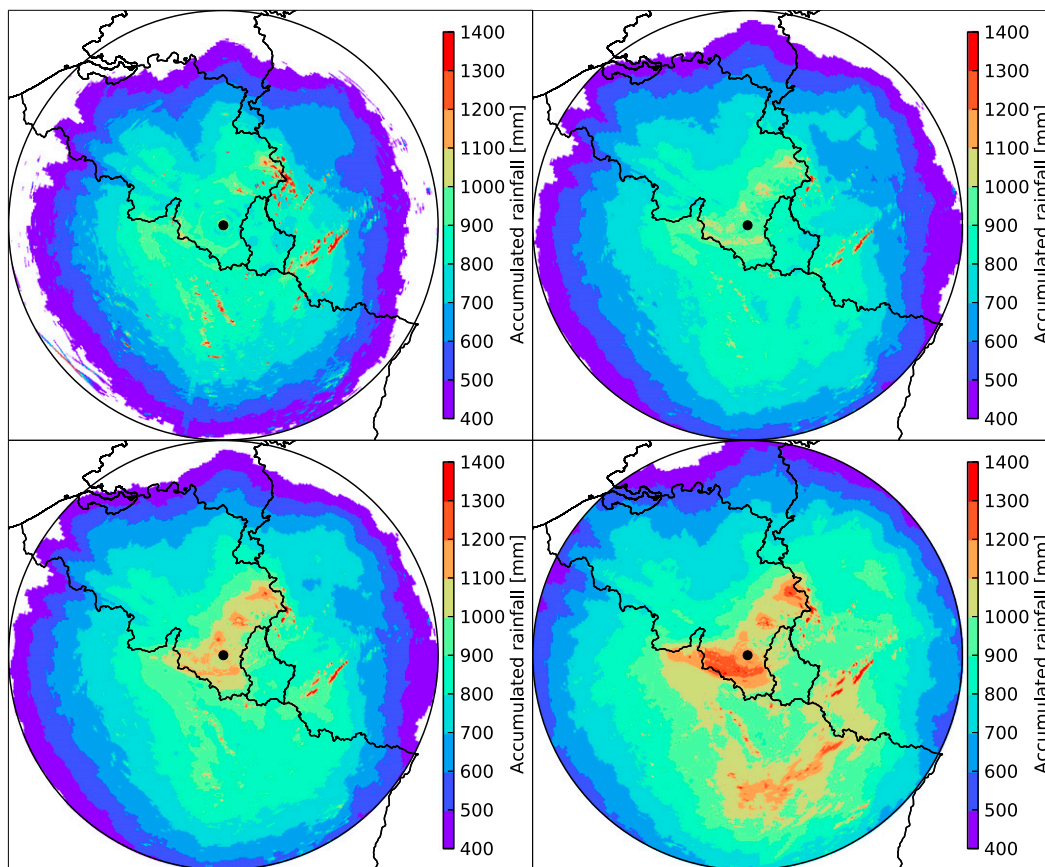


FIG. 9. Annual total mean for the 2005–15 period up to 220-km range based on (top left) QPE1, (top right) QPE2, (bottom left) QPE2 + MFB, and (bottom right) QPE2 + KED.
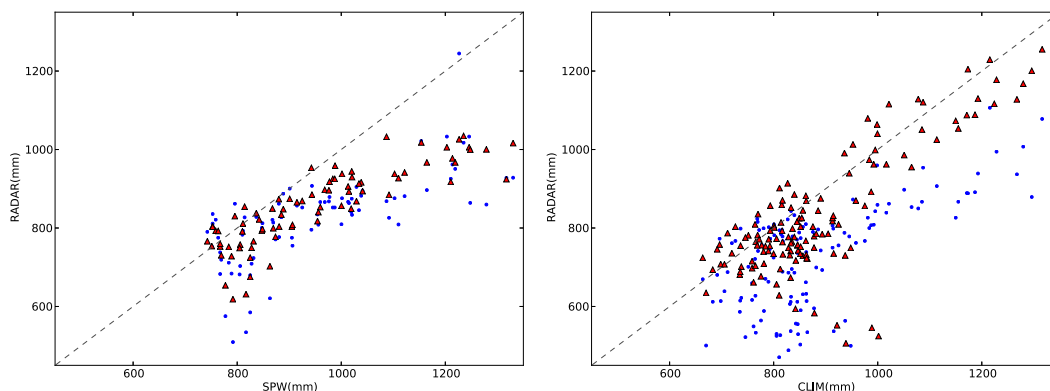
FIG. 10. Scatterplots of mean annual total (mm) (left) from QPE1 (blue points) and QPE2 (red triangles) against the SPW network and (right) from QPE1 (blue points) and QPE2 + KED (red triangles) against CLIM.

that region and the effect of orographic enhancement. The correlation with CLIM increases from 0.54 (QPE1) to 0.66 (QPE2) and 0.80 (QPE2 + KED). The lower correlation with CLIM can be explained by the greater proportion of gauges located far from the radar. However, a lower quality of some manual measurements cannot be ruled out. In the study region, it appears that the mean annual totals range between 600 and 1300 mm. There is a general correlation of the mean rainfall with topography (Fig. 2), and the maximum is found in the Belgian Ardennes. However, one notes a significant decrease of precipitation in the region with the highest topography when crossing the border toward Germany. This might be due to the positive and negative orographic effects on precipitation associated with the dominant southwesterly wind. The actual effect of topography is more difficult to interpret beyond 150 km because of systematic radar underestimation, as seen in the gauge verification. Those results are in good agreement with the precipitation maps for the period 1981–2010 obtained by Journée et al. (2015) using interpolated data from CLIM. Using 8 years of operational radar data in the United Kingdom, Fairman et al. (2015) also suggest that overshooting and orographic enhancement are the two main sources of difference with the gauges.

In Fig. 11, one can see the probability of hourly and daily accumulation to exceed 10 mm. For both statistics, using QPE2 + KED allows us to remove most radar artifacts appearing when using QPE1. The correlation with daily gauge values from CLIM (not shown) increases from 0.42 to 0.72. The probability of daily rainfall to exceed 10 mm ranges between 4% and 12% of the time, and it is correlated with the mean annual total. The probability to exceed 10 mm in 1 h is a less frequent event and exhibits higher variations. It is larger in the southeastern part of the domain, where it reaches a probability

of about 0.1%. This area is known for a higher probability of convective storms (Weckwerth et al. 2011).

Figure 12 shows the maximum values of hourly and daily accumulation with QPE1 and QPE2 without merging. The effect of merging on the extremes is limited given the fact that convective cells are poorly resolved by rain gauge networks and is therefore not shown. It is known that precipitation maxima on a 500-m square are lower than rain gauge maxima because of rainfall small-scale variability. However, for QPE1, one finds values higher than expected from extreme precipitation models for Belgium based on rain gauges (Van de Vyver 2012). Those unrealistic extremes, which are caused by clutter or by a wrong $Z$–$R$ relationship applied to hail, are mitigated in QPE2. The radar maxima become slightly lower than rain gauge maxima (not shown) as expected because of the averaging over a much bigger area. The hourly accumulation maximum exhibits high small-scale variations but no large-scale trend. The daily accumulation maximum also shows variations but at larger spatial scales. A region with higher values is still present in the southeast.

## 5. Conclusions

The polar volumetric data from a single-polarization weather radar (Wideumont, Belgium) have been reanalyzed from 2005 to 2015 using a new processing chain (i.e., QPE2) to generate rainfall rates. A static clutter map and a beam blockage map are used to characterize the quality of the data. Clutter is identified dynamically using three different techniques: a comparison with a cloud-type product from satellite observations, the identification of unrealistic vertical gradient of reflectivity, and the identification of abnormal horizontal reflectivity texture. A robust VPR correction has been applied to mitigate the effect of the bright band and the
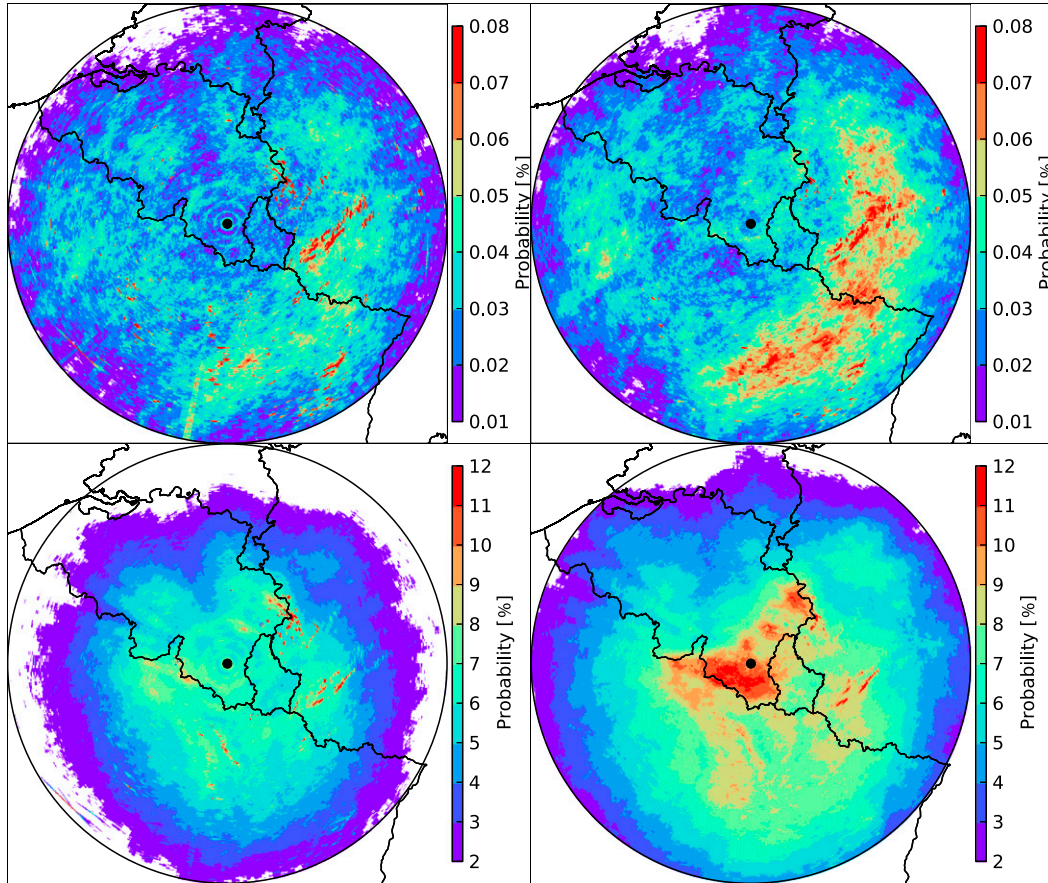
FIG. 11. Probability of (top) hourly and (bottom) daily accumulation to exceed 10 mm for 2005–15 based on (left) QPE1 up to 240-km range and (right) QPE2 + KED up to 220-km range.

underestimation at long ranges. A distinction between stratiform and convective precipitation is made using an enhanced version of the Steiner algorithm. Different VPR corrections and $Z$–$R$ relationships are applied for the two precipitation regimes. The maximum reflectivity is set to 55 dB$Z$ for mitigating contamination by hail. Using a dense hourly rain gauge network, linearly accumulated radar rainfall rates are adjusted by mean field bias. A kriging with external drift using a climatological exponential variogram with 10-km range and a square-root transformation is also applied to radar and rain gauge hourly rainfall estimates.

All algorithms have been fine-tuned on a selection of 13 cases with various meteorological situations. A validation of the cases is performed using an independent daily rain gauge network. The results for three different scores (MAD, SCS, and $F_{20}$) reveal a small to high benefit of QPE2 and the merging methods when compared to interpolated elevations at 800 m above radar level (QPE1). The VPR correction is responsible for most of the benefit while beam blockage correction and the alternative $Z$–$R$ relationships hardly improved the

results. It is interesting to note that the quality-controlled lowest elevation exhibits similar performance as QPE1.

A 10-yr verification of daily precipitation amounts has been performed against the independent daily network using the same scores. A categorical verification of values exceeding 1 mm shows that the false alarm ratio is reduced and the probability of detection is increased with QPE2, which are mainly due to clutter mitigation and VPR profile correction, respectively. Computing three different scores at each gauge location for values exceeding 1 mm reveals a clear benefit of QPE2 versus QPE1, especially at short and long range. In addition, QPE2 allows us to increase the range with good-quality estimates from 100 to 125 km or roughly 50% more coverage. Using the merging methods allows further improvements but reduces the benefit of QPE2, except at short range. The average improvements from QPE1 to QPE2 + KED for gauges up to 150 km are 38%, 35%, and 80% for MAD, SCS, and $F_{20}$, respectively. The average of daily statistics exhibits smaller improvements with 32%, 14%, and 52% obtained for the three scores with QPE2 + KED. This suggests that larger benefits
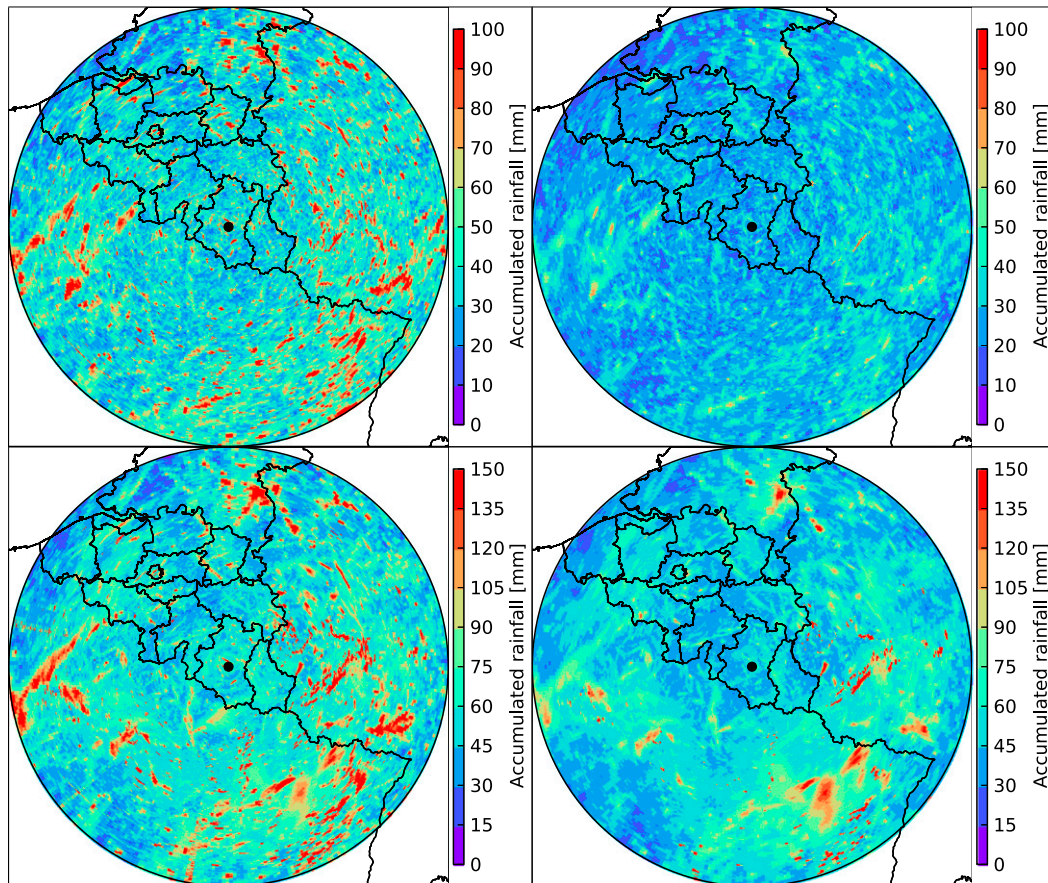
FIG. 12. Max (top) hourly and (bottom) daily accumulation for 2005–15 based on (left) QPE1 and (right) QPE2 up to 240-km range.

are obtained for widespread situations. The benefit increases when considering higher rainfall thresholds, and with 10-mm daily amounts the improvements are 36%, 18%, and 89%, respectively.

The 10-yr mean annual total clearly shows the mitigation of radar artifacts by QPE2, especially the clutter and the underestimation at long range. The correlation with rain gauge values increases from 0.54 for QPE1 to 0.80 for QPE2 + KED. The probability to exceed 10 mm for hourly and daily accumulation further highlights the improvements made by QPE2 and the merging methods with mitigation of artifacts and increased correlation with rain gauge values. Looking at the 10-yr maximum of hourly and daily accumulation reveals unrealistic high values for QPE1 compared to gauges because of clutter and hail. Using QPE2 mitigates those artifacts with more realistic hourly values slightly lower than rain gauge maxima as expected because of the bigger area of the radar estimate.

We would like to stress that finding parameters (e.g., reflectivity thresholds) that perform well in all conditions is particularly difficult. This task represents a significant part of the time spent on this study. For some situations with strong horizontal temperature gradients, the global VPR correction might not work properly, but this effect is limited by using data from a 40-min time window. While the most important errors have been mitigated, the processing of radar volume data could be further improved. One thinks of using dynamic beam blockage correction, using other radars to deal with attenuation, taking into account precipitation advection (e.g., using optical flow) when computing accumulation and refining the kriging with external drift (e.g., with a real-time variogram estimation). The use of model output like temperature could also help in the validation of the brightband identification. The verification of the precipitation dataset against rain gauge values gives a lower bound on the quality of the estimates. Therefore, quantifying the uncertainty of the new estimates remains a challenge. A new dual-polarization radar has been installed in Belgium since 2013 and volume data are archived. It will be interesting to see if the recently developed QPE algorithms making use of polarimetric information can significantly improve the scores of QPE2.

Given the good results obtained by the new radar processing chain, it is going to be implemented in real time. Since the radar now performs a 15 elevation scan every 5 min, further improvements are expected. The enhanced operational QPE will propose a better input for the nowcasting systems running at RMIB, Integrated Nowcasting through Comprehensive Analysis in Belgium (INCA-BE), and Short-Term Ensemble Prediction System (STEPS; Foresti et al. 2015). The validated 10-yr dataset produced here, which will be extended each year, can be used as the basis for further studies, especially in hydrology. A direct application is to generate a rainfall climatology and compare the results with conventional climatology based on rain gauges. One of the main benefits of using radar observations is the ability to compute areal rainfall extreme statistics for different area sizes.

## REFERENCES

Bech, J., U. Gjertsen, and G. Haase, 2007: Modelling weather radar beam propagation and topographical blockage at northern high latitudes. *Quart. J. Roy. Meteor. Soc.*, **133**, 1191–1204, doi:10.1002/qj.98.

Bellon, A., G. W. Lee, A. Kilambi, and I. Zawadzki, 2007: Real-time comparisons of VPR-corrected daily rainfall estimates with a gauge Mesonet. *J. Appl. Meteor. Climatol.*, **46**, 726–741, doi:10.1175/JAM2502.1.

Berenguer, M., D. Sempere-Torres, C. Corral, and R. Sánchez-Diezma, 2006: A fuzzy logic technique for identifying non-precipitating echoes in radar scans. *J. Atmos. Oceanic Technol.*, **23**, 1157–1180, doi:10.1175/JTECH1914.1.

Berndt, C., E. Rabiei, and U. Haberlandt, 2014: Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *J. Hydrol.*, **508**, 88–101, doi:10.1016/j.jhydrol.2013.10.028.

Biggerstaff, M. I., and S. A. Listemaa, 2000: An improved scheme for convective/stratiform echo classification using radar reflectivity. *J. Appl. Meteor.*, **39**, 2129–2150, doi:10.1175/1520-0450(2001)040<2129:AISFCS>2.0.CO;2.

de Leeuw, J., J. Methven, and M. Blackburn, 2015: Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations. *Quart. J. Roy. Meteor. Soc.*, **141**, 798–806, doi:10.1002/qj.2395.

Delobbe, L., and I. Holleman, 2006: Uncertainties in radar echo top heights used for hail detection. *Meteor. Appl.*, **13**, 361–374, doi:10.1017/S1350482706002374.

Delrieu, G., A. Wijbrans, B. Boudevillain, D. Faure, L. Bonnifait, and P.-E. Kirstetter, 2014: Geostatistical radar–raingauge merging: A novel method for the quantification of rain estimation accuracy. *Adv. Water Resour.*, **71**, 110–124, doi:10.1016/j.advwatres.2014.06.005.

Derrien, M., and H. Le Gléau, 2005: MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.*, **26**, 4707–4732, doi:10.1080/01431160500166128.

Erdin, R., C. Frei, and H. R. Künsch, 2012: Data transformation and uncertainty in geostatistical combination of radar and rain gauges. *J. Hydrometeor.*, **13**, 1332–1346, doi:10.1175/JHM-D-11-096.1.

Fairman, J. G., D. M. Schultz, D. J. Kirshbaum, S. L. Gray, and A. I. Barrett, 2015: A radar-based rainfall climatology of Great Britain and Ireland. *Weather*, **70**, 153–158, doi:10.1002/wea.2486.

Farr, T. G., and Coauthors, 2007: The Shuttle Radar Topography Mission. *Rev. Geophys.*, **45**, RG2004, doi:10.1029/2005RG000183.

Foresti, L., M. Reyniers, A. Seed, and L. Delobbe, 2015: Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium. *Hydrol. Earth Syst. Sci.*, **20**, 505–527, doi:10.5194/hess-20-505-2016.

Gabella, M., and R. Notarpietro, 2002: Ground clutter characterization and elimination in mountainous terrain. *Proc. European Conf. on Radar in Meteorology and Hydrology*, Delft, Netherlands, Delft University of Technology, 305–311. [Available online at http://www.copernicus.org/erad/online/erad-305.pdf.]

Germann, U., and J. Joss, 2002: Mesobeta profiles to extrapolate radar precipitation measurements above the Alps to the ground level. *J. Appl. Meteor.*, **41**, 542–557, doi:10.1175/1520-0450(2002)041<0542:MPTERP>2.0.CO;2.

——, G. Galli, M. Boscacci, and M. Bolliger, 2006: Radar precipitation measurement in a mountainous region. *Quart. J. Roy. Meteor. Soc.*, **132**, 1669, doi:10.1256/qj.05.190.

Goudenhoofdt, E., and L. Delobbe, 2009: Evaluation of radar–gauge merging methods for quantitative precipitation estimates. *Hydrol. Earth Syst. Sci.*, **13**, 195–203, doi:10.5194/hess-13-195-2009.

——, and ——, 2012: Long-term evaluation of radar QPE using VPR correction and radar–gauge merging. *IAHS Publ.*, **351**, 249–254.

——, and ——, 2013: Statistical characteristics of convective storms in Belgium derived from volumetric weather radar observations. *J. Appl. Meteor. Climatol.*, **52**, 918–934, doi:10.1175/JAMC-D-12-079.1.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.*, **113**, D20119, doi:10.1029/2008JD010201.

Hazenberg, P., P. J. J. F. Torfs, H. Leijnse, G. Delrieu, and R. Uijlenhoet, 2013: Identification and uncertainty estimation of vertical reflectivity profiles using a Lagrangian approach to support quantitative precipitation measurements by weather radar. *J. Geophys. Res. Atmos.*, **118**, 10 243–10 261, doi:10.1002/jgrd.50726.

Hubbert, J. C., M. Dixon, and S. M. Ellis, 2009: Weather radar ground clutter. Part II: Real-time identification and filtering. *J. Atmos. Oceanic Technol.*, **26**, 1181–1197, doi:10.1175/2009JTECHA1160.1.

Joss, J., and R. Lee, 1995: The application of radar–gauge comparisons to operational precipitation profile corrections. *J. Appl. Meteor.*, **34**, 2612–2630, doi:10.1175/1520-0450(1995)034<2612:TAORCT>2.0.CO;2.

Journée, M., C. Delvaux, and C. Bertrand, 2015: Precipitation climate maps of Belgium. *Adv. Sci. Res.*, **12**, 73–78, doi:10.5194/asr-12-73-2015.

Kirstetter, P.-E., H. Andrieu, G. Delrieu, and B. Boudevillain, 2010: Identification of vertical profiles of reflectivity for correction of volumetric radar data using rainfall classification. *J. Appl. Meteor. Climatol.*, **49**, 2167–2180, doi:10.1175/2010JAMC2369.1.

Koistinen, J., and H. Pohjola, 2014: Estimation of ground-level reflectivity factor in operational weather radar networks using VPR-based correction ensembles. *J. Appl. Meteor. Climatol.*, **53**, 2394, doi:10.1175/JAMC-D-13-0343.1.

Krajewski, W. F., 1987: Cokriging radar-rainfall and rain gage data. *J. Geophys. Res.*, **92**, 9571–9580, doi:10.1029/JD092iD08p09571.

——, and Coauthors, 2011: Towards better utilization of NEXRAD data in hydrology: An overview of Hydro-NEXRAD. *J. Hydroinf.*, **13**, 255–266, doi:10.2166/hydro.2010.056.

Lobligeois, F., V. Andréassian, C. Perrin, P. Tabary, and C. Loumagne, 2014: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrol. Earth Syst. Sci.*, **18**, 575–594, doi:10.5194/hess-18-575-2014.

Nelson, B. R., D.-J. Seo, and D. Kim, 2010: Multisensor precipitation reanalysis. *J. Hydrometeor.*, **11**, 666–682, doi:10.1175/2010JHM1210.1.

Overeem, A., I. Holleman, and A. Buishand, 2009: Derivation of a 10-year radar-based climatology of rainfall. *J. Appl. Meteor. Climatol.*, **48**, 1448–1463, doi:10.1175/2009JAMC1954.1.

Schiemann, R., R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, 2011: Geostatistical radar–raingauge combination with nonparametric correlograms: Methodological considerations and application in Switzerland. *Hydrol. Earth Syst. Sci.*, **15**, 1515–1536, doi:10.5194/hess-15-1515-2011.

Sideris, I. V., M. Gabella, R. Erdin, and U. Germann, 2013: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Quart. J. Roy. Meteor. Soc.*, **140**, 1097–1111, doi:10.1002/qj.2188.

Smith, J. A., M. L. Baeck, G. Villarini, C. Welty, A. J. Miller, and W. F. Krajewski, 2012: Analyses of a long-term, high-resolution radar rainfall data set for the Baltimore metropolitan region. *Water Resour. Res.*, **48**, W04504, doi:10.1029/2011WR010641.

Steiner, M., and J. A. Smith, 2002: Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data. *J. Atmos. Oceanic Technol.*, **19**, 673–686, doi:10.1175/1520-0426(2002)019<0673:UOTDRS>2.0.CO;2.

Tabary, P., 2007: The new French operational radar rainfall product. Part I: Methodology. *Wea. Forecasting*, **22**, 393–408, doi:10.1175/WAF1004.1.

Tang, L., Y. Tian, and X. Lin, 2014: Validation of precipitation retrievals over land from satellite-based passive microwave sensors. *J. Geophys. Res. Atmos.*, **119**, 4546–4567, doi:10.1002/2013JD020933.

Thorndahl, S., J. E. Nielsen, and M. R. Rasmussen, 2014: Bias adjustment and advection interpolation of long-term high resolution radar rainfall series. *J. Hydrol.*, **508**, 214–226, doi:10.1016/j.jhydrol.2013.10.056.

Uijlenhoet, R., and A. Berne, 2008: Stochastic simulation experiment to assess radar rainfall retrieval uncertainties associated with attenuation and its correction. *Hydrol. Earth Syst. Sci.*, **12**, 587–601, doi:10.5194/hess-12-587-2008.

Van de Vyver, H., 2012: Spatial regression models for extreme precipitation in Belgium. *Water Resour. Res.*, **48**, W09549, doi:10.1029/2011WR011707.

Vignal, B., H. Andrieu, G. Delrieu, and J. D. Creutin, 2003: Identification of rain-rate profiles from radar returns at attenuating wavelengths using an inverse method: A feasibility study. *J. Appl. Meteor.*, **42**, 1014–1030, doi:10.1175/1520-0450(2003)042<1014:IORPFR>2.0.CO;2.

Villarini, G., and W. F. Krajewski, 2010: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv. Geophys.*, **31**, 107–129, doi:10.1007/s10712-009-9079-x.

Wagner, A., J. Seltmann, and H. Kunstmann, 2012: Joint statistical correction of clutters, spokes and beam height for a radar derived precipitation climatology in southern Germany. *Hydrol. Earth Syst. Sci.*, **16**, 4101–4117, doi:10.5194/hess-16-4101-2012.

Weckwerth, T. M., J. W. Wilson, M. Hagen, T. J. Emerson, J. O. Pinto, D. L. Rife, and L. Grebe, 2011: Radar climatology of the COPS region. *Quart. J. Roy. Meteor. Soc.*, **137**, 31–41, doi:10.1002/qj.747.

Wilks, D., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction.* International Geophysics Series, Vol. 59, Elsevier, 467 pp.

Yang, S., and S. W. Nesbitt, 2014: Statistical properties of precipitation as observed by the TRMM Precipitation Radar. *Geophys. Res. Lett.*, **41**, 5636–5643, doi:10.1002/2014GL060683.

Zhang, J., and Y. Qi, 2010: A real-time algorithm for the correction of brightband effects in radar-derived QPE. *J. Hydrometeor.*, **11**, 1157–1171, doi:10.1175/2010JHM1201.1.
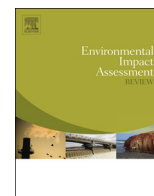
# Well-being, life satisfaction and capabilities of flood disaster victims☆

Luc Van Ootegem [a,b,*], &, Elsy Verhofstadt [b]

[a] HIVA–University of Louvain, Belgium
[b] SHERPPA–Ghent University, Belgium

## ARTICLE INFO

## ABSTRACT

The individual well-being of flood disaster victims is examined making use of two concepts: life satisfaction and perceived capabilities in life. These concepts are compared in two samples: a representative sample of Flemish respondents and a specific sample of people that have been the victim of a pluvial flood. Well-being as life satisfaction is found not to be related to past or expected future flooding, whereas well-being as capabilities in life is negatively related to both past and expected future flooding.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

Natural events and disasters can lead to very different kinds of damage. Typically, damage types are classified into direct vs indirect and tangible vs intangible effects (Jonkman et al., 2008). Indirect damage occurs as the result of direct damage and has a different time frame and/or space dimension; that is, it occurs after the event has passed and/or outside the disaster area (Smith and Ward, 1998). Tangible damage can easily be monetized, while putting a monetary value on intangible damage (not traded on the market) is more difficult (Smith and Ward, 1998). Examples of intangible, direct damage are fatalities and injuries, moral damage and inconvenience such as transportation problems or environmental losses. Intangible, indirect types of damage include psychological or (mental) health problems and political, societal or environmental consequences (Jonkman et al., 2008). Neglecting indirect, intangible effects is problematic (see among others Parker et al., 2007; Messner and Meyer, 2006; Murphy and Gardoni, 2006) as it strongly and wrongly decreases the estimated benefits when protection and investment decisions are made.

This paper investigates an intangible, indirect effect of one specific kind of natural disaster, namely, flooding. Most of the research on flood damage focuses on fluvial floods caused by a river overflowing (for Flanders, see Kellens et al., 2013) and uses "depth–damage" functions to measure short-term direct and tangible damage. We analyze the well-being of victims of pluvial floods, which are floods caused by extreme rainfall events that cannot be processed by existing urban drainage systems. This flood type may be less spectacular than fluvial floods, but is more common in urbanized areas.

There are very few studies on risk assessment which use individuals' well-being as the outcome variable. Most of the research addressing intangible and indirect effects concentrates on mental health as outcome variable (for a warning about the effect of climate change on mental health, see Berry et al., 2010). In the field of epidemiology, many scholars have demonstrated important mortality and health effects both shortly and long after floods took place (for an overview, see, e.g., Ahern et al., 2005; for an application of the effect of floods on mental health in Brisbane, see Alderman et al., 2013). Already in 1970, a famous study by Bennet (1970) showed more psychological problems for victims of the flooding in Bristol during the 12 months after the flood compared to a control group, in addition to increased deaths and hospital referrals. In many later studies, comparable effects were found on anxiety, depressions, and posttraumatic stress (Adeola, 2009; Bourque et al., 2006; Liu et al., 2006; Tapsell and Tunstall, 2008; Tobin and Ollenburger, 1996). These negative psychological effects have been shown to linger for years after a flood event (Hajat et al., 2003; Tapsell and Tunstall, 2008). Reacher et al. (2004) discovered that victims of floods suffer more than other people from diseases and other physical issues such as gastrointestinal problems or earache, which clearly cannot be explained directly by the flood itself. This research on the mental health effects of disasters can be complementary to studies which take individual well-being as the outcome variable. The main difference is that (mental) health effects can be more directly linked to the event while effects on well-being are much more indirect and therefore less clearly attributable to the flood event.

For this study, we consider subjective well-being, making use of self-reported information, and we compare two indicators: a traditional

satisfaction-with-life indicator (as in Diener, 2000; Blanchflower and Oswald, 2004, 2008; Dolan et al., 2008; Stiglitz et al., 2009; Helliwell et al., 2012) and a "perceived capabilities" indicator (Van Ootegem and Verhofstadt, 2012 and forthcoming). Capabilities are defined as the options or opportunities individuals have in life, which is essential to evaluate individual well-being (Alkire, 2005; Fleurbaey, 2006; Gasper, 2007; Kuklys, 2005; Robeyns, 2006; Schokkaert, 2009; Sen, 1985; 1993). The capabilities framework is theoretically and ethically appealing, but implementation is a real challenge. Not only do researchers need a lot of data, but they also have to make choices in order to first define and measure the different dimensions of an individual's set of capabilities and then aggregate these dimensions to obtain a composite index. This is particularly challenging as part of an individual's capabilities is – by definition – not observable. As far as we know, only one study applied the capability approach to the impact of natural disasters: Gardoni and Murphy's (2010) Disaster Impact Index (DII). The DII was illustrated for four disasters: earthquakes in Japan, Pakistan and the United States, and hurricane Katrina. However, from a theoretical point of view, a capabilities approach should focus on people (individuals or households), not countries. Therefore, we opted to directly ask people to evaluate their opportunities or capabilities (Van Ootegem and Verhofstadt, forthcoming). Proceeding as such, we sacrificed some of the objectivity of the concept of capabilities in order to have one composite indicator at the level of individuals. The dimensions or weights are then chosen by the people themselves, not the researchers.

The flood events selected for this study took place at various moments for various people, but individual well-being was measured for one specific year (2013). Consequently, it is possible that some time has passed and other events have occurred in between the flood event and the measurement of well-being in 2013. Similar to the literature on the impact of unemployment on well-being (e.g., Clark et al., 2010; Carr et al., 2011; Lange, 2013), we hypothesized that (past) flooding "scars" because it "scares" (terminology introduced by Knabe and Rätzel, 2011). It is then the fear of more (future) flooding that influences people's well-being rather than having experienced a flood as such. The specific relation between a prior experience of a natural disaster and subsequent behaviour and quality of life was studied in detail for the Katrina catastrophe in Adeola (2009). O'Donnell et al. (2014) stated that life evaluation questions such as on life satisfaction "capture a reflective assessment of how one's life is going" and "are the result of a cognitive evaluation on the part of the subject rather than a description of current emotional state" (p. 28). In contrast, the capabilities concept has – by definition – a more forward-looking perspective as it reflects opportunities in life and "respects the individual's ability to pursue and realise the goals that he or she values" (Stiglitz et al., 2009, p. 152). This distinction was also confirmed by the analysis in Van Ootegem and Verhofstadt (forthcoming): answering the satisfaction question is a more backward-looking reflection, thinking about capabilities and opportunities refers to a more forward-looking exercise. Therefore, the difference between past and (fear for) future flooding may be relevant when comparing satisfaction and capabilities.

In the next section, we explain the data collection methodology and compare the well-being of flood victims with that of non-victims using information on satisfaction and perceived capabilities. Section three examines the determinants (multivariate) of the well-being of flood victims and non-victims, and section four concludes.

## 2. Comparing flood victims with non-victims

We compared data from two sources. First, in 2013, a survey was sent out to identified victims of pluvial floods in Flanders (the northern Dutch-speaking part of Belgium). This survey asked the participants to evaluate several aspects of their well-being and to provide information about the flood disaster they were confronted with. Specifically, they were asked about the severity of the flood (depth, duration and tangible

damage they suffered), the recurrence of floods (how often they have been the victim of floods in the past) and their fear of future flooding. This data collection is part of the Plurisk project about pluvial risks (see infra) and is therefore referred to as the Plurisk survey. Our second source is a representative survey (LEVO[1] 2013) of 1291 Flemish respondents, most of whom have of course not experienced flooding. Only 6.1% of the LEVO respondents have been a victim of a pluvial flood, while 5.3% reported another (non-pluvial) flood-related problem. These participants were asked to assess the same aspects of their well-being as in the Plurisk survey, and thus act as the control or reference group.

The Plurisk survey was distributed among private households that were presumed to be affected by one or more pluvial floods. The survey was sent to 3963 addresses all across Flanders. The majority of these addresses come from a database of the Belgian national disaster relief fund. This fund collected the addresses of pluvial flood victims, but only until 2007. The reason for this is a change in the legislation that year, obliging insurance companies to provide fire and flood insurance in one package, thus ending the need for government compensations that were provided by the disaster relief fund (Portaal Belgische overheid, 2012). The lack of recent data was tackled in two ways. First, we included 260 addresses from records of fire and police departments as well as local authorities in villages and cities that were flooded in recent years during a pluvial event. Second, we asked people to fill out the questionnaire for the most severe flood since 2000 at their address. Many of the addresses in the national disaster relief fund were expected to be quite prone to pluvial floods, for instance, because they are close to malfunctioning sewer systems or in lower parts of the village or city. As such, a fairly high number of victims were confronted with flooding on a regular basis. Since participants were asked to take the worst flood as a reference, they could also report floods after 2007. Twenty-one percent of the reported floods occurred after the year 2007. A total of 973 households completed the survey (24.6% of our sample). In a first step, 353 files were deleted, 260 of which were deemed not useful because respondents claimed never to have suffered damage from pluvial floods. A number of respondents suffered from damage caused by hail or winds. This type of damage was not separated from flood damage in the disaster fund database. In addition, some people moved to the address found in the disaster fund database after the recorded pluvial event took place at that address. The files of 93 other respondents were also deleted, mainly because they turned out to be small shopkeepers, farmers or other self-employed businessmen reporting the damage to their business. The data of the remaining 620 respondents were used to perform further analyses. In Van Ootegem et al. (2015), we examine the reported monetary damage and estimate (depth) damage functions. We found flood depth to be an important predictor of damage, but with a different impact depending on whether the flood occurred on the ground floor or in the basement. Non-hazard indicators (e.g., risk awareness) are also important for the predicted damage, revealing that warning systems and policies can be valuable.

The data from the Plurisk survey are compared to self-reported information obtained from the LEVO survey. As the LEVO respondents are used as a control group of non-victims of flooding, we excluded 147 respondents that report that they have experienced flooding. However, including these respondents does not alter our conclusions. Since the Plurisk sample does not contain any students, this group (100 respondents) was also deleted from the control group. Another five respondents were deleted because too many variables were missing. This resulted in a sample of 1039 respondents. These data were then weighted to obtain a sample representative for the Flemish population according to life situation, gender and age distribution. The weighted

---

[1] LEVO is the Dutch acronym for "LEvensomstandigheden in Vlaanderen Onderzocht" (research on living circumstances in Flanders). It is a yearly large-scale survey organized in the context of a research seminar at Ghent University. The field work is carried out by Ghent University students. Organization, supervision, controlling and cleaning is performed by the authors.

**Table 1**
Socio-economic indicators of respondents in the two samples.

| Indicator | Plurisk | LEVO-weighted |
|---|---|---|
| *Gender* | | |
| Male | 55.4% | 49.6% |
| Female | 44.6% | 50.4% |
| | | |
| *Occupational situation* | | |
| Working full-time | 37.5% | 44.0% |
| Working part-time | 8.2% | 14.5% |
| Retired | 47.6% | 28.0% |
| Unemployed | 1.5% | 4.4% |
| Disabled with benefits | 2.5% | 3.9% |
| Unpaid homemaker | 2.7% | 5.2% |
| | | |
| *Education* | | |
| Basic education | 10.9% | 12.6% |
| Lower secondary | 17.0% | 17.1% |
| Higher secondary | 28.6% | 33.9% |
| Bachelor | 25.1% | 20.8% |
| Master or post-university | 18.3% | 15.6% |
| Age (mean) | 60.03 | 51.83 |

sample contains an equivalent of 847 cases. Table 1 compares some socio-economic characteristics of the two samples. The average age of 60 in the Plurisk database shows that many elderly sent back the questionnaire. This higher average age is also reflected in the fact that over 45% of the respondents are retired.

We used two variables as proxies for subjective well-being. Respondents were asked about their general satisfaction: "How is your satisfaction with life in general?", and about their perceived capabilities: "How do you consider your possibilities/opportunities in life in general?" Both questions had to be rated on a scale from 0 ("very unsatisfied" and "totally insufficient") to 10 ("very satisfied" and "excellent"). In Table 2, flood victims (Plurisk) are compared with people that have never been exposed to floods (LEVO). The average life satisfaction scores are not significantly different. However, there is a difference in perceived capabilities: the average capabilities are significantly ($p = 0000$) lower for the flood victims (7.01 < 7.45).

All respondents were asked to estimate the probability of becoming victims of a flood in the next five years. On average, the LEVO respondents reported probabilities of 4.5%, 4.9% and 5.5% for a flooding of the ground floor, the garage and the basement, respectively. For Plurisk respondents, the average probabilities are much higher, 14.0%, 21.4% and 27.5%, respectively. Combining the Plurisk and LEVO data reveals that the probability of future flooding is not correlated with life satisfaction. For perceived capabilities, however, the negative correlation is significant (see table 3). This probability of future flooding can be considered as a proxy for the fear of future floods (i.e., the idea that past flooding scars because it scares).

## 3. Determinants of the well-being of flood victims and non-victims

We now know that flood victims report lower perceived capabilities and a higher probability of future flooding. However, the well-being of people defined and measured in 2013 will be influenced by other factors and events than only a flood disaster. Therefore, our interpretation must be cautious. While depth–damage correlations can be interpreted as real and direct effects of the flood, the well-being correlations do not

**Table 2**
Mean well-being (satisfaction and capabilities) according to flood history.

| | | N | Mean |
|---|---|---|---|
| **Satisfaction** | LEVO | 846 | 7.25 |
| | Plurisk | 536 | 7.12 |
| **Capabilities** | LEVO | 834 | **7.45** |
| | Plurisk | 536 | **7.01** |

**Table 3**
Pearson correlation between the estimated future flood probability and well-being (satisfaction and capabilities).

| | Satisfaction | Capabilities |
|---|---|---|
| Ground floor | −.020 | **−.120**[***] |
| Garage | −.040 | **−.097**[***] |
| Basement | −.033 | **−.128**[***] |

[***] Correlation significant at the 0.01 level (data LEVO and Plurisk merged).

imply causality. The difference in capabilities may be caused by other differences in the sample characteristics. Therefore, we proceed with a multivariate analysis, which enables us to evaluate the relative importance of a flood event for individual well-being, taking into account the effect of gender, age, income and the reported situation on health, social life and expectations (the robustness analysis also includes personality traits). First, the variables for the multivariate analysis were selected based on a short overview of the empirical literature on the determinants of subjective well-being.

Many studies have summarized the main factors influencing subjective well-being (among others, Blanchflower and Oswald, 2004; Dolan et al., 2008; Stiglitz et al., 2009; Helliwell et al., 2012; O'Donnell et al., 2014). The key determinants mostly identified are gender, age, income, education, work, family life, social capital, religion, environment, mental and physical health. For the regression analysis in this study (see infra), the following indicators were selected: gender, age and age squared, personal income, social life and health. For "social life" and "health," respondents were asked to indicate to what extent they agreed with the following questions using a scale from 0 (completely disagree) to 10 (completely agree): "I have a good social life (e.g., having friends, being part of associations)" and "I consider myself to be in good physical health." "Age squared" was included to capture the possibility that happiness is U-shaped over the life cycle (see, for example, Blanchflower & Oswald, 2008). As an income variable, we use "personal income." Alternatively for personal income, we used models including educational level and the current socio-economic position. When we do this, the results for the flood variables remained unchanged. Given the correlation between income and education, on the one hand, and between age, socio-economic position, social life and health, on the other hand, only income was retained as variable in the regression models that are presented.

The importance of expectations is indisputable in happiness research (e.g., Veenhoven, 1991; Diener, 2000; Helliwell et al., 2012). Therefore, we included a measure for expectations in the analysis (rating having high expectations on a numeric 7-point scale). In a robustness analysis, we also included traits related to personality: emotional, extravert, conscientious, altruistic and creative. This did not affect the results of the study.

After controlling for gender, age, income, health situation, social life and expectations, we found that both past and possible future floods have a significant importance for perceived capabilities but not for life satisfaction (Table 4). When both past and possible future floods are included in the same model, both remain significant, although at a lower significance level because of their correlation. Those with a past flooding experience have an average estimated future probability of 14%, while this is only 4% for the group of non-victims.

There is no effect of gender on any of the well-being proxies, in contrast to health and social life, which significantly influence both well-being indicators. Income is especially important for the capabilities indicator. All the estimated coefficients are rather stable when past flooding is substituted by future expected flooding. Alternative estimations (replacing health or social life with the educational level and the current socio-economic situation, including personality traits) produce similar results for both flood variables[2].

---

[2] Available on request from the authors.

**Table 4**
OLS regressions for satisfaction and capabilities—unstandardized coefficients.

| | Satisfaction | | Capabilities | |
|---|---|---|---|---|
| Constant | 1.938*** | 1.801*** | 2.839*** | 2.921*** |
| **Dummy for being a flood victim** | 0.037 | | −0.324*** | |
| **Future probability of flood on ground floor** | | 0.000 | | −0.006*** |
| Gender (dummy for men) | 0.048 | 0.046 | 0.050 | 0.031 |
| Age | 0.029* | 0.033** | 0.018 | 0.016 |
| Age² | −0.000 | 0.000 | 0.000 | 0.000 |
| Health | 0.199*** | 0.208*** | 0.173*** | 0.178*** |
| Social life | 0.222*** | 0.216*** | 0.255*** | 0.247*** |
| Personal income (in thousands, €) | 0.110* | 0.110* | 0.227*** | 0.212*** |
| High expectations | 0.222*** | 0.225*** | 0.127*** | 0.128*** |
| | Prob > F = 0.000 | Prob > F = 0.000 | Prob > F = 0.000 | Prob > F = 0.000 |
| | Adj. R² = 0.184 | Adj. R² = 0.186 | Adj. R² = 0.295 | Adj. R² = 0.286 |

*** $p < 0.01$ (data LEVO and Plurisk merged).
** $p < 0.05$ (data LEVO and Plurisk merged) .
* $p < 0.10$ (data LEVO and Plurisk merged).

In summary, when comparing flood victims with non-victim respondents, we found that past and expected future flooding are correlated with a lack of reported capabilities in life. The question on perceived capabilities appears to be interpreted as a question on the person's opinion about future possibilities and opportunities in life. In contrast, flooding information does not correlate with reported life satisfaction.

## 4. Conclusion

We compared two samples of Flemish respondents (2013), one comprising victims of a flood disaster in the past (Plurisk) and the second comprising non-victims of a flood (LEVO). We used two proxies for subjective well-being: the well-known satisfaction-with-life indicator and an alternative indicator that asks for the perceived capabilities. The latter is different from the capabilities application to disasters by Gardoni and Murphy (2010) because they investigated the societal impact of natural disasters whereas our study is based on self-reported data at the individual level.

When comparing flood disaster victims with non-victim respondents, we observed that past flooding is not correlated with life satisfaction. Conversely, having been a victim of floods is associated with a lack of capabilities in life. Particularly, there is a significant negative correlation between perceived capabilities and the fear of future flooding. This indicates that a flood "scares." These findings confirm that "there is a major negative impact on general quality of life" (Adeola, 2009, p. 483), quality of life being conceptualized here as reported capabilities. Looking back to the past, people have adapted to what happened, and therefore reported satisfaction is not influenced by the flood event. Looking forward, being afraid of a next flood, is translated into lower reported capabilities.

These results suggest that monetary damage and life satisfaction are not the only variables needed to assess the possible impact of a disaster. Using satisfaction as metric for subjective well-being would imply that victims do not need to be compensated since our findings show no significant difference in life satisfaction. Therefore, compensation based on a capabilities deficiency should be considered, as this does yield a significant result. As such, these findings call for a reflection on the use of well-being concepts when addressing the possible impact of disasters, also when it comes to compensation. There must be a re-evaluation of which notions are appropriate for which application for which kind of disaster.

The significant differences between life satisfaction and perceived capabilities found in Van Ootegem and Verhofstadt (forthcoming) are

confirmed. It is intuitively obvious that the two concepts as such provoke different life evaluations. We can now add that a disaster (a flood) can be correlated with significant differences between the two concepts. The question remains whether this is also the case for other events or disasters. For the design of a compensation policy for insurances or governmental support, it is crucial to know which kind of potential damage is relevant when going beyond direct and monetary damage.

Another important route for further research is a longitudinal investigation of the exact nature of the causality between disasters and different well-being concepts. Our findings support the idea that a (flood) disaster can make people "scared" for the future because of the "scars" that it provokes. This scarring effect of certain events on subjective well-being might thus stretch much further than its traditional use in research about the impact of unemployment on well-being.

## References

Adeola, F.O., 2009. Katrina cataclysm, does duration of residency and prior experience affect impacts, evacuation, and adaptation behavior among survivors. *Environ. Behav.* 41 (4), 459–489.

Ahern, M., Kovats, R.S., Wilkinson, P., Few, R., Matthies, F., 2005. Global health impacts of floods: epidemiologic evidence. Epidemiol. Rev. 27 (1), 36–46.

Alderman, K., Turner, L.R., Tong, S., 2013. Assessment of the health impacts of the 2011 summer floods in Brisbane. Disaster Med. Public Health Prep. 7, 380–386.

Alkire, S., 2005. Why the capability approach. J. Hum. Dev. 6 (1), 115–134.

Bennet, G., 1970. Bristol floods 1968. Controlled survey of effects on health of local community disaster'. Br. Med. J. 3 (5720), 454–458.

Berry, H.L., Bowen, K., Kjellstrom, T., 2010. Climate change and mental health: a causal pathways framework. *Int. J. Public Health* 55 (2), 123–132.

Blanchflower, D., Oswald, A., 2004. Well-being over time in Britain and the United States. J. Public Econ. 88 (7/8), 1359–1386.

Blanchflower, D., Oswald, A., 2008. 'Is well-being U-shaped over the life cycle? *Soc. Sci. Med.* 66, 1733–1749.

Bourque, L.B., Siegel, J.M., Kano, M., Wood, M.M., 2006. Weathering the storm: the impact of hurricanes on physical and mental health. Ann. Am. Acad. Pol. Soc. Sci. 604 (1), 129–151.

Carr, E., Elliot, M., Tranmer, M., 2011. 'A multilevel analysis of the relationship between national economic conditions, an individual's job insecurity and well-being in Western Europe'. CCSR Paper 2011–05.

Clark, A., Knabe, A., Rätzel, S., 2010. Boon or bane? Others' unemployment, well-being and job insecurity'. *Labour Econ.* 17, 52–61.

Diener, E., 2000. Subjective well-being: the science of happiness and a proposal for a national index. *Am. Psychol.* 55, 34–43.

Dolan, P., Peasgood, T., White, M., 2008. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. J. Econ. Psychol. 29 (1), 94–122.

Fleurbaey, M., 2006. Capabilities, functionings and refined functionings. J. Hum. Dev. 7 (3), 299–310.

Gardoni, P., Murphy, C., 2010. Gauging the societal impacts of natural disasters using a capability approach. *Disasters* 34 (3), 619–636.

Gasper, D., 2007. What is the capability approach? Its core, rationale, partners and dangers. J. Socio-Econ. 36 (3), 335–359.

Hajat, S., Ebi, K.L., Kovats, S., Menne, B., Edwards, S., Haines, A., 2003. The human health consequences of flooding in Europe and the implications for public health: a review of the evidence. Appl. Environ. Sci. Public Health 1 (1), 13–21.

Helliwell, J., Layard, R., Sachs, J., 2012. World Happiness Report. Columbia University, New York, The Earth Institute.

Jonkman, S.N., Bockarjova, M., Kok, M., Bernardini, P., 2008. Integrated hydrodynamic and economic modelling of flood damage in The Netherlands. *Ecol. Econ.* 66 (1), 77–90.

Kellens, W., Vanneuville, W., Verfaillie, E., Meire, E., Deckers, P., De Maeyer, P., 2013. Flood risk management in Flanders: past developments and future challenges. *Water Resour. Manag.* 27, 3585–3606.

Knabe, A., Rätzel, S., 2011. Scarring or scaring? The psychological impact of past unemployment and future unemployment risk. *Economica* 78, 283–293.

Kuklys, W., 2005. *Amartya Sen's Capability Approach*. Springer, Berlin.

Lange, T., 2013. Scarred from the past or afraid of the future? Unemployment and job satisfaction across European labour markets. Int. J. Hum. Resour. Manag. 24 (6), 1096–1112.

Liu, A., Tan, H., Zhou, J., et al., 2006. An epidemiologic study of posttraumatic stress disorder in flood victims in Hunan China. Can. J. Psychiatr. 51 (6), 350–354.

Messner, F., Meyer, V., 2006. 'Flood damage, vulnerability and risk perception—challenges for flood damage research. In: Schanze, J., Zeman, E., Marsalek, J. (Eds.), *Flood risk management*: hazards, vulnerability and mitigation measures. NATO Science Series vol 67.

Murphy, C., Gardoni, P., 2006. The role of society in engineering risk analysis: a capabilities-based approach. *Risk Anal.* 26 (4), 1073–1083.

O'Donnell, G., Deaton, A., Durand, M., Halpern, D., Layard, R., 2014. Wellbeing and Policy—Full Report. Legatum Institute.

Parker, D., Tapsell, S., McCarthy, S., 2007. Enhancing the human benefits of flood warnings. *Nat. Hazards* 43, 397–414.

Reacher, M., McKenzie, K., Lane, C., et al., 2004. Health impacts of flooding in lewes: a comparison of reported gastrointestinal and other illness and mental health in flooded and non-flooded households. *Commun. Dis. Public Health* 7 (1), 39–46.

Robeyns, I., 2006. The capability approach in practice. *J. Political Philos.* 14 (3), 351–376.

Schokkaert, E., 2009. The capabilities approach. In: Anand, P., Puppe, C., Pattanaik, P. (Eds.), The Handbook of Rational and Social Choice. Oxford University Press, Oxford, pp. 542–566.

Sen, A., 1985. Commodities and Capabilities. North Holland, Amsterdam.

Sen, A., 1993. 'Capability and well-being'. In: Nussbaum, M., Sen, A. (Eds.), *The Quality of Life*. Clarendon Press, Oxford.

Smith, K., Ward, R., 1998. *Floods: Physical Processes and Human Impacts*. Wiley, Chichester.

Stiglitz, J.E., Sen, A., Fitoussi, J.-P., 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress. Commission on the Measurement of Economic Performance and Social Progress.

Tapsell, S.M., Tunstall, S.M., 2008. "I wish I'd never heard of banbury": the relationship between 'place' and the health impacts from flooding. Health Place 14 (2), 133–154.

Tobin, G.A., Ollenburger, J.C., 1996. Predicting levels of postdisaster stress in adults following the 1993 floods in the upper midwest. Environ. Behav. 28 (3), 340–357.

Van Ootegem, L., Verhofstadt, E., 2012. Using capabilities as an alternative indicator for well-being. Soc. Indic. Res. 106 (1), 133–152.

Van Ootegem, L., Verhofstadt, E., 2015. Perceived capabilities as an aggregated indicator for well-being'. *Appl. Res. Qual. Life* http://dx.doi.org/10.1007/s11482-014-9343-1 (forthcoming).

Van Ootegem, L., Verhofstadt, E., Van Herck, K., Creten, T., 2015. Multivariate pluvial flood damage models. Environ. Impact Assess. Rev. 54, 91–100.

Veenhoven, R., 1991. Is happiness relative? Soc. Indic. Res. 24 (1), 1–34.

**Luc Van Ootegem** is a professor in Economics at Ghent University (Faculty of Economics and Business Administration) where he teaches Welfare Economics, Environmental Economics and Social and Economic Policy. He has a PhD in economics from the KU Leuven (1995). From 2003 till 2008, he was responsible for the master's program in Applied Economics and Business Administration at the University College Gent. Since then, he is combining teaching and research on well-being (measurement and evaluation) and applied welfare economics. He is a part-time project manager in the domain of research on well-being and sustainable development and environmental policy at the Research Institute for Work and Society (HIVA, KU Leuven) in the Research Team on "Environmental Policy and Sustainable Development".

**Elsy Verhofstadt** (PhD in economics) is an associate professor at Ghent University, Department of Social Economics, and affiliated to SHERPPA. She teaches Research Methods, Social and Economic Policy and Welfare Economics. Her research interests include among others the measurement and evaluation of well-being and the relation with sustainability. Her contributions to research have been published, among others, in *Journal of Happiness Studies, Work & Stress* and *Social Indicators Research*.

# Enhanced object-based tracking algorithm for convective rain storms and cells

Carlos Muñoz[a], Li-Pen Wang[a,b], Patrick Willems[a]

[a] *KU Leuven, Dept. of Civil Engineering, Hydraulics Section, Leuven 3001, Belgium*
[b] *RainPlusPlus Ltd., Derby DE1 3RL, UK*

ABSTRACT

This paper proposes a new object-based storm tracking algorithm, based upon TITAN (Thunderstorm Identification, Tracking, Analysis and Nowcasting). TITAN is a widely-used convective storm tracking algorithm but has limitations in handling small-scale yet high-intensity storm entities due to its single-threshold identification approach. It also has difficulties to effectively track fast-moving storms because of the employed matching approach that largely relies on the overlapping areas between successive storm entities. To address these deficiencies, a number of modifications are proposed and tested in this paper. These include a two-stage multi-threshold storm identification, a new formulation for characterizing storm's physical features, and an enhanced matching technique in synergy with an optical-flow storm field tracker, as well as, according to these modifications, a more complex merging and splitting scheme. High-resolution (5-min and 529-m) radar reflectivity data for 18 storm events over Belgium are used to calibrate and evaluate the algorithm. The performance of the proposed algorithm is compared with that of the original TITAN. The results suggest that the proposed algorithm can better isolate and match convective rainfall entities, as well as to provide more reliable and detailed motion estimates. Furthermore, the improvement is found to be more significant for higher rainfall intensities. The new algorithm has the potential to serve as a basis for further applications, such as storm nowcasting and long-term stochastic spatial and temporal rainfall generation.

## 1. Introduction

Extreme convective rain storms may cause pluvial flooding and potentially severe socio-economic consequences. For this reason, convective events have been a fundamental object of study in the hydrological and meteorological field. Convective weather systems imply highly dynamic spatial and temporal processes which understanding remains a challenging issue.

The development of high resolution radar products in recent years (Einfalt et al., 2004; Seo et al., 2015; Thorndahl et al., 2016) has significantly advanced the observations of these phenomena. Due to its ability to better capture the spatial and temporal variability of the rainfall fields, high resolution weather radars allow for a more detailed analysis of the forming and movement of convective storms (Emmanuel et al., 2012; Liguori et al., 2012; Sebastianelli et al., 2013; Vulpiani et al., 2015).

Proper study of the spatial organization patterns and the temporal evolution of convective precipitation fields demands accurate tracking algorithms. There are two main types of methods that have been extensively developed to approach the storm tracking issue when using radar data; these are field and object based methods (Reyniers, 2008). The general idea of the former (also known as field trackers), is to compute a field of displacements/movement vectors over a continuous spatial grid by comparing two successive reflectivity/rainfall rate radar images. The images are often divided into grid blocks with an identical size. Then, the advection vectors of the blocks are obtained by using a certain measure of similarity between these two images. These methods have proven to provide a good 'global' motion estimation for the entire radar image field. For example, TREC (Tracking Radar Echoes by Correlation) method employs 'spatial' correlation coefficients to identify the movement of each of the pre-defined grid blocks (Rinehart and Garvey, 1978; Tuttle and Foote, 1990); VET (Variational Echo Tracking) and optical flow based methods derive the storm movement for each pixel location by determining the field of displacements that minimize the overall 'variation/difference' between two successive radar images (Bowler et al., 2004; Germann and Zawadzki, 2002; Laroche and Zawadzki, 1994; Laroche and Zawadzki, 1995; Pierce et al., 2012; Pierce et al., 2004).

In contrast to the field trackers, the object based tracking methods (or cell trackers) derive motion by first detecting storm objects (or

entities) following a particular identification process, and then, associating similar entities on two successive images based upon specific matching techniques. The final motion estimate is thus obtained by computing the associated object centroid displacements. The fact that storm entities can be identified and tracked provides this types of methods a great potentiality to deal with the smaller high intensity rainfall details shown in high resolution radar images (Lakshmanan et al., 2003). One of the well-known object based tracking algorithms in the literature is the TITAN (Thunderstorm Identification, Tracking, Analysis and Nowcasting) algorithm (Dixon and Wiener, 1993). TITAN was specifically built to identify, track and further forecast convective storms and has been routinely used during the last two decades (García-Ortega et al., 2009; Gascón et al., 2015; Goudenhoofdt and Delobbe, 2013; Han et al., 2009; Potts, 1993; Thorndahl et al., 2014; Yang et al., 2014). Despite its good performance in general, the settings of the original TITAN has been found unable to fully exploit the advantages that high resolution radar data sets can provide, as further explained in Section 3.1. For that reason, a new enhanced adapted version of the original TITAN algorithm able to better handle high resolution data is implemented in this work.

This paper is organized as follows. The data used for testing the new algorithm are described in Section 2. An overview of the TITAN algorithm, the deficiencies identified in its original structure and the treatments included to tackle them are introduced in Section 3, as well as the methodology applied to evaluate the new algorithm performance. Results are shown and discussed in Section 4. Lastly, in Section 5, main conclusions are presented, and future expectations of this work are proposed.

## 2. Data

High-resolution weather radar data provided by the Royal Meteorological Institute of Belgium (RMI) are used to develop and to evaluate the proposed storm tracking algorithm. Also known as the Belgian radar composite (Reyniers, 2008), the RMI radar product, is a composite of the Wideumont and Zaventem radar observations. Both the Wideumont and Zaventem radars, which location is depicted in Fig. 1, are single-polarization C-band Doppler radars that utilize 5 and 1 PPI elevation scanning strategies, respectively, to provide horizontal reflectivity data up to a range of 240 km. After applying a number of quality-control (QC) procedures, a pseudo-CAPPI composite reflectivity product at a height of 1500 m is produced at 5-min temporal and 529-m spatial resolution. For a detailed description of QC procedures that have been applied to the Belgian composite, readers are referred to Delobbe and Holleman (2006) and Foresti et al. (2016).

Radar reflectivity (dBZ) data from a total of 18 convective storm events over the entire Belgium area between 2012 and 2014 were employed in this paper. These events were selected based upon the convective storm classification methodology developed by Goudenhoofdt and Delobbe (2013). The selection process was focused on choosing storm events containing clear 'convective systems', namely with well-defined clustering structures with high reflectivity storm cores (Goudenhoofdt and Delobbe, 2013). Among these selected events, six of them (hereafter called, 'calibration events') were randomly chosen to tune the key parameters of the proposed tracking algorithm, and the rest of 12 events (hereafter, 'evaluation events') were used for evaluation. The list of events is summarized in Table 1.

## 3. Methodology

### 3.1. Original TITAN scheme

In the implementation of the original TITAN algorithm, convective storm entities are identified by defining contiguous regions that exceed a single reflectivity threshold value (i.e. 35 dBZ). Afterwards, TITAN employs a combinatorial method to match these identified entities

between two successive time steps, followed by an additional scheme able to handle splitting and merging processes.

The combinatorial method is formulated as an optimization problem, of which the aim is to minimize a cost (objective) function that indicates the similarity of storm entities. This cost function includes a couple of geometrical and physical features of storm entities: the velocity vectors (distance between entity centroids) and the difference in volume of the storm entities under consideration. In addition, some restrictions are imposed to narrow down and to ensure the feasibility of the solution domain. Constraints on the velocity parameter, and overlapping techniques (i.e. it is assumed that, if two storm entities overlap between two successive time steps, it is very likely that they are same entities at different time steps) are used for such a purpose.

Finally, the Hungarian numerical method is employed to solve this optimization (minimization) problem. Movement vectors representing the displacements of the centroids between each pair of 'matched' entities are eventually obtained. Afterwards, merging and splitting situations are handled using a 'short-term' forecast scheme. A merging situation occurs when the centroids of two or more storm entities at current time step are expected to be located within the same storm entity at the next time step. In contrast, a splitting situation takes place when a storm entity at the current time step is expected to overlap with the centroids of two or more storm entities at the next time step.

In spite of a generally good performance in tracking convective storm entities, a number of deficiencies can be identified in TITAN.

First, the single thresholding setting can effectively identify large convective systems, but appears to be insufficient to well isolate small-sized yet high-intensity storm details that can be observed in high resolution radar data images. In turn, single-thresholding approaches often lead to false merging problems in the identification process (Han et al., 2009; Handwerker, 2002), i.e. when two or more smaller independent storm entities are wrongly identified as a larger unique one.

Second, overlapping techniques may show difficulties in matching storm entities in fast moving storm cases since it is likely that no overlapping at all may occur (Dixon and Seed, 2014). Similarly, this lack of overlapping may also be largely magnified if high-intensity storm details are captured because smaller sizes are expected.

Third, the current criteria used to quantify the similarity of storm entities at successive time steps (merely dependent on the volume differences of storm entities) may not be able to cope with the highly dynamic behavior of small-sized storm entities and therefore it may not be able to provide satisfactory results when matching them.

Finally, motion estimates obtained from object centroid displacement may lack accuracy due to, first, object mismatching outcomes resulting from the use of overlapping techniques and the oversimplified criteria of similarity and second, the random centroid displacement problem (Han et al., 2009). This problem occurs in cases where a storm entity changes drastically its size or shape at two successive images due to the threshold application and it specially affects to larger size and false merged storm entities. As a consequence of these motion estimates inaccuracies, the algorithm may also show difficulties when handling merging and splitting situations, since possible unreliable motion estimates may be employed in the 'short-term' forecast scheme used to identify splits and mergers.

In order to overcome these known deficiencies, this study applies three main modifications to the original TITAN structure. Techniques that benefit from the advantages of having a high resolution data set, are incorporated. As shown in Table 2, first, a multi-threshold identification method is implemented to substitute the original single-threshold identification approach. Second, a more sophisticated matching scheme that avoids overlapping techniques by integrating a field tracker to optimize the combinatorial method, is constructed. It also makes use of a more comprehensive cost function and includes a new approach to deal with merging/splitting situations. Third, a more elaborated methodology to obtain motion estimates is developed.

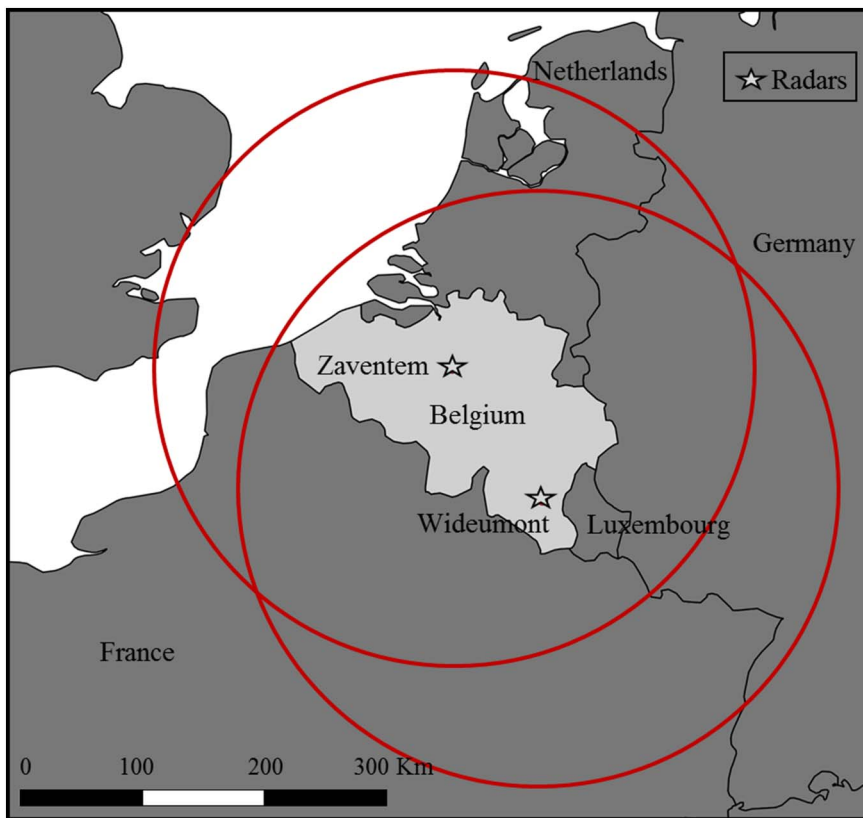These proposed changes aim at 1) creating an enhanced

Fig. 1. Location of the C-band weather radars.

identification methodology capable of better capturing convective storm details, 2) implementing a more robust matching scheme able to improve the matching association, and 3) obtaining more accurate motion estimates.

The implemented identification and matching schemes as well as the methodology used to obtain motion estimates are presented below. Right after, the methodology employed to evaluate the new tracking algorithm performance is introduced.

### 3.2. Two step heuristic multi-threshold identification method

During the last years, multi-threshold approaches have proven to provide added values to the storm identification process (Han et al., 2009; Handwerker, 2002; Hering et al., 2004; Johnson et al., 1998; Jung and Lee, 2015; Kyznarová and Novák, 2009; Novo et al., 2014). These methods enable better capturing the small-sized yet high-intensity storm details contained in high resolution data images. In this work, a two-step multi-threshold method is proposed to replace the original single-threshold method in the original TITAN. As shown in Fig. 2, this method consists of two main steps.

In the 1st step, the original radar image is decomposed into two different structural levels by using two different threshold values. Thus, by identifying contiguous regions of pixels that exceed 35 dBZ and 40 dBZ, two different sets of rainfall storm entities are initially isolated.

**Table 2**

Main modifications applied to the original TITAN algorithm.

|  | Original TITAN | New algorithm |
|---|---|---|
| Identification scheme | Single threshold method | Multi-threshold method |
| Matching schemes | Optimization method |  |
|  | Velocity constrains and overlapping techniques | Field tracker integration |
|  | Cost function |  |
|  | Volume and distance of storm entities | More measure indexes that quantify the similarity between rainfall entities |
|  | Merging/splitting method |  |
|  | Based upon a short-term forecast scheme | Based upon a field tracker |
| Motion estimation | Centroid based method | Mix of centroid and field tracker methods |

These sets are referred hereafter as large and small structural level entities (LSLE$_s$, SSLE$_s$). The selected thresholds are in accordance with the values used in Dixon and Wiener (1993) to define 'convective storms' and 'individual convective cells'.

In the 2nd step, a multi-threshold segmentation (MTS) technique is applied to each of the structural level entities with the purpose of, first,
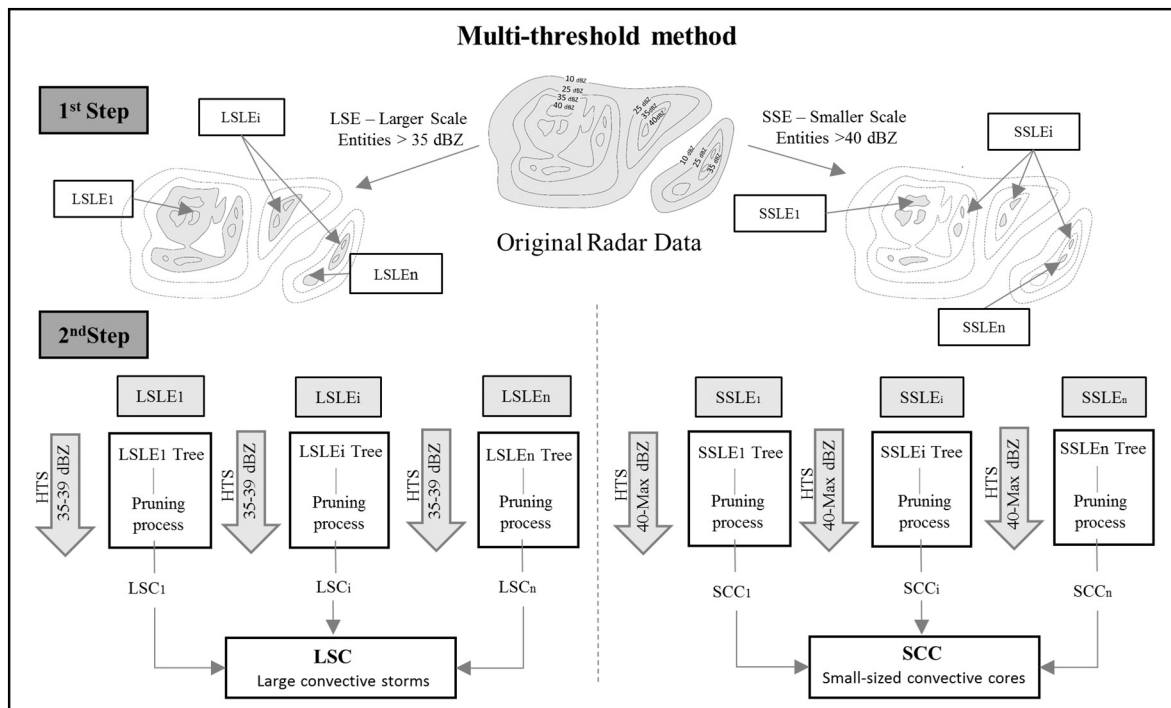
**Table 1**

Selected events employed for calibration and evaluation.

| Calibration events | Period | Evaluation events | Period | Evaluation events | Period |
|---|---|---|---|---|---|
| A | 10/8/2014 | 13:15–18:15 | 1 | 9–10/6/2014 | 22:30–03:45 | 7 | 10/6/2014 | 19:15–23:50 |
| B | 9/6/2014 | 13:30–19:40 | 2 | 9/6/2014 | 5:40–12:00 | 8 | 4/10/2014 | 15:30–18:20 |
| C | 27/7/2013 | 4:40–7:30 | 3 | 8/5/2013 | 14:00–22:30 | 9 | 3/10/2012 | 14:00–19:30 |
| D | 27/7/2013 | 17:50–23:30 | 4 | 10/7/2014 | 19:20–23:50 | 10 | 19/7/2014 | 16:50–20:30 |
| E | 23/9/2012 | 20:45–23:45 | 5 | 13/7/2014 | 13:00–20:00 | 11 | 20/5/2014 | 16:30–23:10 |
| F | 8/3/2014 | 13:15–20:00 | 6 | 27/6/2014 | 13:50–20:05 | 12 | 18/7/2014 | 17:30–22:40 |

**Fig. 2.** Steps of the multi-threshold method.

reducing the false merging problem in the LSLE$_s$, and second, isolating small high-intensity storm entities within the SSLE$_s$. The technique is implemented based on the hierarchical threshold segmentation (HTS) method proposed by Peak and Tag (1994). This method was originally created to isolate cloud features from satellite images by applying different 'grey shade' threshold levels to the original image in order to obtain a hierarchy of isolated regions. These isolated regions constitute a hierarchical tree structure, of which the nodes at a given level represent the isolated regions identified by a given threshold. Those nodes sharing the same parent are smaller isolated regions that split from a common larger region identified with a lower threshold. The tree structure is finally pruned by using decision tree analysis, which takes into account size and shape of the different nodes. The selected regions, obtained from the pruning process, are eventually collected and combined in a final segmented image. In this work, an adapted version of the HTS method is used to implement the MTS technique. A detailed explanation of the methodology followed is presented in Appendix 1, where the processes to construct and to prune the different structural level trees are described.

After applying the MTS technique, a final composite of storm entities, hereafter called large convective storms (LCS$_s$) and small-sized convective cores (SCC$_s$), are respectively obtained from the LSLE and SSLE groups.

### 3.3. Modified matching scheme

The original TITAN matching methodology is employed as a starting point to develop an enhanced scheme able to properly match the final composite of entities (LCS$_s$ and SCC$_s$) obtained by the new identification technique. As mentioned before, the drawbacks presented in the way TITAN matches storm entities may be largely magnified for SCC$_s$ due to its highly dynamic clustered nature. Furthermore, the short lifetimes associated to SCC$_s$ may lead to less robust global results if the original short-term forecast based scheme is used to handle mergers and splits, since more storm entities with short history or no history at all would be involved.

In order to tackle these deficiencies and with the intention to add

robustness when matching LCS$_s$ and SCC$_s$, the aforementioned three main modifications shown in Fig. 2 are implemented and incorporated to the original TITAN matching scheme. These are described hereafter.

#### 3.3.1. Field tracker integration

A storm field tracker, based on the optical flow technique, is applied to the high resolution data set and employed to optimize the process of constraining the feasible sets of solution when solving the matching problem. Optical flow was originally developed for image and video compression (Sadek et al., 2012) and has been largely used in the meteorology field as a field tracker (Bowler et al., 2006; Bowler et al., 2004; Cheung and Yeung, 2012; Wang et al., 2015). In this work, the multi-scale variational optical flow technique proposed by Brox et al. (2004) is used. Through comparing two successive radar images and imposing the so-called optical flow (OF) constraint together with other physical constraints (see Brox et al., 2004 and Wang et al., 2015 for details), one can obtain a 'smooth' field of movement vectors.

This field of movement vectors is then mapped with the image of the identified storm entities at time step $t$. For each entity, a single movement vector can thus be obtained by averaging the mapped movement vectors obtained from the field tracker. These averaged movement vectors are then used to extrapolate the entities one time step in order to provide their first guess location at time step $t + 1$. The possible set of entities at time step $t + 1$ that can be considered as potential candidates to be match to another entity at time step $t$, is limited by imposing a simple criterion as shown in Fig. 3. That is, the distance between the extrapolated entity's centroid ($\hat{C}_{t+1}$, in Fig. 3) and the possible candidate entities' centroids ($C_{t+1}$, $C'_{t+1}$) must be shorter than half of the maximum of the major axis lengths of the ellipses that best fit the entities under consideration (see $R$ and $R'$ in Fig. 3). As can be seen in Fig. 3, the distance $|\hat{C}_{t+1} - C_{t+1}|$ is shorter than R, while $|\hat{C}_{t+1} - C'_{t+1}|$ is larger than $R'$. Therefore, the entity with the centroid placed at $C'_{t+1}$ is discarded from the set of potential matching entities of the entity at $C_t$. This can effectively reduce the solution domain when solving the matching problem.
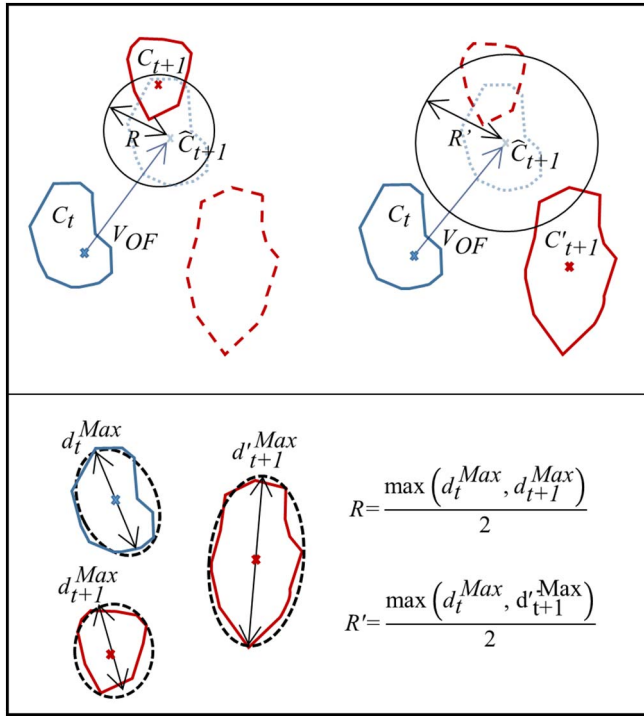
**Fig. 3.** Schematic representation of the matching criteria. $C_{t+1}$ and $C'_{t+1}$ are the centroids of two possible candidates to be matched with the object at $C_t$. $V_{OF}$ is the first guess movement of the object at $C_t$ obtained by the optical flow technique and used to obtain the extrapolated object at $\widehat{C}_{t+1}$. R is $\max(d_t^{Max}, d_{t+1}^{Max}) / 2$ and R' is $\max(d_t^{Max}, d_{t+1}'^{Max}) / 2$.

### 3.3.2. Cost function reformulation

As compared to the original TITAN, a more complex cost (objective) function, is proposed in this work. The cost function is based upon the forecast verification methodology proposed by Wernli et al. (2008) and also used in Shah et al. (2013), in which a number of measure indexes that quantify the similarity between objects, are employed as follows.

(a) The distance component ($L_c$). This measure accounts for the normalized distance of the centroids between the extrapolated and the possible matching entities. It is termed:

$$L_c = \frac{|\widehat{C}_{t+1} - C_{t+1}|}{R},\tag{1}$$

where $\widehat{C}_{t+1}$, $C_{t+1}$ and $R$ are the previously defined parameters in Fig. 3.

(b) The area component ($A_c$). This measure normalizes the size difference of the possible matching entities as follows:

$$A_c = \frac{|A_t - A_{t+1}|}{0.5[A_t + A_{t+1}]}.\tag{2}$$

Here, $A$ represents the areas of the entities under consideration.

(c) The structure component ($S_c$). This measure normalizes the difference of the structure index values for the possible matching alternatives:

$$S_c = \frac{S_t - S_{t+1}}{0.5[S_t + S_{t+1}]}; \quad S = \sum_{(i,j)\in A} \frac{R_{i,j}}{R^{max}}.\tag{3}$$

The structure index, $S$, is obtained by scaling the total amount of reflectivity values ($R_{i,j}$) with its maximum value ($R^{max}$). The structure component, $S_c$ describes how widespread or peaked the reflectivity spatial organization is within the area of the entity (Wernli et al., 2008).

(d) The eccentricity component ($E_c$). This measures represents the

eccentricity difference between the possible matching entities by:

$$E_c = |e_t - e_{t+1}|\tag{4}$$

where $e$ denotes the eccentricity of the ellipses that best fit the candidates to be matched.

These indexes (components) are then combined into a single cost value between two possible matching entities:

$$Cost = w_1\, L_c + w_2\, A_c + w_4\, S_c + w_4\, E_c.\tag{5}$$

where $w_1$, ..., $w_4$ respectively represent weighting factors for the four defined components. Based upon sensitivity analysis, the following values are assigned to these weighting factors: $w_1 = 0.4$, $w_2 = 0.25$, $w_3 = 0.25$ and $w_4 = 0.1$. Similarly to the original TITAN algorithm, the Hungarian method is employed to solve a cost function matrix built on the basis of all the possible entity combinations (cost values) at two successive time steps. After the matching problem is solved, a first association of matched entities can be obtained.

### 3.3.3. Merging and splitting situations handling

Special attention has been given to storm merging and splitting processes in the meteorological field during the last decades. It is well known, that splitting and merging processes generally occur during extreme weather episodes and therefore play an important role in convective dynamics (Carbunaru et al., 2013; Ćurić et al., 2009; Spiridonov et al., 2010). For this reason, special emphasis is given in this work on improving the handling of these phenomena.

An alternative scheme that makes use of the first movement guess ($V_{OF}$ in Fig. 3), is proposed to handle merging and splitting situations as opposed to the 'short-term' forecast based methodology employed in the original TITAN. As shown in the schematic example of Fig. 4, after solving the combinatorial matching problem defined in Section 3.3.2, a first group of matched storm entities are associated (i.e. A → G, B → H, E → J, N → S, O → U, Q → V, R → W). Then, in order to identify splitting situations, entities at time step $t + 1$ that have already been matched (G, H, J, S, U, V and W) are discarded from the matching problem. Then, the combinatorial optimization problem is again solved and new matching entities obtained (E → K, O → T). Since one entity may split into more than two entities (e.g. E), the matching problem is solved iteratively, until no more splitting associations are found (E → L). Afterwards, merging situations are handled. For storm entities that remain unmatched at time step $t$ (C, D, F, M and P) and those entities at time step $t + 1$ that do not take part in a splitting situation (G, H, I, S and V), the matching problem is again solved. Similarly, the matching problem is iteratively solved until no more merging cases are found (C → H, P → V, M → S, D → H).

Similarly to the original TITAN algorithm, the above approach handles merely merging and splitting situations, individually. In other words, it cannot deal with the case when splitting and merging situations occurs simultaneously. To cope with these situations, an additional matching step is incorporated. In this last step, all the entities at both time steps $t$ and $t + 1$ can be potentially associated as long as they have not been matched together in any of the previous steps. Similarly, the combinatorial optimization problem is solved iteratively until no more of these 'simultaneous' situations are identified in such a way that, all the splits involved in merging situations (O → S and R → V) and merging entities that take part on splitting situations (P → U), are eventually associated. Once the splitting and merging matching last step completes, the remaining unmatched entities at time step $t$ and $t + 1$ are respectively defined as 'deadly' entities (e.g. F), and 'newly born' entities (e.g. I).

In the example of Fig. 4, after all the steps involved in the matching scheme are finished, a total of six different storm entity matching situations are captured for both LCS$_s$ and SCC$_s$: continuity, splitting, merging, simultaneous splitting/merging, new born and dead matching situations.
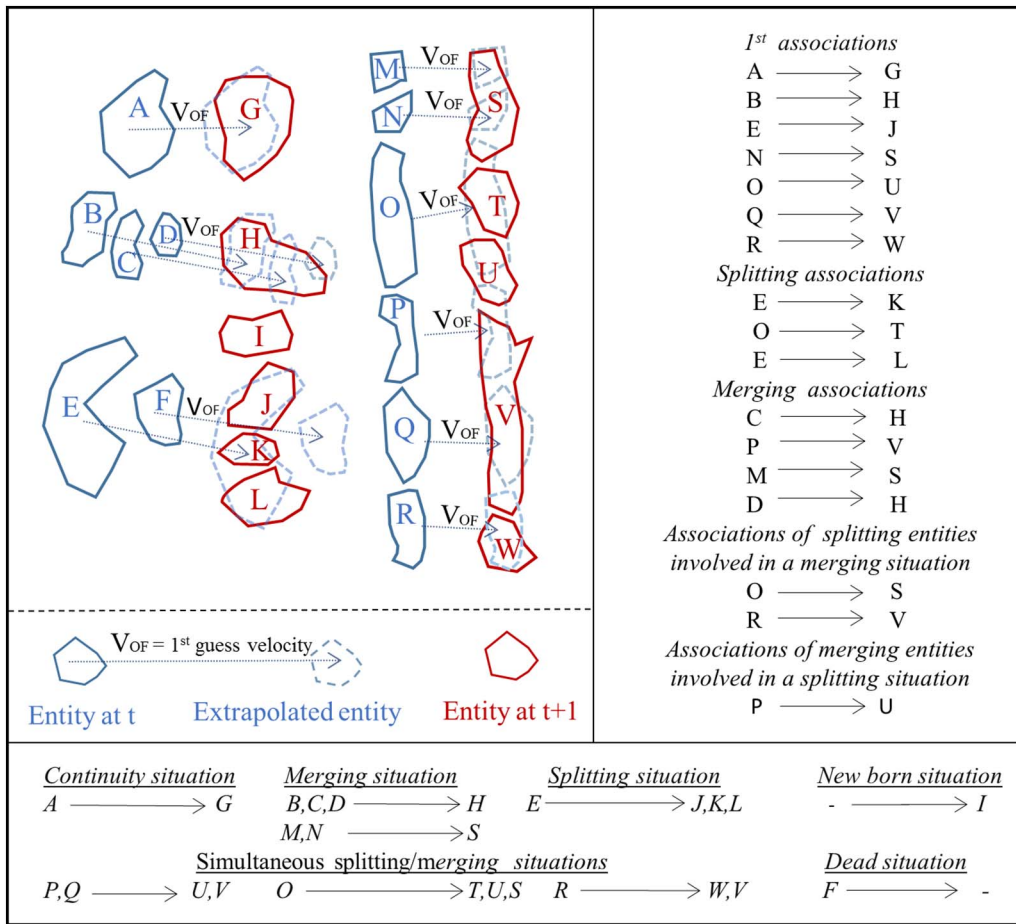
## 3.4. Motion estimation

After the matching process, it is critical to assign correct motion estimates for each storm entity. For the LCS$_s$, the previously defined first guess movement ($V_{OF}$), is assigned to each entity. As for SCC$_s$, motion is obtained by taking into account the corresponding next time step matched entities obtained by the matching scheme. As shown in

Fig. 5, depending on the particular matching situation, different approaches are employed. A combination of centroid displacements, weighted with the areas of the storm entities involved in the process, is employed for continuity, splitting and merging situations (Fig. 5a, b, c) to derive their motion vectors (i.e. $V_A$, $V_C$, $V_F$ and $V_G$).

For splitting entities that also contribute in merging processes (Fig. 5d), a different centroid displacement based approach is
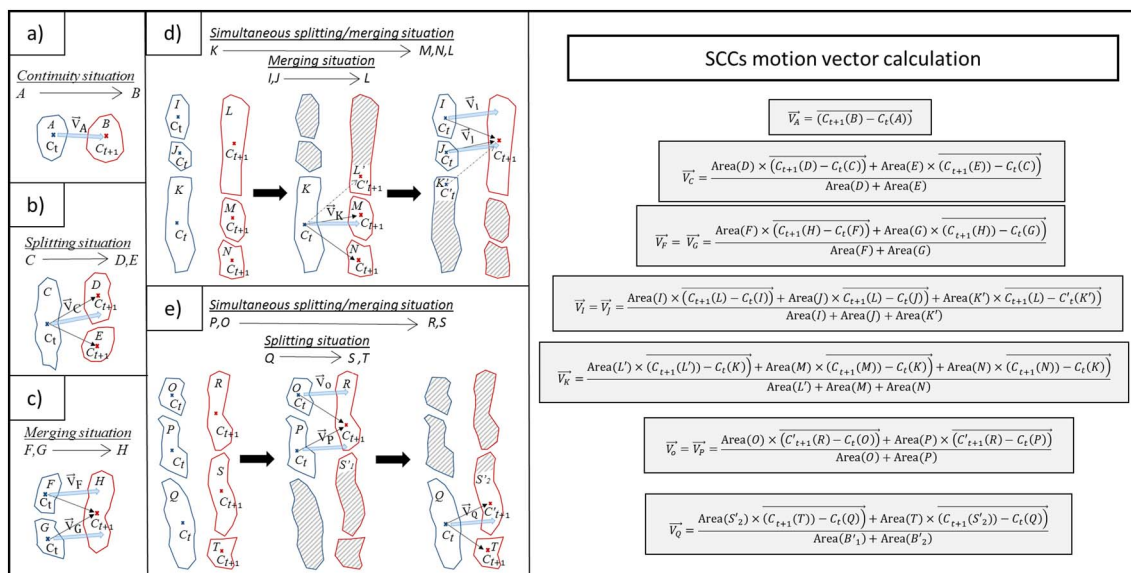


**Fig. 5.** Schematic representation of the motion estimation calculation for the different SCC matching situations.

employed. First, the purely splitting/merging situations assignments are identified $((I, J) \rightarrow L; K \rightarrow (M, N))$. Then, by means of accounting for the geometric characteristics and location of the entities that originally take part of the purely merging process (I and J), a remaining area $(L')$ within the merger object (L) is defined, and a new centroid $(C'_{t+1} (L'))$ can be derived and used to compute the final motion vector for the split object (i.e. $V_K$). As for the entities that originally take part in the merging (I and J), a similar approach is utilized. A remaining area $(K')$ in the split object (K) can be defined and the new centroid $(C'_t (K'))$ is used to calculate the final motion vectors $V_I$ and $V_J$. Finally, when entities involved in a merging situation take also part of a splitting process (Fig. 5e), the same methodology accounting for proportional contribution areas and new appearing centroids is again used to calculate and assign motion vectors to each of the storm entities ($V_O$, $V_P$ and $V_Q$).

### 3.5. Evaluation of the new tracking algorithm

The assessment is conducted by comparing the proposed tracking algorithm against the original TITAN by using the above defined 'evaluation events' (see Table 1). The different evaluation methods employed for the comparison are presented below.

### 3.5.1. Visual inspection of storm entity identification

Visual inspection is one of the most commonly used approaches when evaluating storm identification methods in literature (Fiolleau and Roca, 2013; Han et al., 2009; Handwerker, 2002; Johnson et al., 1998, Kober and Tafferner, 2009; Liu et al., 2014; Shah et al., 2013; Wilson et al., 2004). In this work, visual inspection is used first, to assess how well the new identification method handles the false merging problem (this is done, by comparing the LCS$_s$ with the storm entities identified by the original TITAN), and second, to evaluate its ability to isolate SCC$_s$ within convective systems.

### 3.5.2. Assessments of storm entity matching situations

Comparing the results obtained automatically from the matching scheme with those given by visual investigation is a common approach to assess storm entity matching performances (Handwerker, 2002; Johnson et al., 1998; Jung and Lee, 2015; Kober and Tafferner, 2009; Kyznarová and Novák, 2009; Shah et al., 2013; Shimizu and Uyeda, 2012). For such a purpose, a contingency table approach is employed to quantify the ability of the new algorithm to 'correctly' associate the different storm entity matching situations defined in Fig. 5.

A two-class contingency table based on a yes/no binary classification problem is used in this work. It reflects the occurrence and non-occurrence of what is previously defined as an 'event' in both the algorithm results and the visual inspection observations. As seen in Table 3, four different outcomes can be obtained from the binary classifier ($n_{hits}$, $n_{failure}$, $n_{false\ alarm}$ and $n_{correct\ rejection}$): $n_{hits}$ occurs when both the algorithm and the visual inspection equally capture an 'event'; $n_{failure}$ is obtained when human criteria identifies an 'event' that the new tracking algorithm is not able to detect; $n_{false\ alarm}$ is produced when the algorithm consider an 'event' that is not identified as such by visual inspection; $n_{correct\ rejection}$ is accounted when both the algorithm and the observation rejects the occurrence of an 'event'.

In this case, an 'event' is defined as a matching association situation.

**Table 3**
Two class binary contingency table.

|            | Observed |          |
|------------|----------|----------|
| Algorithm  | Yes      | No       |
| Yes        | $n_{hits}$ | $n_{false\ alarm}$ |
| No         | $n_{misses}$ | $n_{correct\ rejection}$ |

Once the contingency table is built, categorical statistics from its elements are calculated. Probability of detection (POD) (also called hit rate (HR)), false alarm ratio (FAR) and critical success index (CSI) scores are computed in order to quantify the ability of the new matching scheme to correctly match different matching association situations at two successive time steps. Continuity, new born, dead and merging/splitting situations (simultaneous merging/splitting situations included) are differentiated in the calculation of the scores. POD, FAR and CSI score values are calculated as follows:

$$POD = HR = \frac{n_{hits}}{n_{hits} + n_{misses}}, \quad CSI = \frac{n_{hits}}{n_{hits} + n_{misses} + n_{false\ alarm}}, \quad FAR$$

$$= \frac{n_{false\ alarm}}{n_{hits} + n_{false\ alarm}}. \tag{6}$$

In order to evaluate the new matching scheme performance, a comparison with the original TITAN scheme scores is carried out.

### 3.5.3. Statistical analysis of the overall tracking algorithm performance

In order to assess the overall performance of the proposed tracking algorithm against the original TITAN, a deterministic persistence forecast approach is employed. Hence, by neglecting growth and decay processes, the forecast performance is expected to improve when more accurate motion estimates are used (Germann et al., 2006). In this approach, the storm entities obtained from the respective identification schemes are extrapolated by using the motion vectors derived from their corresponding matching schemes.

A schematic representation of the TITAN based forecast scheme (TFS) and the new algorithm based forecast scheme (NFS) is illustrated in Fig. 6. As can be seen, for the TFS case, the forecasted image is obtained straightforward. On the contrary, for the NFS case, since different motion estimates are obtained at two different structural levels, relative movements between SCC$_s$ and LCS$_s$ are expected. As a consequence, overlapped and missing pixel values may appear in the forecasted image. To tackle this issue, when more than one values are obtained at the same pixel location, the maximum of them is used while when no pixel values are assigned, inverse distance squared weighted interpolation method is conducted to fill the missing values. The forecasted images, obtained from TFS and NFS respectively, are then compared with the observed radar images to evaluate the forecast performances from 1 to 4 time steps ahead (i.e. 5–20 lead time minutes). The performance measures used include:

1) (Spatial) correlation coefficients

A traditional statistical pixel by pixel verification method is applied to compare the general performances of the forecast schemes. Correlation coefficients are computed to measure the correspondence
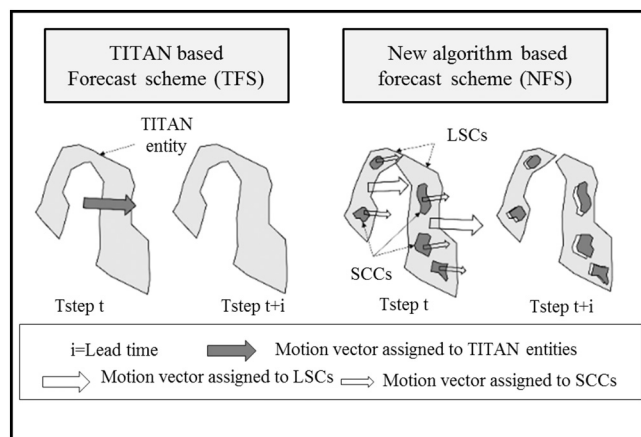


**Fig. 6.** Schematic representation of the TFS and NFS.

between forecast and observed pixel values. Only pixels that belong to the extrapolated objects are considered in the calculations.

2) ROC (Receiver Operating Characteristic) curve analysis

The new algorithm performance is evaluated at different intensity levels by comparing the skills of the two forecast schemes (TFS and NFS) at different reflectivity threshold levels. This is done by the ROC (Receiver Operating Characteristic) curve analysis. This analysis involves a pixel by pixel evaluation that makes use of the contingency table approach (Table 3). The contingency table is used here to determine the skills of the forecast scheme to discriminate between what is considered as an 'event' and 'no event'. In this case, an event occurs when the reflectivity of the pixel under consideration is larger than an established reflectivity threshold. A contingency table based on yes/no forecasted and yes/no observed events can be then constructed. As with correlation coefficient calculations, all the pixels belonging to forecasted entities are considered to calculate the final number of hits, false alarms, misses, and correct rejections. These values, are then used to calculate a hit rate (HR) as defined in Eq. (6) and a false alarm rate (FR) as follows:

$$FR = \frac{n_{false\ alarm}}{n_{false\ alarm} + n_{correct\ rejection}}. \tag{7}$$

Finally, the ROC curve is constructed by plotting against each other all the pairs of HR and FR values obtained when the established reflectivity threshold changes. Reflectivity threshold values from 35 to 54 dBZ are used in this work.

## 4. Results and discussion

### 4.1. Storm entity identification

Visual inspection of the new identification scheme performance is carried out for all the 'evaluation events'. The results reveal that, the way the false merging problem is handled when identifying $LSS_s$, gives more satisfactory results in comparison with the single threshold approach used by the original TITAN scheme. Two examples of this are shown in Fig. 7. As can be seen, entities weakly connected and identified as a single object by TITAN are now separated forming clusters of independent $LCS_s$. The ability of the identification technique to separate clusters of $SCC_s$ within $LCS_s$, is also confirmed. Examples of how these

entities are isolated are shown in Fig. 8.

### 4.2. Storm entity matching situations

Matching situations captured from the 12 'evaluation events', are gathered together and used to calculate the final *POD*, *FAR* and *CSI* score values. Since this is a highly time-consuming task, only hours that show important convective activity, are taken into account. The scoring results obtained for both the new and original TITAN algorithm are given in Tables 4 and 5, respectively.

A first look at the results on Table 4 indicate a good fit between the visual inspection and what the new algorithm captures for both $SCS_s$ and $LSS_s$. When it comes to comparing these results with those obtained by the original TITAN algorithm, it should be noticed that the results obtained in Table 5, can only be contrasted with the ones obtained for $LCS_s$, as they deal with entities of similar structural level. The increased *POD* and *CSI* values obtained by the new matching scheme, reflect an important enhancement in the matching performance with respect to the original TITAN. Moreover, its lower FAR values also prove better capability to avoid wrong entity associations. As for $SCC_s$, although the results cannot be directly compared with TITAN, the score values obtained indicate that, despite a presumably greater difficulty when matching $SCC_s$, due to their higher dynamic changing nature with respect to the entities identified by TITAN, a similar performance is accomplished and therefore good matching skills are achieved.

### 4.3. Tracking algorithm performance

#### 4.3.1. Correlation coefficients

Correlation coefficients between motion estimate based forecasts and observed pixel values are computed for the whole duration of the 'evaluation events' for lead times up to 20 min. Results for a lead time of 5 min are presented in Fig. 9. Percentage change between the correlation coefficient means of NFS with respect to TFS is computed for all the leading times. Results for the 12 evaluation events are plotted in a box plot and presented in Fig. 10. Figs. 9 and 10 show a significant mean percentage increase (20.1%) in correlation coefficients for NFS in comparison with TFS. This reveals that, the new tracking algorithm performs better and therefore produces more reliable motion vectors.

#### 4.3.2. ROC (Receiver Operating Characteristic) curve analysis

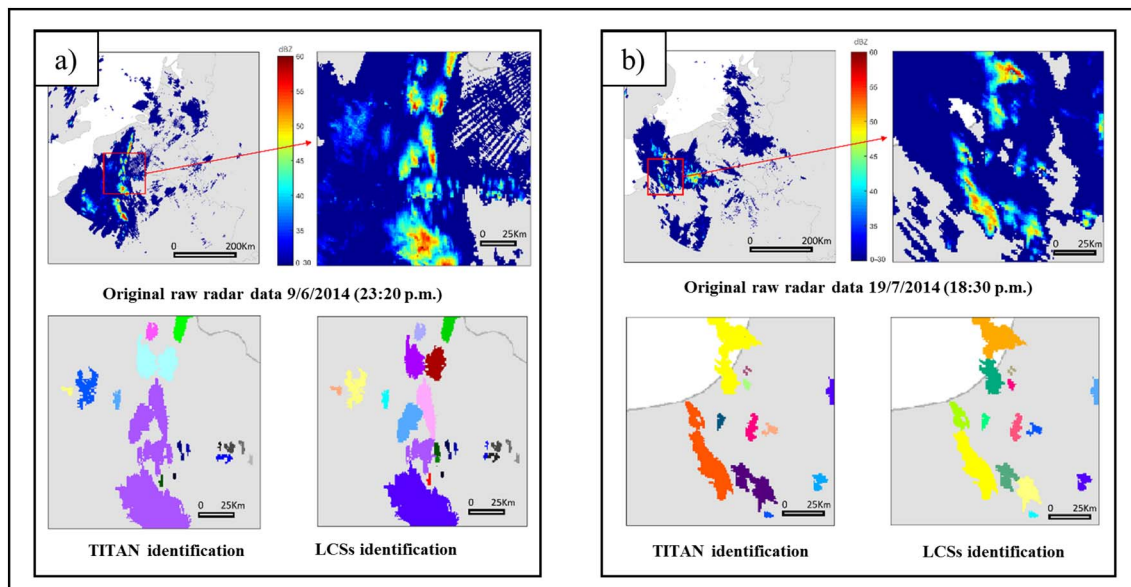The resulting ROC curves constructed from both NFS and TFS, for



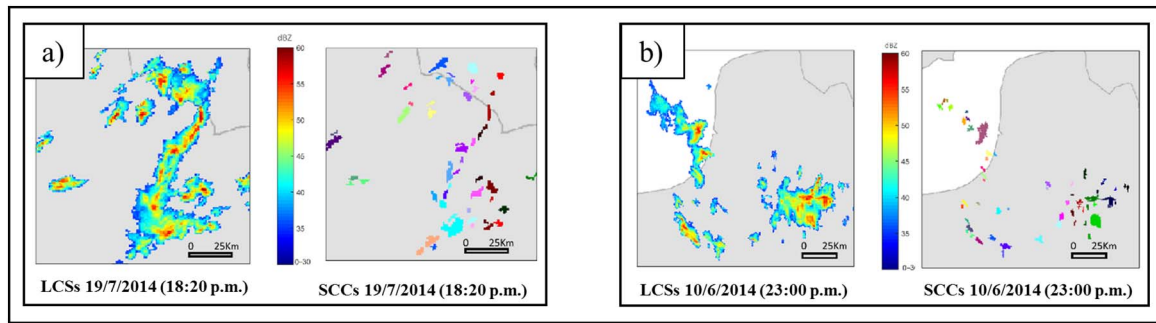**Fig. 7.** TITAN Vs LSS identification results for event 1 at 23:20 pm (a) and event 10 at 18:30 pm (b).

**Fig. 8.** SCCs identification results for events on event 10 at 18:20 pm (a) and event 7 at 23:00 pm (b).

**Table 4**
POD, FAR and CSI scores for LCSs and SCCs.

| LCS | $n_{hits}$ | $n_{failure}$ | $n_{false\ alarm}$ | POD | FAR | CSI |
|---|---|---|---|---|---|---|
| Continuity | 543 | 30 | 27 | 0.948 | 0.047 | 0.905 |
| New | 129 | 9 | 21 | 0.935 | 0.140 | 0.811 |
| Dead | 156 | 12 | 15 | 0.929 | 0.088 | 0.852 |
| Splitting/merging | 156 | 15 | 18 | 0.912 | 0.103 | 0.825 |

| SCC | $n_{hits}$ | $n_{failure}$ | $n_{false\ alarm}$ | POD | FAR | CSI |
|---|---|---|---|---|---|---|
| Continuity | 1008 | 90 | 63 | 0.918 | 0.059 | 0.868 |
| New | 378 | 42 | 54 | 0.900 | 0.125 | 0.797 |
| Dead | 300 | 39 | 42 | 0.885 | 0.123 | 0.787 |
| Splitting/merging | 399 | 54 | 69 | 0.881 | 0.147 | 0.764 |

**Table 5**
POD, FAR and CSI scores for TITAN identified entities.

| TITAN | $n_{hits}$ | $n_{failure}$ | $n_{false\ alarm}$ | POD | FAR | CSI |
|---|---|---|---|---|---|---|
| Continuity | 450 | 48 | 60 | 0.904 | 0.118 | 0.806 |
| New | 120 | 18 | 24 | 0.870 | 0.167 | 0.741 |
| Dead | 105 | 12 | 15 | 0.897 | 0.125 | 0.795 |
| Splitting/merging | 96 | 15 | 18 | 0.865 | 0.158 | 0.744 |

each of the 12 'evaluation events', are obtained for the different lead times. Fig. 11 shows an example of ROC curve for the event of 20/05/2014 and lead time of 5 min. Then, *HR* and *FR* scores for three representative reflectivity threshold levels (37, 45 and 50 dBZ) are extracted for both NFS and TFS ROC curves.
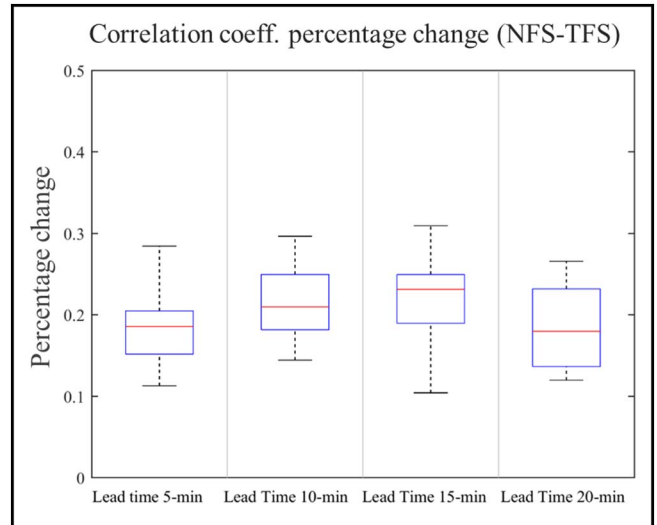


**Fig. 10.** Box plots representing the percentage increase between correlation coefficient means of NFS with respect to TFS for 4 leading times and the 12 evaluation events.

Better forecast performances provide higher *HR* and lower *FR* scores. A no skill forecast would lay on the diagonal line represented in Fig. 11. The area under the ROC is used as a verification measure of the forecast skill. An area equal to 1 indicate a perfect score whereas 0.5 implies no skills. The results of the areas and the extracted *HR* and *FR* scores for a lead time of 5 min, are shown in Fig. 12 for the 12 'evaluation events'. Percentage change between areas under the ROC curves
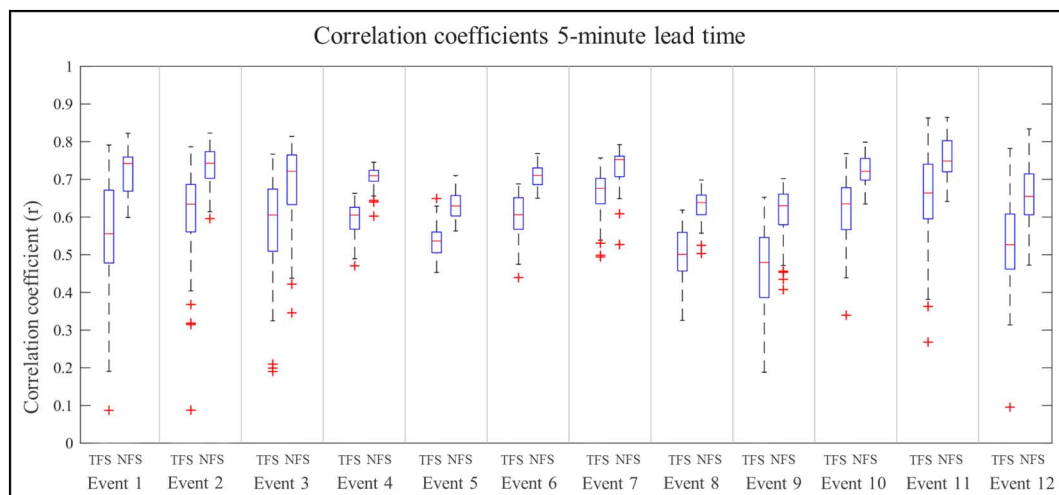


**Fig. 9.** Box plots representing correlation coefficients between motion estimates based forecasts (NFS and TFS) and observed pixel values for the 12 evaluation events and leading time of 5 min.
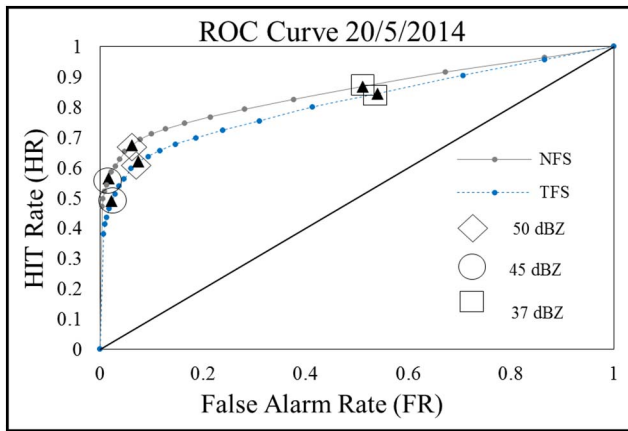
Fig. 11. ROC curve associated to 20/5/2014 event.

of the NFS and TFS, as well as the percentage change between *HIT* and *FR* scores at the three reflectivity threshold levels, are calculated for all the lead times. Results for the 12 'evaluation events' are plotted in Fig. 13.

The higher areas under the ROC curves shown in Fig. 13a corroborate that the performance increases when motion vectors given by the new algorithm are employed. Regarding the extracted *HR* and *FR* scores (Fig. 13b and c), the mean *HR* percentage increases with 25.1% (50 dBZ), 14.2% (45 dBZ) and 3.3% (37 dBZ), and the mean *FR* percentage decreases *FAR* values of 26% (50 dBZ), 20.9% (45 dBZ) and 9.6% (37 dBZ). These results exhibit a substantial improvement in the *HR* and *FR* scores for high reflectivity thresholds, as well as a not negligible enhancement at the lower reflectivity thresholds. The FR, moreover noticeable improves for higher lead times, especially at the highest threshold levels.

The presented results reveal, the ability of the new algorithm to better handle and track high reflectivity storm details. It also shows the added value of using, for LCS$_s$, motion vectors derived from the optical flow technique, which confirms the ability of the optical flow technique
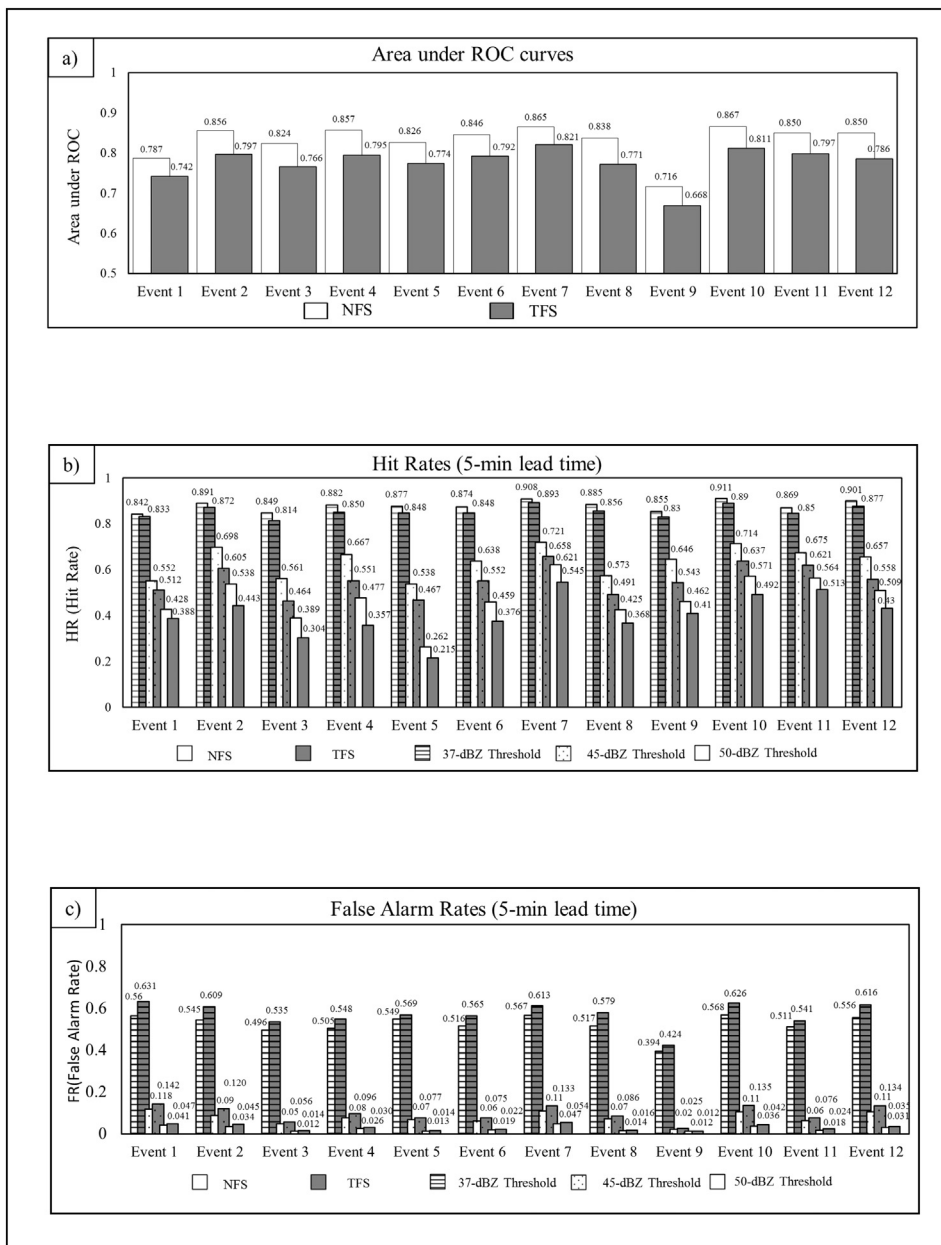


Fig. 12. Areas under ROC curves (a), ROC HIT rates (b) and FR rates (c) at 37,45, 50 dBZ thresholds for NFS and TFS.
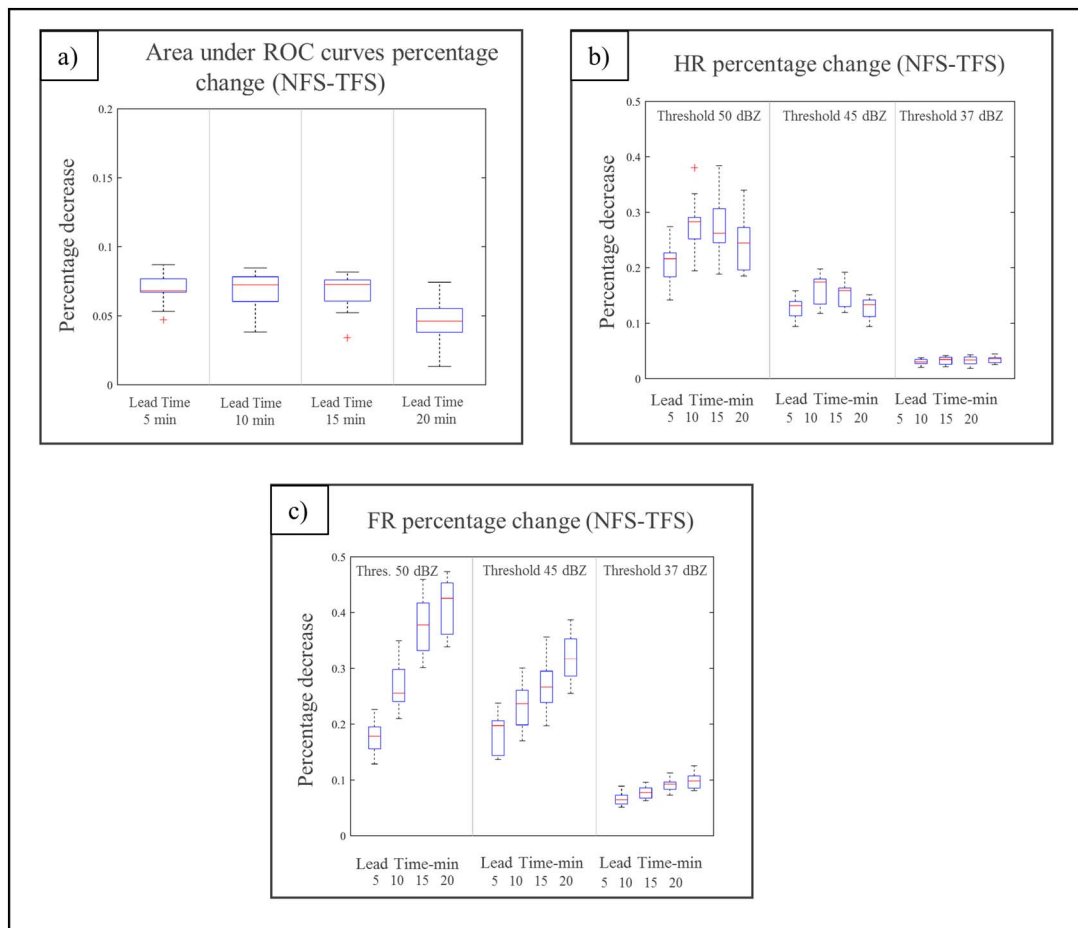
**Fig. 13.** Box plots representing the percentage change of the NFS with respect to the TFS for 4 leading times and 12 evaluation events. Percentage increase between areas under ROC curves (a). Percentage increase between HIT rates at 37, 45, 50 dBZ thresholds levels (b). The percentage decrease between FAR rates at 37, 45, 50 dBZ thresholds levels (c).

to provide better global motion estimates.

## 5. Conclusions and future work

In this paper, an enhanced object-based storm tracking algorithm, which can better capture convective storm details, is presented. The proposed algorithm is developed based upon the TITAN algorithm, which has been the basis of convective storm tracking and nowcasting systems for many national meteorological services. However, with the development of radar technology, higher resolution radar observations allowing to observe more small-sized rainfall features became available. In this regards, the original TITAN is found to be unable to satisfactorily handle these spatial storms details. A number of modifications are therefore proposed and implemented herein, on the basis of TITAN, to improve its tracking capacity. The main modifications are as follows:

1. A two-step multi-threshold storm identification method is proposed to replace the single-threshold scheme in the original TITAN. In the new method, the spatial structure of a convective storm system is firstly separated into large and small-structure levels of entities (i.e. LSLEs and SSLEs, respectively). A multi-threshold segmentation technique is then applied to each structural level to further identify large convective storm (LCSs) and small-sized convective cores (SCCs) entities.
2. A new matching scheme is constructed to better establish the (temporal) association of the identified entities between two successive time steps. In this new scheme, the following modifications are implemented:
   a. A field storm tracker, based upon a multi-scale variational optical

flow technique (*OF*), is incorporated. The *OF* field tracker is used to conduct an initial motion estimation of storm entities to improve both the efficiency and the accuracy of the matching problem.
   b. The cost function used to quantify the similarity of matching storm entities is reformulated. For such a purpose, more geometric and physical features of storm entities are considered as compared to the original cost function.
   c. An alternative approach to handle storm entity merging and splitting situations is implemented. In this approach, as opposed to the original TITAN algorithm, cases where splitting and merging situations occurs simultaneously are included.
   d. A more sophisticated methodology, which benefits from the fact that more storm matching situations can be captured, is employed to estimate the motion vectors assigned to each identified storm entity.

High-resolution radar data from a total of 18 convective storm events in Belgium during the period of 2012–2014 have been employed to test the proposed modifications. Six of these events were randomly selected and used to conduct the analysis of the key parameters of the proposed algorithm, and the other 12 events were used to evaluate its performance. As compared to the original TITAN algorithm, the following improvements are obtained:

1. The two-step identification scheme shows a better capacity of isolating and capturing small-sized yet high-intensity storm cell clusters (SSC$_s$) as well as of reducing the false merging problem for LCS$_s$.
2. Object matching association is improved. When comparing POD,

FAR and CSI scores for LCSs matching situations with respect to TITAN, an average percentage increase of 5.3%, 5.3% and 8.69% is respectively obtained. For $SCC_s$ matching situations, a direct comparison with TITAN cannot be established. Nevertheless, and taking into account the already assumed good performance of TITAN, the similar scores obtained for $SCC_s$ with respect to TITAN, indicates a good matching performance considering the highly dynamic nature of these high intensity small entities. It should also be remarked the ability of the new algorithm to better track splitting and merging situations.

3. More reliable motion estimates are obtained by the new algorithm. It is shown that by capturing $SCC_s$ movements and deriving $LCS_s$ motion vectors by using the *OF* technique, better motion fields are obtained with respect to TITAN. This is reflected when the performance of the simple deterministic persistence forecast schemes constructed on the basis of their respective motion fields are compared. Average percentage increase of 20.1% is found when comparing correlation obtained from the NFS (new algorithm based forecast scheme) with the ones from TFS (TITAN based forecast scheme). Average HR percentage increase values of 25.1% (50 dBZ), 14.2% (45 dBZ) and 3.3% (37 dBZ), and average FR percentage decreases values of 26% (50 dBZ), 20.9% (45 dBZ) and 9.6% (37 dBZ) are obtained when comparing the ROC curves constructed from the NFS with respect to the ones from the TFS. These results also reflect that, as several authors have already pointed out, object based tracking methods and field trackers should not be seen as adversaries but as methods that may complement each other (Han et al., 2009; Pierce et al., 2012; Dixon and Seed, 2014).

Despite the promising results obtained in this work, additional validation of the proposed algorithm is necessary. In order to assess the robustness of the algorithm, its performance must be further evaluated with more storm events and different radar data sets. In addition, the proposed matching scheme that integrates object based tracking with a field tracker can be further explored. More specifically, apart from optical flow techniques, other field trackers or image matching methods may be worth investigating. For example, some coherent structure recognition methods (based upon POD and PCA techniques) that have been widely used in computer vision (Liu et al., 2010; Verbeke and Vincent, 2007), can be tested and compared with the optical flow field tracker used in this work.

In terms of future applications, the proposed algorithm is intended to be used as the basis for characterizing the spatial-temporal properties of convective storms. This in turn can be used to improve the forecasts of convective storms (e.g. Dixon and Wiener, 1993; Mueller et al., 2003; Berenguer et al., 2005; Novak, 2007; Kober and Tafferner, 2009; Liang et al., 2010; Zahraei et al., 2012), the generation of stochastic spatial-temporal rainfall fields (e.g. Willems, 2001; McRobie et al., 2013; Niemi et al., 2014) and the temporal interpolation of high-resolution radar images (Nielsen et al., 2014; Wang et al., 2015).

In addition, the features of the identified convective storm entities can be used to map or to cross compare with other atmospheric variables (e.g. temperature and soil moisture) or with those resulting from other numerical models such as cloud resolving models (CRM) (Ćurić and Janc, 2011a; Ćurić and Janc, 2011b). It is worth noting that, in the past few years, CRMs have gained an increasing interest by the meteorological research community due to their ability to reproduce convective cloud dynamics and to generate rainfall fields (Ćurić and Janc, 2012; Ćurić et al., 2007; Wilhelmson and Klemp, 1978; Wu and Li, 2008). In this regard, the enhanced tracking algorithm can be used to evaluate given CRM's abilities to reproduce specific storm characteristics and to simulate the splitting/merging processes by means of comparing the modeled outputs with radar observations (Karacostas et al., 2013; Caine et al., 2013). In this way, deficiencies in the model physics might be found and changes can be proposed.

### Acknowledgements

### Appendix 1

In this appendix, a description of the MTS technique introduced in Section 3.2 is presented. In order to better explain the whole methodology, and the different settings involved in the way the MTS technique is applied at each of the two structural level groups, the two representative schematic examples shown in Figs. A.1 (LSLE1) and A.2 (SSLE1) are used as a support.

First, the thresholds used to obtain the hierarchy of isolated regions are defined. The whole range of thresholds from 35 dBZ to 39 dBZ is applied to the $LSLE_s$. Meanwhile, values from 40 dBZ to the maximum reflectivity value of the considered entity are used for the $SSLE_s$.

Second, as represented in Fig. A.1 and A.2, tree structures with nodes representing isolated regions formed by contiguous pixels (connectivity = 4) that exceed the different established thresholds are obtained for each of the structural level entities. Then, all the isolated region areas values are computed and assigned to their representative nodes. Later, these nodes are categorized. The lower reflectivity threshold level region, placed on the top of the tree, is defined as 'roof node' (nodes $1L_1$ and 1S1 in the example). Nodes with two or more children are classified as a 'splitting nodes' (e.g. nodes $2L_1$, $8L_1$, $6L_1$, $3S_1$, $4S_1$ or $13S_1$). On the contrary, nodes with no children are labeled as 'leaf nodes' (e.g. nodes $5L_1$, $14L_1$, $18L_1$, $12S_1$, $7S_1$ or $26S_1$). The remaining ones are categorized as 'simple nodes' (e.g. nodes $4L_1$, $9L_1$, $6S_1$ or $10S_1$).

Finally, a heuristic methodology for pruning the tree structures, based upon the category and the area of the nodes, is implemented. A flow diagram that reflects the whole process is shown in Fig. A.3.

As shown in Fig. A.3, the pruning method is a repetitive process that starts at each of the 'leaf nodes' and goes towards the upper parts of the tree until a pruning decision is made. Whenever a node with a 'simple node' as a parent is found (e.g. nodes $14L_1$, $13L_1$, $5L_1$, $18L_1$, $23S_1$ or $16S_1$), one level upwards movement is done. On the contrary, when a node, which parent is a 'splitting node' or the 'roof node', is found (e.g. nodes $9L_1$, $18L_1$, $21S_1$, $24S_1$ or $6S_1$) the possibility to prune the tree in that location is considered. Two conditions, with different settings for the two defined structural levels, has to be met at the same time to perform the pruning:

1) At least two children/siblings of the considered 'splitting or roof node' are 'significant nodes'. In a LSLE tree case, a node is considered as 'significant' when its area is larger than 10% of the correspondent 'roof node'. As for a SSLE tree case, the area of the node has to be larger than an established threshold and to exceed the 10% of its parent area. A threshold area of 3 km$^2$ is chosen.
2) The sum of the areas of all the significant nodes must be larger than 50% of its parent area.

When the conditions are not met [e.g. nodes $4L_1$, $5L_1$, $11L_1$ or $12L_1$ (first condition is not met since only one of the siblings is 'significant' ($> 10\%$ of 'roof node' area)); nodes $24S_1$ or $25S_1$ (the first condition is not met since none of the siblings are 'significant' (Area $< 3$ km$^2$)); $11S_1$, $12S_1$ (second condition is not met, since the sum of both areas is lower than 50% of the parent area)], we keep moving levels upwards until the next node where the pruning can be considered is found. The process is repeated until the conditions are fulfilled (e.g. nodes $2L_1$, $9L_1$, $10L_1$, $7L_1$, $7S_1$, $15S_1$, $17S_1$

or 6S₁) or the 'roof node' is reached. In the first case, if the node under consideration is 'significant' (e.g. nodes $2L_1$, $7L_1$, $15S_1$, $6S_1$ or $19S_1$), the tree is pruned under it, unless the node is a 'loaf node' (e.g. nodes $12S_1$ or $21S_1$) where there is no need to prune. On the contrary, if the node is not 'significant' (e.g. nodes $14S_1$ or $7S_1$) the pruning is performed above it. In the second case, when the 'roof node' is reached, the tree is pruned under it.

Once the process is repeated for all the 'loaf nodes' and the whole tree is pruned, a set of selected regions is obtained ($2L_1$, $9L_1$, $10L_1$ and $7L_1$ for the $LSLE_1$ tree and $12L_1$, $21L_1$, $22L_1$, $15L_1$, $17L_1$,$6L_1$, $18L_1$, $19L_1$, $4S_1$ for the $SSLE_1$). At this point, the pruning process for the $SLSE_s$ is considered finished while for the $LSLE_s$ a third condition is imposed. In such a case, the sum of the selected regions has be larger than 75% of the 'roof node' area. When the condition is met, the pruning process ends. If not, which is the case for $LSLE_1$ (the sum of initial set of selected regions ($2L_1$, $9L_1$, $10L_1$, $7L_1$) is lower than 75% of the roof node) the regions at the lowest level of the tree are discarded ($9L_1$ and $10L_1$) and their parent ($10L_1$), is categorized as a new 'loaf node' for which the whole pruning process is restarted. As a consequence, a new set of selected regions are obtained ($2L_1$, $6L_1$ and $7L_1$) and the third condition is reexamined. This is repeatedly done until the condition is met (In the example, the sum of the new set of selected regions areas ($2L_1$, $6L_1$ and $7L_1$) fulfill the condition) or the roof node is reached. At that moment, the pruning process is finished and the last set of obtained regions is eventually selected. Then, as represented in Fig. A.1, a final dilation treatment, consisting of expanding these region shapes until the borders of its roof region, is applied.

After the whole pruning process is finished for both $LSLE_s$ and $SSLE_s$, a final composite of selected entities, large convective storms ($LCS_s$) and small convective cores ($SCC_s$), are respectively obtained from for both structural level entities.

The heuristics involved in the MTS technique were established after performing a sensitivity analysis for different values within the following intervals: For $SCC_s$, values within 1–8 km² and 5–20% (Condition 1), and 40–75% (Condition 2) were tested. Values within an interval of 15–30% (Condition 1), and 30–70% (Condition 2) and 50–90% (Condition 3) were analyzed for the $LCS_s$ case. The final employed values were calibrated by visually checking how satisfactory the identification processes results were in the calibration events (Events A to F, Table 1).
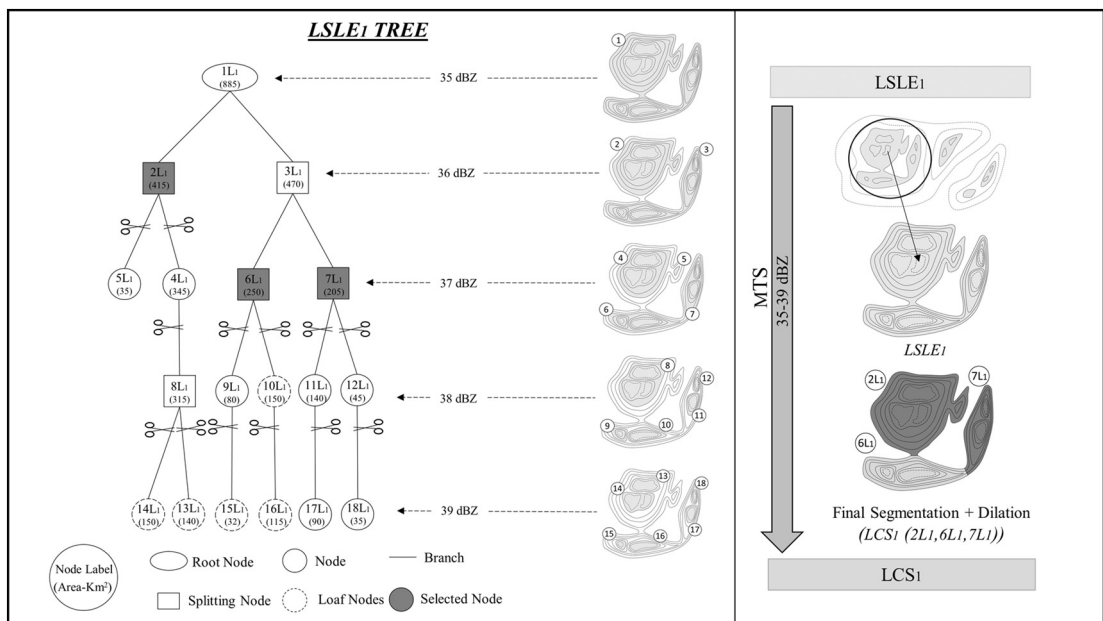


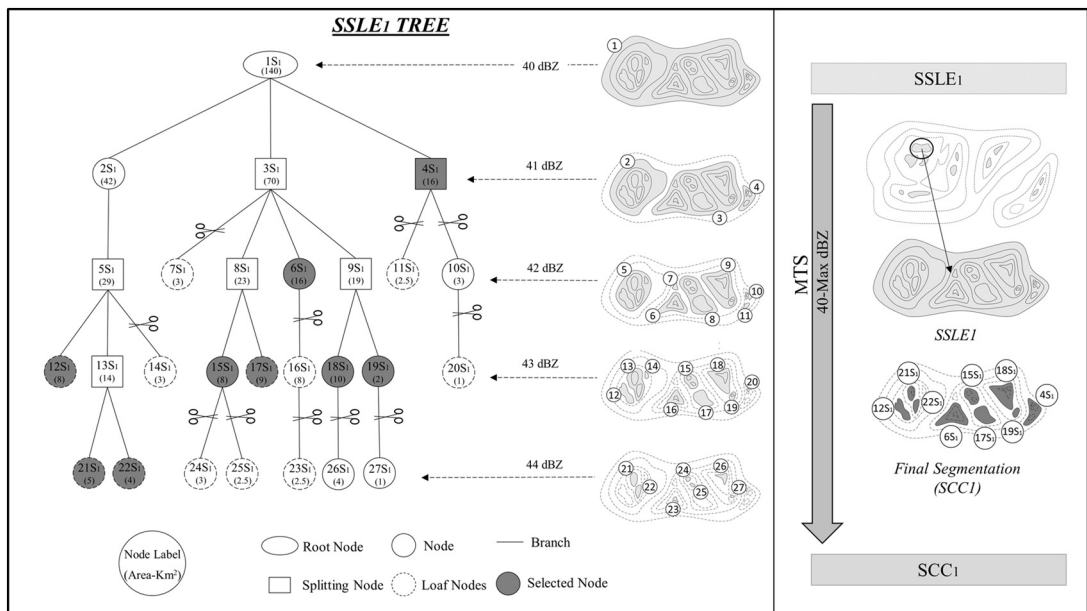**Fig. A.1.** Schematic example of the MTS technique applied to a LSLE.

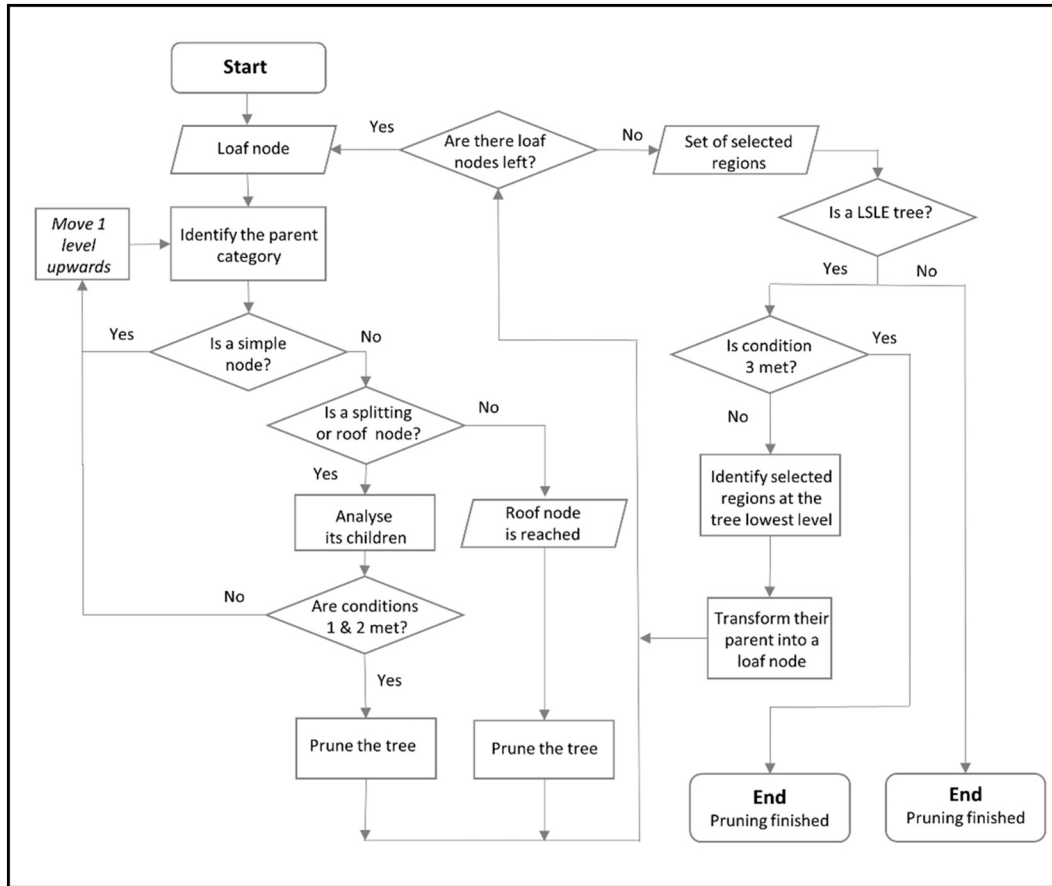**Fig. A.2.** Schematic example of the MTS technique applied to a SSLE.



**Fig. A.3.** The flowchart showing the general steps in the pruning process.

# References

Berenguer, M., Corral, C., Sánchez-Diezma, R., Sempere-Torres, D., 2005. Hydrological validation of a radar-based nowcasting technique. J. Hydrometeorol. 6, 532–549. http://dx.doi.org/10.1175/JHM433.1.

Bowler, N.E., Pierce, C.E., Seed, A., 2004. Development of a precipitation nowcasting algorithm based upon optical flow techniques. J. Hydrol. 288, 74–91. http://dx.doi.org/10.1016/j.jhydrol.2003.11.011.

Bowler, N.E., Pierce, C.E., Seed, A., 2006. STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. Q. J. R. Meteorol. Soc. 132, 2127–2155. http://dx.doi.org/10.1256/qj.04.100.

Brox, T., Bruhn, A., Papenberg, N., Weickert, J., 2004. High accuracy optical flow estimation based on a theory for warping. In: Pajdla, T., Matas, J. (Eds.), Computer Vision - ECCV 2004 SE - 3, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 25–36. http://dx.doi.org/10.1007/978-3-540-24673-2_3.

Caine, S., Lane, T.P., May, P.T., Jakob, C., Siems, S.T., Manton, M.J., Pinto, J., 2013. Statistical assessment of tropical convection-permitting model simulations using a cell-tracking algorithm. Mon. Weather Rev. 141, 557–581. http://dx.doi.org/10.1175/MWR-D-11-00274.1.

Carbunaru, D., Stefan, S., Sasu, M., Stefanescu, V., 2013. Analysis of convective thunderstorm split cells in south-eastern Romania. Int. J. Atmos. Sci. 2013, 1–19. http://dx.doi.org/10.1155/2013/162541.

Cheung, P., Yeung, H.Y., 2012. Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong. In: The 3rd WMO International Symposium on Nowcasting and Very Short-range Forecasting(WSN12). Rio de Janeiro, Brasil, Available online at: http://www.labhidro.iag.usp.br/wsn12/papers/tra6.pdf.

Ćurić, M., Janc, D., 2011a. Comparison of modeled and observed accumulated convective precipitation in mountainous and flat land areas. J. Hydrometeorol. 12, 245–261. http://dx.doi.org/10.1175/2010JHM1259.1.

Ćurić, M., Janc, D., 2011b. Analysis of predicted and observed accumulated convective precipitation in the area with frequent split storms. Hydrol. Earth Syst. Sci. 15, 3651–3658. http://dx.doi.org/10.5194/hess-15-3651-2011.

Ćurić, M., Janc, D., 2012. Differential heating influence on hailstorm vortex pair formation. Q. J. R. Meteorol. Soc. 138, 72–80. http://dx.doi.org/10.1002/qj.918.

Ćurić, M., Janc, D., Vučković, V., 2007. Numerical simulation of Cb cloud vorticity. Atmos. Res. 83, 427–434. http://dx.doi.org/10.1016/j.atmosres.2005.10.024.

Ćurić, M., Janc, D., Vučković, V., 2009. The influence of merging and individual storm splitting on mesoscale convective system formation. Atmos. Res. 93, 21–29. http://dx.doi.org/10.1016/j.atmosres.2008.10.018.

Delobbe, L., Holleman, I., 2006. Uncertainties in radar echo top heights used for hail detection. Meteorol. Appl. 13, 361–374. http://dx.doi.org/10.1017/S1350482706002374.

Dixon, M., Seed, A.W., 2014. Developments in echo tracking - enhancing TITAN. In: The Eighth European Conference on Radar in Meteorology and Hydrology. Garmisch-Partenkirchen, Germany, Available online at: www.pa.op.dlr.de/erad2014/programme/ExtendedAbstracts/155_Dixon.pdf.

Dixon, M., Wiener, G., 1993. TITAN: thunderstorm identification, tracking, analysis, and nowcasting—a radar-based methodology. J. Atmos. Ocean. Technol. 10, 785–797. http://dx.doi.org/10.1175/1520-0426(1993)010.

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. J. Hydrol. 299, 186–202. http://dx.doi.org/10.1016/j.jhydrol.2004.08.004.

Emmanuel, I., Andrieu, H., Leblois, E., Flahaut, B., 2012. Temporal and spatial variability of rainfall at the urban hydrological scale. J. Hydrol. 430–431, 162–172. http://dx.doi.org/10.1016/j.jhydrol.2012.02.013.

Fiolleau, T., Roca, R., 2013. An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite. IEEE Trans. Geosci. Remote Sens. 51, 4302–4315. http://dx.doi.org/10.1109/TGRS.2012.2227762.

Foresti, L., Reyniers, M., Seed, A., Delobbe, L., 2016. Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium. Hydrol. Earth Syst. Sci. 20, 505–527. http://dx.doi.org/10.5194/hess-20-505-2016.

García-Ortega, E., López, L., Sánchez, J.L., 2009. Diagnosis and sensitivity study of two severe storm events in the Southeastern Andes. Atmos. Res. 93, 161–178. http://dx.doi.org/10.1016/j.atmosres.2008.10.030.

Gascón, E., Merino, A., Sánchez, J.L., Fernández-González, S., García-Ortega, E., López, L., Hermida, L., 2015. Spatial distribution of thermodynamic conditions of severe storms in southwestern Europe. Atmos. Res. 164–165, 194–209. http://dx.doi.org/10.1016/j.atmosres.2015.05.012.

Germann, U., Zawadzki, I., 2002. Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. Mon. Wea. Rev. 130, 2859–2873. http://dx.doi.org/10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2.

Germann, U., Zawadzki, I., Turner, B., 2006. Predictability of precipitation from

continental radar images. Part IV: limits to prediction. J. Atmos. Sci. 63, 2092–2108. http://dx.doi.org/10.1175/JAS3735.1.

Goudenhoofdt, E., Delobbe, L., 2013. Statistical characteristics of convective storms in Belgium derived from volumetric weather radar observations. J. Appl. Meteorol. Climatol. 52, 918–934. http://dx.doi.org/10.1175/JAMC-D-12-079.1.

Han, L., Fu, S., Zhao, L., Zheng, Y., Wang, H., Lin, Y., 2009. 3D convective storm identification, tracking, and forecasting - an enhanced TITAN algorithm. J. Atmos. Ocean. Technol. 26, 719–732. http://dx.doi.org/10.1175/2008JTECHA1084.1.

Handwerker, J., 2002. Cell tracking with TRACE3D - a new algorithm. Atmos. Res. 61, 15–34. http://dx.doi.org/10.1016/S0169-8095(01)00100-4.

Hering, A.M., Morel, C., Galli, G., Ambrosetti, P., Boscacci, M., 2004. Nowcasting thunderstorms in the Alpine Region using a radar based adaptive thresholding scheme. In: Proc. ERAD Conf. 2004, Visby, Sweden, pp. 206–211.

Johnson, J.T., MacKeen, P.L., Witt, A., Mitchell, E.D.W., Stumpf, G.J., Eilts, M.D., Thomas, K.W., 1998. The storm cell identification and tracking algorithm: an enhanced WSR-88D algorithm. Weather Forecast. 13, 263–276. http://dx.doi.org/10.1175/1520-0434(1998)013.

Jung, S.-H., Lee, G., 2015. Radar-based cell tracking with fuzzy logic approach. Meteorol. Appl. 22, 716–730. http://dx.doi.org/10.1002/met.1509.

Karacostas, T., Spiridonov, V., Stolaki, S., Pytharoulis, I., Tegoulias, I., 2013. A three-dimensional simulation of the 10th August 2008 storm occurred over Greece: AgI seeding of cell merger by using a cloud resolving model. Atmos. Res. 169 (Part B), 547–555. http://dx.doi.org/10.1155/2010/234731.

Kober, K., Tafferner, A., 2009. Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. Meteorol. Z. 18, 75–84. http://dx.doi.org/10.1127/0941-2948/2009/359.

Kyznarová, H., Novák, P., 2009. CELLTRACK - convective cell tracking algorithm and its use for deriving life cycle characteristics. Atmos. Res. 93, 317–327. http://dx.doi.org/10.1016/j.atmosres.2008.09.019.

Lakshmanan, V., Rabin, R., DeBrunner, V., 2003. Multiscale storm identification and forecast. Atmos. Res. 67–68, 367–380. http://dx.doi.org/10.1016/S0169-8095(03)00068-1.

Laroche, S., Zawadzki, I., 1994. A variational analysis method for retrieval of three-dimensional wind field from single-Doppler radar data. J. Atmos. Sci. 51 (18), 2664–2682. http://dx.doi.org/10.1175/1520-0469(1994)051.

Laroche, S., Zawadzki, I., 1995. Retrievals of horizontal winds from single-Doppler clear-air data by methods of cross correlation and variational analysis. J. Atmos. Ocean. Technol. 12, 721–738. http://dx.doi.org/10.1175/1520-0426(1995)012.

Liang, Q.Q., Feng, Y.R., Deng, W.J., Hu, S., Huang, Y.Y., Zeng, Q., Chen, Z.T., 2010. A composite approach of radar echo extrapolation based on TREC vectors in combination with model-predicted winds. Adv. Atmos. Sci. 27, 1119–1130. http://dx.doi.org/10.1007/s00376-009-9093-4.

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. http://dx.doi.org/10.1016/j.atmosres.2011.05.004.

Liu, K., Du, Q., Yang, H., Ma, B., 2010. Optical flow and principal component analysis-based motion detection in outdoor videos. EURASIP J. Adv. Signal. Process. 2010http://dx.doi.org/10.1155/2010/680623. (680–623).

Liu, J., Ma, C., Liu, C., Qin, D., Gu, X., 2014. An extended maxima transform-based region growing algorithm for convective cell detection on satellite images. Remote Sens. Lett. 5, 971–980. http://dx.doi.org/10.1080/2150704X.2014.980917.

McRobie, F.H., Wang, L.P., Onof, C., Kenney, S., 2013. A spatial-temporal rainfall generator for urban drainage design. Water Sci. Technol. 68, 240–249. http://dx.doi.org/10.2166/wst.2013.241.

Mueller, C., Saxen, T., Roberts, R., Wilson, J., Betancourt, T., Dettling, S., Oien, N., Yee, J., 2003. NCAR auto-nowcast system. Weather Forecast. 18, 545–561. http://dx.doi.org/10.1175/1520-0434(2003)018<0545:NAS>2.0.CO;2.

Nielsen, J.E., Thorndahl, S., Rasmussen, M.R., 2014. A numerical method to generate high temporal resolution precipitation time series by combining weather radar measurements with a nowcast model. Atmos. Res. 138, 1–12. http://dx.doi.org/10.1016/j.atmosres.2013.10.015.

Niemi, T.J., Kokkonen, T., Seed, A.W., 2014. A simple and effective method for quantifying spatial anisotropy of time series of precipitation fields. Water Resour. Res. 50, 5906–5925. http://dx.doi.org/10.1002/2013WR015190.

Novak, P., 2007. The Czech Hydrometeorological Institute's severe storm nowcasting system. Atmos. Res. 83, 450–457. http://dx.doi.org/10.1016/j.atmosres.2005.09.014.

Novo, S., Martínez, D., Puentes, O., 2014. Tracking, analysis, and nowcasting of cuban convective cells as seen by radar. Meteorol. Appl. 21, 585–595. http://dx.doi.org/10.1002/met.1380.

Peak, J.E., Tag, P.M., 1994. Segmentation of satellite imagery using hierarchical thresholding and neural networks. J. Appl. Meteor. 33, 605–616. http://dx.doi.org/10.1175/1520-0450(1994)033<0605:SOSIUH>2.0.CO;2.

Pierce, C.E., Ebert, E., Seed, A.W., Sleigh, M., Collier, C.G., Fox, N.I., Donaldson, N., Wilson, J.W., Roberts, R., Mueller, C.K., 2004. The nowcasting of precipitation during Sydney 2000: an appraisal of the QPF algorithms. Weather Forecast. 19, 7–21. http://dx.doi.org/10.1175/1520-0434(2004)019.

Pierce, C., Seed, A., Ballard, Sue, Simonin, D., Li, Zhihong, 2012. Nowcasting. In: Bech, J. (Ed.), Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications. InTech, pp. 97–142. http://dx.doi.org/10.5772/39054.

Potts, R., 1993. The application of TITAN for thunderstorm nowcast operations. In: 17th BMRC Modelling Workshop, Melbourne, Australia, pp. 147–151.

Reyniers, M., 2008. Quantitative Precipitation Forecasts Based on Radar Observations: Principles, Algorithms and Operational Systems. vol. 62 R. Meteorol. Inst. Belgium Available online at: https://www.meteo.be/meteo/download/fr/3040165/pdf/rmi_scpub-1261.pdf.

Rinehart, R.E., Garvey, E.T., 1978. Three-dimensional storm motion detection by conventional weather radar. Nature 273, 287–289. http://dx.doi.org/10.1038/273287a0.

Sadek, R., Ballester, C., Garrido, L., Meinhardt, E., Caselles, V., 2012. Frame interpolation with occlusion detection using a time coherent segmentation. In: International Conference on Computer Vision Theory and Applications (VISAPP 2012). Roma, Italy, pp. 367–372.

Sebastianelli, S., Russo, F., Napolitano, F., Baldini, L., 2013. On precipitation measurements collected by a weather radar and a rain gauge network. Nat. Hazards Earth Syst. Sci. 13, 605–623. http://dx.doi.org/10.5194/nhess-13-605-2013.

Seo, D.-J., Habib, E., Andrieu, H., Morin, E., 2015. Hydrologic applications of weather radar. J. Hydrol. 531, 231–233. http://dx.doi.org/10.1016/j.jhydrol.2015.11.010.

Shah, S., Notarpietro, R., Bertoldo, S., Branca, M., Lucianaz, C., Rorato, O., Allegretti, M., 2013. Automatic storm(s) identification in high resolution, short range, X-band radar images. In: Proceedings of the 2013 International Conference on Electromagnetics in Advanced Applications, ICEAA 2013. Torino, Italy, pp. 945–948. http://dx.doi.org/10.1109/ICEAA.2013.6632379.

Shimizu, S., Uyeda, H., 2012. Algorithm for the identification and tracking of convective cells based on constant and adaptive threshold methods using a new cell-merging and -splitting scheme. J. Meteorol. Soc. Japan 90, 869–889. http://dx.doi.org/10.2151/jmsj.2012-602.

Spiridonov, V., Dimitrovski, Z., Ćurić, M., 2010. A three-dimensional simulation of supercell convective storm. Adv. Meteorol. 2010. http://dx.doi.org/10.1155/2010/234731.

Thorndahl, S., Smith, J.A., Baeck, M.L., Krajewski, W.F., 2014. Analyses of the temporal and spatial structures of heavy rainfall from a catalog of high-resolution radar rainfall fields. Atmos. Res. 144, 111–125. http://dx.doi.org/10.1016/j.atmosres.2014.03.013.

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J.E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M.R., Molnar, P., 2016. Weather radar rainfall data in urban hydrology. Hydrol. Earth Syst. Sci. Discuss. 21, 1359–1380. http://dx.doi.org/10.5194/hess-2016-517.

Tuttle, J.D., Foote, G.B., 1990. Determination of the boundary layer airflow from a single Doppler radar. J. Atmos. Ocean. Technol. 7, 218–232. http://dx.doi.org/10.1175/1520-0426(1990)007<0218:DOTBLA>2.0.CO;2.

Verbeke, N., Vincent, N., 2007. A PCA-based technique to detect moving objects. In: Ersbøll, B.K., Pedersen, K.S. (Eds.), Image Analysis. SCIA 2007. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 641–650. http://dx.doi.org/10.1007/978-3-540-73040-8_65.

Vulpiani, G., Baldini, L., Roberto, N., 2015. Characterization of Mediterranean hail-bearing storms using an operational polarimetric X-band radar. Atmos. Meas. Tech. 8, 4681–4698. http://dx.doi.org/10.5194/amt-8-4681-2015.

Wang, L.-P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R.D., Pessemier, M., Kroll, S., Willems, P., Onof, C., 2015. Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment. J. Hydrol. 531, 408–426. http://dx.doi.org/10.1016/j.jhydrol.2015.05.049.

Wernli, H., Paulat, M., Hagen, M., Frei, C., 2008. SAL - a novel quality measure for the verification of quantitative precipitation forecasts. Mon. Weather Rev. 136, 4470–4487. http://dx.doi.org/10.1175/2008MWR2415.1.

Wilhelmson, R.B., Klemp, J.B., 1978. A numerical study of storm splitting that leads to long-lived storms. J. Atmos. Sci. 35, 1974–1986. http://dx.doi.org/10.1175/1520-0469(1978)035.

Willems, P., 2001. A spatial rainfall generator for small spatial scales. J. Hydrol. 252, 126–144. http://dx.doi.org/10.1016/S0022-1694(01)00446-2.

Wilson, J.W., Ebert, E.E., Saxen, T.R., Roberts, R.D., Mueller, C.K., Sleigh, M., Pierce, C.E., Seed, A., 2004. Sydney 2000 forecast demonstration project: convective storm nowcasting. Weather Forecast. 19, 131–150. http://dx.doi.org/10.1175/1520-0434(2004)019.

Wu, X., Li, X., 2008. A review of cloud-resolving model studies of convective processes. Adv. Atmos. Sci. 25, 202–212. http://dx.doi.org/10.1007/s00376-008-0202-6.

Yang, L., Tian, F., Smith, J.A., Hu, H., 2014. Urban signatures in the spatial clustering of summer heavy rainfall events over the Beijing metropolitan region. J. Geophys. Res. Atmos. 119, 1203–1217. http://dx.doi.org/10.1002/2013JD020762.

Zahraei, A., Hsu, K. lin, Sorooshian, S., Gourley, J.J., Lakshmanan, V., Hong, Y., Bellerby, T., 2012. Quantitative precipitation nowcasting: a lagrangian pixel-based approach. Atmos. Res. 118, 418–434. http://dx.doi.org/10.1016/j.atmosres.2012.07.001.

# Exploring the potential of multivariate depth-damage and rainfall-damage models

L. Van Ootegem[1,2], K. Van Herck[1], T. Creten[1], E. Verhofstadt[2], L. Foresti[3,4], E. Goudenhoofdt[3], M. Reyniers[3], L. Delobbe[3], D. Murla Tuyls[5] and P. Willems[5]

1 HIVA – Research Institute for Work and Society, KU Leuven, Leuven, Belgium
2 Ghent University, Gent, Belgium
3 Royal Meteorological Institute of Belgium (RMI), Brussels, Belgium
4 Radar, Satellite and Nowcasting Division, MeteoSwiss, Locarno-Monti, Switzerland
5 Hydraulics Section, KU Leuven, Leuven, Belgium

**Correspondence**
Luc Van Ootegem, HIVA – Research Institute for Work and Society, KU Leuven, Parkstraat 47, 3000 Leuven, Belgium
Email: Luc.VanOotegem@UGent.be

## Abstract

In Europe, floods are among the natural catastrophes that cause the largest economic damage. This article explores the potential of two distinct types of multivariate flood damage models: 'depth-damage' models and 'rainfall-damage' models. We use survey data of 346 Flemish households that were victim of pluvial floods complemented with rainfall data from both rain gauges and weather radars. In the econometrical analysis, a Tobit estimation technique is used to deal with the issue of zero damage observations. The results show that in the 'depth-damage' models flood depth has a significant impact on the damage. In the 'rainfall-damage' models there is a significant impact of rainfall accumulation on the damage when using the gauge rainfall data as predictor, but not when using the radar rainfall data. Finally, non-hazard indicators are found to be important for explaining pluvial flood damage in both 'depth-damage' and 'rainfall-damage' models.

## Introduction

In Europe floods are among the natural catastrophes that cause the largest economic damage (European Environment Agency, 2012). The prediction of damage is essential for risk assessments and cost-benefit analyses to decide about e.g. government investments in specific infrastructure or for setting up flood warning systems. Most of the existing studies use data from floods caused by rivers overflowing (called fluvial floods) (Messner and Meyer, 2006; Jonkman *et al.*, 2008). However, their generalisability to other floods is unclear (Kellens *et al.*, 2013). In this article, the focus is on damage caused by a type of flood for which damage assessment has been rarely conducted, i.e. floods in urban areas that are caused by extreme rainfall events during which the water cannot be sufficiently processed by existing urban drainage systems (called pluvial floods). Pluvial floods come with less damage, but occur frequently and the cumulative damage over the years can be just as high as with fluvial events (ten Veldhuis, 2011). Moreover, projections on the impact of climate change indicate that in the future many regions of the world will face more extreme weather events

in summer, such as thunderstorms and heavy rainfall, which may increase even further the occurrence of pluvial floods (National Climate Commission, 2010; IPCC, 2012; Willems *et al.*, 2012; Willems, 2013).

Researchers often combine simulated flood depths with existing depth-damage curves to obtain damage predictions (Ernst *et al.*, 2008; Pistrika and Jonkman, 2009; De Moel and Aerts, 2011). An increasing number of authors pointed out the shortcomings of the existing models used for flood damage estimations (Freni *et al.*, 2010; Hurford *et al.*, 2012; Spekkers *et al.*, 2013). Often these depth-damage curves do not include any control variables and are not specific regarding the type of damage (damage to buildings versus damage to the content) nor the type of flood (fluvial versus pluvial). Therefore, other studies have estimated the relationship between the depth of the flood and the monetary damage, including a number of control variables (Freni *et al.*, 2010; Van Ootegem *et al.*, 2015). This type of model requires ex-post survey or insurance information on the depth of the flood or the availability of simulated depths, e.g. by hydraulic models or by probabilistic or stochastic simulation approaches (Van Dyck and Willems, 2013).

An alternative approach is to estimate the relationship between the rainfall at the time of the flood and the monetary damage (so called 'rainfall-damage' models instead of 'depth-damage' models). This approach does not require ex-post information or the simulation of flood depths using time-demanding hydraulic modelling. Instead, it derives the direct relationship between the rainfall and the monetary damage. In addition, 'rainfall-damage' models have the important benefit that rainfall data are completely exogenous in the regression. This is in contrast with the depth data which are potentially endogenous meaning that there can be correlation between the depth data and the error term in the regression. Endogeneity can arise as a result of an omitted variable which affects both the flood depth and the damage. The model coefficients then visualise not only the relation between depth and damage (what they should be representing), but also the (unknown) effect of the omitted variable. Examples of such omitted variables could be certain characteristics of the house (e.g. wall around the property) or socio-economic characteristics (e.g. composition of the household). A disadvantage of the 'rainfall-damage' relationship is that it is less direct; the same rainfall intensity on sewer systems with different hydraulic characteristics may lead to different impacts. The sewer system characteristics are not accounted for (paved areas, pumps, and other hydraulic regulation structures, system configuration, etc.), whereas they may cause differences in sewer flood impacts for the same rainfall intensity.

Few authors have estimated 'rainfall-damage' models. For example, Zhou *et al.* (2012) analysed the impact of the rainfall on 1000 insurance claims related to pluvial floods in Aarhus (Denmark). They concluded that the rainfall depth does not significantly affect the cost per claim. However, the rainfall rate affects the total cost per day. Ririassa and Hoen (2010) analysed insurance claims in the Netherlands, but did not find a significant relationship between the total level of monetary damage in a province and the hourly rainfall data. Spekkers *et al.* (2013) found that rainfall intensity has a significant impact on the monetary damage. Blanc *et al.* (2012) estimate pluvial flood damage using rainfall simulations to compute flood depth. However, overall the explained variance in case of the 'rainfall-damage' models is rather low. This shows the need to further study these types of models and include additional explanatory variables such as the socio-economic characteristics of the households that have been flooded. As illustrated in Figure 1, the relationship described by 'depth-damage' models is in fact part of the 'rainfall-damage' relationship.

In this article, we study both 'depth-damage' and 'rainfall-damage' models. For the 'depth-damage' models, we used survey data of 346 Flemish households that were victims of pluvial floods. We collected data on a wide range of variables, including the reported depth of the flood (in the
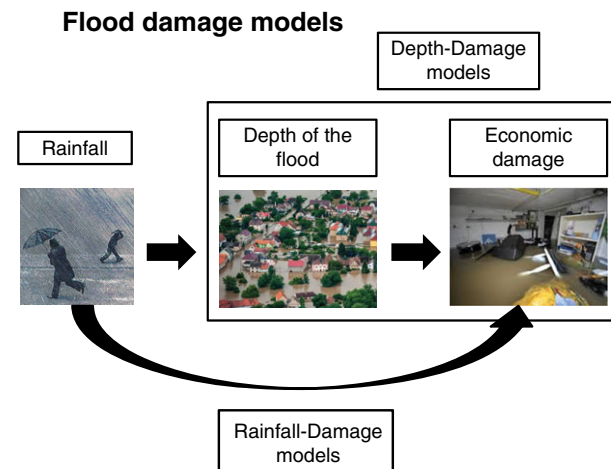
**Flood damage models**



**Figure 1** 'Depth-damage' versus 'rainfall-damage' models.

basement and at the ground floor), the monetary damage to different parts of the house (building, content, and total damage), the prevention measures taken by households, socio-economic characteristics such as the per capita income of the individuals that have been flooded etc. Based on these survey data we estimated 'depth-damage' models with self-reported depth as the main independent variable and self-reported damage as the dependent variable. In addition, we included a number of control variables obtained from the survey. For the 'rainfall-damage' models, we combined damage data and control variables obtained from the survey with rainfall data from two sources, namely rain gauges and weather radars. The contribution of our research is threefold.

First, in addition to estimating a 'depth-damage' model, we also estimate a 'rainfall-damage' model. Estimating the direct relationship between the rainfall and monetary damage has – at least – two major advantages. The first is that instead of using ex-post survey data or simulating the expected depth of a flood based on hydraulic models, we estimate the direct relationship between rainfall accumulation and damage. This is an important time-saving step, which is particularly important for the development of early-warning systems for pluvial floods which in general have short lead times. Moreover, hydraulic models are location-specific and do not exist for every urban area. An additional advantage is that we are able to exclude endogeneity problems by using exogenous data on the rainfall intensity, (cfr. supra).

Second, by using a Tobit model estimation, we deal with the issue of reported zero damage (Tobin, 1958). As a result, we are able to include 'no damage cases' of people who were flooded but did not suffer any damage. Those people would not appear in insurance or disaster fund records. Our questionnaire reveals that a certain fraction of flood victims suffer no damage to buildings and/or content.

Third, we extend the traditional bivariate depth/rainfall-damage relations by constructing more detailed multivariate flood damage models including a number of non-hazard control variables. This shift to more complex multivariate models has taken place in the literature in recent years (Apel *et al.*, 2009; Elmer *et al.*, 2010; Merz *et al.*, 2013). We include characteristics of the buildings (type of building, size of the dwelling), behavioural predictors related to the behaviour of victims before and during the flood (recurrence, risk awareness, emergency measures) and income.

The remaining of the article is organised as follows. The section on *Data and methodology* describes the data (collection) and econometrical methodology. In the section on *Discussion of the results*, we discuss the results. Finally, we provide conclusions and policy implications in the section on *Conclusions*.

## Data and methodology

In this article, we use a data set of 346 Flemish households that experienced floods in the past two decades. This data set is based on two sources. First, we collected survey data on the damage and the depth of the floods as well as on a number of socio-economic characteristics of the households that have been flooded. Second, we supplemented these survey data with detailed rainfall data measured by rain gauges and estimated by weather radars.

### Survey data

In 2013 we conducted a mail survey among private households that were affected by one or more pluvial floods. The survey was sent to 3963 addresses in Flanders (the northern region of Belgium). The vast majority of those addresses come from a database of the national disaster fund. Also 260 addresses were included from records of fire departments in villages and cities that were flooded recently. A total of 973 people sent back the questionnaire, which corresponds to a response rate of 24.5%. First, 353 questionnaires could not be used because they never suffered damage from floods (people moved to the address after the event took place) or because they report damage to their business. Second, in the remaining 620 cases, quite a large amount of people were not capable of estimating the

damage they suffered. Another, less important, reason for fall-out was missing data on the flood year (making it impossible to calculate damage figures corrected for inflation). Respondents were asked to self-report on the damage and the flood depth. With some of the respondents, there was a telephone call to check the validity and reliability of the answers. We have performed a robustness check in which we restricted the sample to the past 10 years since we expect that respondents are more able to remember recent events. That check did not affect our results significantly. In Van Ootegem *et al.* (2015), the authors provide a more detailed description of the sample and the survey methodology.

The main variable of interest is the damage that the households experienced because of the flood. Respondents were asked to report about their worst flood damage experience, which implies that the damage figures may be related to different years and hence have to be corrected for inflation. For this correction, we used the Belgian Consumer Price Index. We distinguish between three types of damage. First, we asked the respondents to report the financial damage to several parts of the building. The reported amounts for the different parts of the building are added to construct a total building damage variable. For damage to the content, we asked victims to report immediately on the total damage to their content. Finally, we aggregate the damage to the building and to the content to get a total damage figure. Table 1 shows the summary statistics on the damage levels: 283 respondents report a damage figure for their building, while for content 321 respondents report a damage figure. On average, the damage to the building is higher than the damage to the content. Zero damage is frequently the case, especially for the content damage. In addition, we also asked information on the total damage that has been recovered by the respondents from the disaster fund or insurance company. The total damage recovered from the disaster fund or insurance company is substantially lower for both buildings and content.

### Rainfall data

Rainfall data are obtained from two sources. First, we used rainfall data from the rain gauge network of the Flemish Environment Agency (VMM). For each pluvial flood event

**Table 1** Reported damage figures and zero damages (euro corrected for inflation)

|  | Observations | Zero damage observations | Mean | Max damage |
|---|---|---|---|---|
| *Total damage to* | 346 | 57 | 8001.7 | 91 657.6 |
|   Building | 283 | 24 | 6262.8 | 77 499.3 |
|   Content | 321 | 132 | 3103.5 | 65 576.3 |
| Total damage recovered from the disaster fund or insurance | 175 | 0 | 2070.5 | 19 307.8 |

in the survey, we derived the daily rainfall accumulation of the closest rain gauge. Although sub-daily rainfall accumulations are expected to be more indicative for pluvial sewer system floods, the network of rain gauges with sub-daily rainfall registrations is much coarser than the network with daily registrations. Moreover, the days with the highest sub-daily rainfall maxima most often also show the highest daily accumulations. Based on the geographical location of the respondents' residence (geographical coordinates) and information on the timing of the flood (we have asked for the month and the year of the flood), we have computed the maximum daily rainfall accumulation for that month.

Second, we derived rainfall data from the weather radar as an alternative to rain gauges. This has two major advantages. First, the distance between the location of the pluvial flood and the closest rain gauge may be too large to find a good correlation, in particular when the thunderstorms are highly localised in space. Second, the temporal resolution of daily rainfall accumulations may be too coarse for our application, in particular because the typical response times of sewer systems are often below 1 h. Weather radars are able to provide detailed information about the spatial and temporal distribution of rainfall at resolutions of 1 km and 5 min respectively. Therefore, there is potential to find a better agreement between the rainfall accumulation derived from the radar and the location of the pluvial flood. In this study, we used rainfall rates estimated from the C-band weather radar of the Royal Meteorological Institute (RMI) of Belgium located at Wideumont (Ardennes). The radar data processing chain at RMI includes a number of steps to

convert the measured reflectivity $Z$ into rainfall rates $R$, in particular to remove non-meteorological echoes (flying objects, ground clutter, radio interferences) and to extrapolate the reflectivity aloft to the ground level. A detailed description of the radar-based quantitative precipitation estimation and its verification at RMI is presented in Goudenhoofdt and Delobbe (2016). Such data processing steps are important to increase the accuracy of radar-based rainfall estimation, which is affected by various sources of errors and uncertainty (Villarini and Krajewski, 2010).

Based on the geographical location of the respondents' residence (geographical coordinates) and information on the timing of the flood (month and year of the flood), we have computed the maximum hourly radar rainfall accumulation on each month and registered its timing of occurrence (see Figure 2). In order to account for the size of the sewer system and the mismatch between the exact location of the flood and the thunderstorm, the maximum hourly rainfall accumulations were computed in circular radii of 1, 5, 10, and 15 km from the respondents' residence. In the multivariate regression (infra) we have only included the maximum rainfall accumulation in a radius of 1 km. The results of the multivariate regression remain stable also if we use radii of 5, 10, or 15 km (results available upon request from the authors).

## Econometrical approach

We report on the estimation of two types of models: 'depth-damage' models and 'rainfall-damage' models which
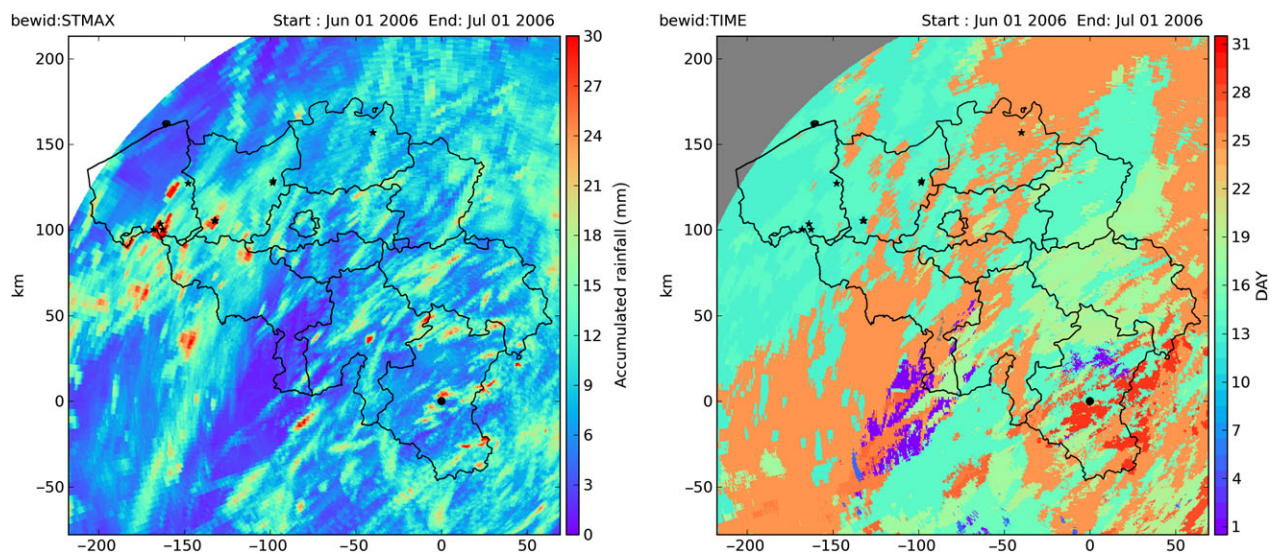


**Figure 2** Spatial distribution of the maximum hourly rainfall derived from the Wideumont weather radar (located at coordinates (0,0) km) for the month of June 2006. The location of sewer system floods is marked with stars. The borders of the Belgian districts are displayed in black. (a) Maximum hourly rainfall accumulation. (b) Corresponding day of occurrence (the hour of occurrence is also available but not shown).

analyse the impact of the depth of the flood, respectively the rainfall, at the time of the flood on the damage, controlling for a number of non-hazard indicators. For each of these models, we distinguish between six distinct model specifications. First, for both types of models we distinguish between three dependent variables, namely the damage to the content, the damage to the building and the total damage. Second, we include for the 'depth-damage' models two measures of flood depth, namely flood depth in the basement and at the ground floor. For the 'rainfall-damage' models, we include the rain gauge data and radar data as the main independent rainfall variable of interest. The damage data are skewed to the left as the majority of the households report relative low levels of damage, while few of them report high levels of damage. In order to obtain a less skewed distribution, we did a log-transformation of the damage data, which allows us to estimate a linear relation between damage and depth.

In a first step, we estimate a simple ordinary least squares (OLS) model. The model specification and results of the OLS are discussed in the appendix. However, given that a substantial number of respondents reported zero damage, especially for content damage, OLS estimates may become biased and inefficient depending on the number of zeros in relation to the number of observations in the data set. One solution is simply to delete the zero values, i.e. truncating the sample. However, this would mean ignoring relevant information and therefore would also lead to inconsistent estimations of the parameters (Wooldridge, 2010).

Therefore, in a second step, we used a Tobit model, developed by James Tobin to deal with the problem of the many zero values. The model was originally used to explain purchases of luxury goods. Because of income restraints, many survey respondents would report zero purchases of a luxury good, leading to problems when using OLS to identify the factors that explain the purchase of those goods (Tobin, 1958). The methodology is widely applicable for any situation where observations are 'censored' above or below a certain border value $c$. the Tobit model is therefore also known as a specific type of 'censored regression model' (Wooldridge, 2010) and is widely used for analysing diverse micro-economic topics such as e.g. the hours worked by different socio-economic groups (see McDonald and Moffitt, 1980) or explaining charity donations.

The Tobit model is based on the assumption that there is a latent (i.e. unobservable) outcome variable $\text{Damage}_i^*$ for a household $i$. This variable linearly depends on the main variable of interest ($\text{MainVar}_i$) and a vector of $n$ control variables $X_i$ via a parameter (vector) which determines the relationship between the explanatory variables and the latent variable $\text{Damage}_i^*$. In addition, the model includes a normally distributed error term $\varepsilon_i$ to capture random influences on this relationship. The observable variable $\text{Damage}_i$

is defined to be equal to the latent variable whenever the latent variable is above $c$ and zero otherwise. The model looks as follows:

$$\text{Damage}_i = \begin{cases} \text{Damage}_i^* & \text{if } \text{Damage}_i^* > c \\ 0 & \text{if } \text{Damage}_i^* \leq c \end{cases} \tag{1}$$

$$\text{Damage}_i^* = \alpha_0 + \alpha_1 \text{MainVar}_i + \sum_{j=2}^{n} \alpha_j X_{i,j} + \varepsilon_i \tag{2}$$

where $\text{MainVar}_i = \{\text{Depth}_i$ – depth of the flood (in the basement or at the ground floor) for 'depth-damage' models and $\text{Rainfall}_i$ – rainfall (gauge or radar data) for 'rainfall-damage' models$\}$; $X_{i,j}$ represents a vector of $n$ control variables discussed in Table 2 and $\varepsilon_i$ the error term.

The censoring could be seen as follows: in case of very low damage figures, individuals may not report on the actual damage, but rather indicate that the damage equals zero. For example, in case the damage corresponds to mopping the floor, zero damage may be reported. However, when the damage exceeds a specific damage level $c$, it is expected that the actual level of damage has been reported. Hence, it is expected that several damage levels below $c$ are observed as zero damage, but when the damage exceeds the level $c$, the damage reports match better the actual ones.

A necessary condition to use the Tobit model is the assumption that the data generating process that determines the censoring is the same process that determines the outcome (see for example Wooldridge, 2010). In our application to flood damage, it seems reasonable to claim that factors such as flood depth, flood duration, building characteristics, behavioural predictors (such as emergency measures) and socio-economic indicators all have similar contributions to the amount of damage on the one hand and to whether or not there is any damage reported on the other hand.

## Discussion of the results

### Depth-damage model

Table 2 presents summary statistics on the variables included in the different model specifications. Table 3 presents the marginal effects of the estimated Tobit model (Eqn (2) with depth as the main variable of interest).[1] The

---

[1]Note that the Tobit coefficients refer to the latent variable, and this is not the effect we are interested in. We do want to take into account the zero values in the estimations, but in the end we only want to know the effect of the independent variables on the actual observed damage (comparable to the interpretation of an OLS coefficient). Therefore, we calculate the average marginal effects of the hazard and non-hazard indicators on the observed (or censored) damage (Bartus, 2005; Collis *et al.*, 2010).

**Table 2** Dependent and independent variables included in the 'depth-damage' model and the 'rainfall-damage' model

| | Description | Mean | Std. dev. |
|---|---|---|---|
| *Damage variables* | | | |
| Damage to building | Natural logarithm of the monetary damage to the households' building increased by 1 euro,* corrected for inflation (in euros) | 7.28 | 2.62 |
| Damage to content | Natural logarithm of the monetary damage to the households' content increased by 1 euro, corrected for inflation (in euros) | 4.53 | 3.95 |
| Total damage | Natural logarithm of the monetary damage to the households' building and content increased by 1 euro, corrected for inflation (in euros) | | |
| *Depth variables* | | | |
| Depth: basement | Depth of the flood in the basement (in cm) | 87.53 | 67.20 |
| Depth: ground floor | Depth of the flood at the ground floor (in cm) | 14.77 | 21.24 |
| *Rainfall variables* | | | |
| Rainfall: gauge | Monthly maximum daily rainfall accumulation at the closest rain gauge (in mm/day) | 53.66 | 34.58 |
| Rainfall: radar | Monthly maximum hourly rainfall accumulation in a radius of 1 km from the households' building location (in mm/h) | 17.91 | 8.94 |
| *Control variables* | | | |
| *Flood characteristics* | | | |
| Topography[†] | Relative height of the property measured as the percentage change in the height of the property as compared to the mean height in a concentric circle of 1 km | −5.69 | 11.77 |
| Part of building affected | Dummy variable that takes a value of one if the basement was flooded and zero otherwise | 0.79 | 0.40 |
| | Dummy variable that takes a value of one if the ground floor was flooded and zero otherwise | 0.71 | 0.46 |
| | Dummy variable that takes a value of one if the garage was flooded and zero otherwise | 0.59 | 0.49 |
| *Characteristics of the house* | | | |
| Type of building[‡] | Categorical variable distinguishing between three types of buildings: (1) detached; (2) semi-detached; and (3) terraced | – | – |
| Size of dwelling | Size of the dwelling (in squared metres) | 131.64 | 163.01 |
| *Behavioural and socio-economic characteristics* | | | |
| Recurrence | Number of times that the household has been flooded (times) | 3.27 | 3.67 |
| Risk awareness | Dummy variable that takes a value of one if the household was aware of their house being at risk just before the water entered the building; zero otherwise | 0.38 | 0.49 |
| Precautionary measures | Index composed of six precautionary measures created by PCA analysis[§] | 0.68 | 1.18 |
| Income | Categorical variable distinguishing between five per capita income categories: (1) below 1000 euros; (2) between 1000 and 1499 euros; (3) between 1500 and 1999 euros; (4) between 2000 and 2499 euros; and (5) above 2500 euros | – | – |

*In order to include the observations where zero damage was reported, we increased the level of damage by 1 euro before taking the natural logarithm.

[†]This variable is only included in the 'rainfall-damage' models.

[‡]This variable is not included in the models that use the damage to the content as the independent variable of interest.

[§]Based on a Principal Component Analysis (PCA) of six precautionary measures we have calculated a 'precautionary index'. The precautionary measures included are (1) moving content to a higher floor; (2) moving content to a higher area; (3) putting sandbags or similar to reduce the inflow of water; (4) purchase of pumping equipment; (5) moving vehicles; and (6) other.

reported standard errors are obtained by bootstrapping. The goodness of fit of the model is given by the squared correlation between the predicted damage levels based on the Tobit model and the observed values in the data set. Depending on the model specification, the predicted values share between 13% and 21% of their variance with the observed damage levels. There are five key findings.

First, the depth of the flood has an important impact on the damage, independently on the type of damage

(to the building, content, or total damage). For the basement models, we find that for each extra centimetre of water depth, there is an increase of 1.25% in building damage, 1.21% in content damage, and 1.05% in total damage.[2] The depth coefficients are higher in the ground

---

[2]Note that in case of a logarithmic transformation of the dependent variable $y$, coefficients $\beta$ should be recalculated by the following formula to obtain the effects in percent changes: $\%\Delta y = 100 \times (e^\beta - 1)$.

**Table 3** Tobit models, based on average marginal effects, using depth data as main variable

| Variables | Damage to building | | Damage to content | | Total damage | |
|---|---|---|---|---|---|---|
| *Flood characteristics* | | | | | | |
| Depth basement | 0.012*** | | 0.011*** | | 0.013*** | |
| | (0.00298) | | (0.00394) | | (0.00378) | |
| Depth ground floor | | 0.031*** | | 0.052*** | | 0.040*** |
| | | (0.00871) | | (0.0115) | | (0.0102) |
| Part: basement | | −0.190 | | 1.511** | | 1.095** |
| | | (0.398) | | (0.678) | | (0.550) |
| Part: ground floor | 0.415 | | 0.381 | | 0.171 | |
| | (0.439) | | (0.586) | | (0.519) | |
| Part: garage | 0.396 | 0.393 | 1.140* | −0.133 | 0.505 | 0.00501 |
| | (0.416) | (0.401) | (0.584) | (0.660) | (0.506) | (0.568) |
| *Characteristics of the house* | | | | | | |
| Type: semi-detached | 0.649 | 0.703* | | | 1.353** | 0.869* |
| | (0.475) | (0.410) | | | (0.537) | (0.513) |
| Type: terraced | −0.406 | −1.106* | | | −0.335 | −1.486** |
| | (0.514) | (0.591) | | | (0.572) | (0.697) |
| Size of the dwelling | 0.002 | 0.002 | 0.002 | 0.003** | 0.002 | 0.002* |
| | (0.00167) | (0.00134) | (0.00222) | (0.00124) | (0.00239) | (0.00104) |
| *Behavioural and socio-economic characteristics* | | | | | | |
| Reoccurrence | 0.027 | −0.081 | −0.192** | −0.085 | −0.043 | −0.088 |
| | (0.0564) | (0.0844) | (0.0842) | (0.120) | (0.0757) | (0.0919) |
| Risk awareness | −0.390 | −1.077 | −2.379*** | −1.423 | −0.899 | −1.140 |
| | (0.630) | (0.679) | (0.842) | (0.908) | (0.811) | (0.810) |
| Precautionary measures | 0.295* | 0.372* | 0.685** | 0.317 | 0.401 | 0.132 |
| | (0.179) | (0.206) | (0.346) | (0.377) | (0.258) | (0.320) |
| Income: 1000–1499 | −1.217* | −0.672 | 0.925 | 0.493 | 0.943 | 1.889 |
| | (0.660) | (0.665) | (1.321) | (1.556) | (1.107) | (1.435) |
| Income: 1500–1999 | −0.969* | −0.506 | 1.418 | 2.524* | 0.738 | 2.450* |
| | (0.557) | (0.549) | (1.276) | (1.474) | (1.035) | (1.366) |
| Income: 2000–2499 | −0.782 | −0.207 | 2.704** | 2.152 | 1.486 | 2.333 |
| | (0.615) | (0.666) | (1.310) | (1.503) | (1.104) | (1.435) |
| Income: 2500+ | −1.108* | −0.250 | 0.661 | 1.180 | 0.913 | 2.426* |
| | (0.600) | (0.629) | (1.379) | (1.513) | (1.038) | (1.451) |
| *Observations* | 220 | 199 | 247 | 221 | 265 | 235 |
| Censored observations | 21 | 14 | 100 | 85 | 21 | 24 |
| Uncensored observations | 199 | 185 | 146 | 136 | 193 | 218 |
| Log likelihood | −511.52 | −444.60 | −537.44 | −495.12 | −506.65 | −575.75 |
| Prob > $\chi^2$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Corr $(y^\sim y)^2$ | 0.1615 | 0.2110 | 0.159 | 0.169 | 0.184 | 0.128 |

Standard errors in parentheses.

***$P < 0.01$.

**$P < 0.05$.

*$P < 0.1$.

floor flood models, with an increase in damage to the building of 3.12%, to content of 5.37%, and to total damage of 3.82% for each additional centimetre of water depth. It is no surprise that the same increase in the depth level causes more severe damage (to buildings and content) in case of ground floor floods compared to basement floods. On the ground floor, the content is likely to be more valuable and more valuable building elements can be damaged.

Second, also other flood characteristics such as the part of the house that has been flooded affect the damage in some model specifications. The results show that in case more parts of the house are flooded, there is more damage. For ground floor floods, we find that in case the basement was also flooded, the damage to the content increases on average by 362%. For basement floods, we find that the damage to the content increases on average by 182% in case the garage was also flooded.

Third, in terms of the characteristics of the house we find that the type of the house affects the damage to the building and the total damage. The reference category of detached houses is compared to semi-detached houses and terraced buildings. For ground floor floods, the building damage for terraced houses is on average 67% lower than the damage to detached houses. For semi-detached houses the damage to the building is found to be 102% higher compared to detached houses. For the total damage, we find that on average the damage to terraced houses is 78% lower as compared to detached houses. Finally, we also find in some model specifications a significant impact of the size of the house: an increase of 1 m$^2$ of the size of the house is expected to lead to an increase of the monetary damage of 0.30% with respect to the damage to the content.

Fourth, we find that the attitude of individuals before and during the flood also affects the damage caused by the flood. First, we find that individuals learn from their experiences (reoccurrence): in case of basement floods the damage to the content decreases on average by 21% for each new flood experience. Second, we find that individuals that were aware of the flood risk experience less damage. For ground floor floods we find that being aware of the flooding risk reduces the damage to the building on average by 66%, while for basement floods we find that risk awareness reduce the damage to the content by 87%. Finally, we identified five possible precautionary measures and asked respondents to indicate whether they took those specific actions: moving (a part of) the contents to another floor, elevating (a part of) the contents at the same floor, placing sandbags or limiting in any other way the water inflow, purchasing a pump and/or preparing it for functioning, moving vehicles to a safer place. We find that taking additional precautionary measures increases the monetary damage: an increase of one in the precautionary measures index leads to an increase which ranges between 45% for the damage to the building caused by basement floods and 86% for the damage to the content caused by basement floods. While this finding may seem contradictory at first sight, it may reflect that, in case people are confronted with very severe floods, they desperately try to protect their belongings by moving them to a higher floor or placing sandbags. However, it could be that they act too late, especially since pluvial floods take place unexpectedly and quickly.

Finally, we analysed the impact of an increase in income, which is likely to be correlated with the level of education and with the labour market situation. The reference category here is respondents with a net monthly income lower than 1000 euro. The findings on the income variable differ depending on the type of damage. In case of damage to the building, we find that the individuals with high income experience less damage than the individuals in the reference category. This is different in case of damage to the content

or total damage, where we find that individuals being rich experience on average more damage than individuals in the reference category.

## Rainfall-damage model

Table 4 presents the marginal effects of the estimated Tobit model (Eqn (2) with rainfall as the main variable of interest). There are six key findings.

First, as compared to the 'depth-damage' models, the 'rainfall-damage' models explain in general a smaller fraction of the variation in the data. Depending on the model specification, the predicted values of the damage levels share between 9% and 19% of their variance with the observed values in the data set. In particular, the models which include the damage to the content and the total damage are found to explain only a small fraction of the variation in the data. This confirms previous results by Spekkers *et al.* (2013).

Second, the rain gauge data are significantly correlated with the damage to the building and the content. An increase of 1 mm in rainfall daily accumulation leads to an increase of 1.4% in building damage and 1.8% in content damage. There is no significant impact of rainfall measured by the radar on the damage. Overall, the correlation between the rainfall data and damage is smaller than for the depth. This finding could relate to the fact that a significant number of pluvial floods may be caused by blockages of inlets and sewer pipes as suggested by ten Veldhuis (2011). In these cases the observed damage is not necessarily related to heavy rainfall. The poor correlation between the rainfall measured by the radar and the damage may also relate to the lower accuracy of radar measurements with respect to rain gauges and the use of hourly instead of daily accumulations. Although the shorter term, sub-daily, rainfall accumulations are expected to be more indicative for pluvial flooding, because of the quick response times of sewer systems to rainfall, the differences in accuracy and statistical variability in hourly versus daily rainfall accumulations play a role as well. The spatial variability of hourly rainfall accumulations derived from the radar is expected to be higher than the one of the smoother daily rain gauge accumulations, which affects the computation of the correlation in the multivariate regression. Therefore, concluding that radar measurements do not have potential for building rainfall-damage models could be misleading. Further investigation could prove them to be useful.

Third, for the models that include the damage to the building as an explanatory variable we find that the topography has a significant impact on the damage and a house that is located higher than the other houses in the neighbourhood is likely to experience less damage. A one percentage point increase in the altitude of the property as

**Table 4** Tobit models, based on average marginal effects, using rainfall data as main variable

| Variables | Damage to building | | Damage to content | | Total damage | |
|---|---|---|---|---|---|---|
| *Rainfall, topography, and flood characteristics* | | | | | | |
| Rainfall: gauge data | 0.014*** | | 0.018** | | 0.008 | |
| | (0.00508) | | (0.00836) | | (0.00528) | |
| Rainfall: radar data | | 0.026 | | −0.013 | | −0.006 |
| | | (0.0193) | | (0.0364) | | (0.0227) |
| Topography | −0.032** | −0.032** | 0.008 | 0.004 | −0.017 | −0.017 |
| | (0.0153) | (0.0159) | (0.0261) | (0.0260) | (0.0168) | (0.0170) |
| Part: basement | −0.886** | −0.846** | 0.587 | 0.688 | 0.0733 | 0.126 |
| | (0.384) | (0.373) | (0.794) | (0.826) | (0.532) | (0.533) |
| Part: ground floor | 0.384 | 0.366 | 0.365 | 0.504 | 0.315 | 0.374 |
| | (0.466) | (0.509) | (0.678) | (0.684) | (0.472) | (0.497) |
| Part: garage | 0.304 | 0.300 | 1.589*** | 1.584*** | 0.487 | 0.455 |
| | (0.383) | (0.377) | (0.595) | (0.582) | (0.455) | (0.444) |
| *Characteristics of the house* | | | | | | |
| Type: semi-detached | 0.371 | 0.350 | | | 0.508 | 0.461 |
| | (0.392) | (0.411) | | | (0.465) | (0.491) |
| Type: terraced | −1.067** | −1.181** | | | −1.045** | −1.174** |
| | (0.500) | (0.478) | | | (0.526) | (0.511) |
| Size of the dwelling | 0.001 | 0.001 | 0.003** | 0.003* | 0.002* | 0.002 |
| | (0.000950) | (0.00104) | (0.00143) | (0.00175) | (0.00106) | (0.00104) |
| *Behavioural and socio-economic characteristics* | | | | | | |
| Reoccurrence | 0.049 | 0.043 | −0.118 | −0.125 | 0.026 | 0.018 |
| | (0.0486) | (0.0517) | (0.0887) | (0.0889) | (0.0534) | (0.0492) |
| Risk awareness | −0.395 | −0.387 | −1.876** | −1.907** | −0.366 | −0.349 |
| | (0.512) | (0.527) | (0.884) | (0.907) | (0.640) | (0.584) |
| Precautionary measures | 0.138 | 0.128 | 0.342 | 0.399 | −0.196 | −0.179 |
| | (0.167) | (0.178) | (0.387) | (0.374) | (0.273) | (0.252) |
| Income 1000–1499 | −1.541*** | −1.236** | 0.510 | 0.817 | 0.00657 | 0.145 |
| | (0.581) | (0.564) | (1.972) | (1.462) | (0.981) | (0.982) |
| Income 1500–1999 | −1.666*** | −1.539*** | 1.094 | 1.312 | −0.220 | −0.141 |
| | (0.565) | (0.536) | (1.915) | (1.416) | (0.984) | (0.956) |
| Income 2000–2499 | −1.297** | −1.180** | 1.450 | 1.752 | 0.0198 | 0.126 |
| | (0.635) | (0.561) | (1.928) | (1.487) | (1.036) | (1.012) |
| Income 2500+ | −1.200** | −1.188** | −0.185 | 0.0752 | −0.126 | −0.0372 |
| | (0.547) | (0.570) | (1.907) | (1.529) | (1.045) | (1.025) |
| *Observations* | 218 | 218 | 236 | 236 | 253 | 253 |
| Censored observations | 15 | 15 | 83 | 83 | 25 | 25 |
| Uncensored observations | 203 | 203 | 153 | 153 | 228 | 228 |
| Log likelihood | −485.55 | −488.61 | −546.80 | −549.27 | −607.10 | −608.22 |
| Prob > $\chi^2$ | 0.03 | 0.04 | 0.00 | 0.07 | 0.06 | 0.11 |
| Corr $(y\tilde{\ }y)^2$ | 0.1867 | 0.1621 | 0.1119 | 0.0887 | 0.0955 | 0.0859 |

Standard errors in parentheses.

***$P < 0.01$.

**$P < 0.05$.

*$P < 0.1$.

compared to the mean altitude in the neighbourhood decreases the damage to the building by 3.13%.

Fourth, also the part of the house that has been flooded affects the damage. As in the 'depth-damage' models damage to the content increases in case the garage was also flooded. Further, we find that also the type of the house affects the damage to the building and the total damage. For terraced houses the damage to the building is on average 66% lower than the damage to detached houses (the reference category), while the total damage on terraced houses is on average 64% lower than the damage to detached houses. Finally, we find that in some model

specifications there is a significant impact of the size of the house and an increase of 1 m$^2$ of the size of the house is expected to lead to an increase of up to 0.31% with respect to the damage to the content.

Fifth, with respect to the behavioural and socio-economic characteristics, the results are similar to the 'depth-damage' models: being aware of the flood risk reduces damage. Risk awareness is found to decrease the content damage by 85% in case rain gauge data are used.

Finally, we analysed the impact of an increase in income. As in the 'depth-damage' models the findings on the income variable differ depending on the type of damage. In case of damage to the building the individuals with high income experience less damage while for damage to the content richer individuals experience on average more damage.

### Limitations and scope for future research

This study is innovative as it is the first to estimate and compare multivariate 'depth-damage' and 'rainfall-damage' models using the same data set, next to the focus on pluvial floods. However, there are some limitations to the study as well as there is a scope for future research.

First, this study is limited to estimating the impact of the depth of the flood and the rainfall at the time of the flood on the monetary damage to residential property caused by pluvial floods. It is important to note that this reflects only a part of the total damage caused by floods and that additional information is required to be incorporated in any cost–benefit analysis on flood management. Other authors have pointed to the damage to commercial and public property (Kreibich *et al.*, 2010) and damage caused by other types of floods, such as ground water floods (Kreibich and Thieken, 2008). Also, pluvial floods can have a large impact on non-monetary damage as they may have a significant impact on the health of those individuals that have been flooded. In addition to a physical non-monetary impact, pluvial floods can also affect the well-being of individuals. In Van Ootegem and Verhofstadt (2016), the impact of pluvial floods on subjective well-being is studied.

Second, the models that we use explain between 13% and 21% (for the 'depth-damage' Tobit models) and between 9% and 19% (for the 'rainfall-damage' Tobit models) of the variation in the damage caused by the flood. The explanatory power of the model could be further improved by including additional control variables that could explain the flood damage, such as variables related to the characteristics of the flood (e.g. velocity or the level of sediment of the water that entered the property) or variables related to the socio-economic characteristics of the households affected by the flood (e.g. total value of the building or content owned or rented by the household or detailed data on the wealth of households instead of income data).

Third, the data have been collected based on recall from an event that may have happened more than 10 years ago. This may raise problems regarding the accuracy of some of the variables. It is unclear to what extent respondents are able to remember in detail the amount of the damage, the depth of the flood or the month in which the flood took place. In order to test for the robustness of our results, we restricted the sample to the past 10 years (period 2002–2012) since we expect that respondents are more able to remember recent events. This did not change the results significantly (results are available from the authors upon request).

Fourth, the sample includes only households that experienced a flood in the past two decades. As result, it is possible that the results cannot be generalised to all Flemish households, but only to the Flemish households that live in regions that are prone to pluvial floods. In future research, we may also include households that have not been flooded in order to make the results generalisable for the entire of Flanders.

## Conclusions

In this article we explore the potential of two distinct types of multivariate flood damage models used to estimate the impact of different drivers for pluvial floods, namely multivariate 'depth-damage' models and 'rainfall-damage' models. In a first step, we estimate multivariate 'depth-damage' models in which we include a number of control variables such as building characteristics, behavioural predictors and income. Up to now; most studies estimating flood damage models have focused on estimating bivariate depth-damage curves and only few have estimated multivariate 'depth-damage' models, including a number of control variables (Freni *et al.*, 2010; Van Ootegem *et al.*, 2015). In a second step, we estimate a 'rainfall-damage' model. The potential of 'rainfall-damage' models is even less explored with as notable exceptions the studies by Ririassa and Hoen (2010), Zhou *et al.* (2012), and Spekkers *et al.* (2013). Estimating the direct relationship between the rainfall and monetary damage has – at least – two major advantages. First, instead of using ex-post survey data or simulating the expected depth of a flood based on hydraulic models, which are time and labour intensive techniques, we can investigate the direct relationship between the rainfall accumulation and the damage. Second, by using exogenous data on the rainfall accumulation, we are able to exclude potential endogeneity problems related to a self-reported depth variable.

We used survey data of 346 Flemish households that were victims of pluvial floods. We collected data on a wide range of variables, including the depth of the flood (in the basement and at the ground floor), the monetary damage

to different parts of the house (building, content, and total damage), the prevention measures taken by the households, socio-economic characteristics such as the income of the individuals that have been flooded, etc. In addition, we included rainfall data from two sources, namely the rain gauge network of the Flemish Environment Agency and radar rainfall data of the RMI. We applied an OLS estimation and a Tobit estimation. The latter is specifically designed to deal with the issue of many zero damage observations.

For the 'depth-damage' models, we find that the depth of the flood has a significant impact on the damage. For the basement models, we find that for an extra centimetre of water depth, there is an increase in damage that ranges between 1.05% and 1.25%, depending on the type of damage (building, content, or total). The depth coefficients are higher in the ground floor flood models and an increase of 1 cm in the depth of the flood at the ground floor is expected to lead to an increase in the damage which ranges between 3.12% and 5.37%. This difference can be attributed to the fact that more valuable content and building elements can be damaged at the ground floor.

For the 'rainfall-damage' models, we find a different impact of rainfall accumulation depending on the source of the data. We distinguish between rainfall gauge data and radar data. For the rainfall gauge data, we find that an increase of 1 mm in rainfall daily accumulation leads to an increase which ranges between 1.4% to 1.8%, depending on the type of damage (building or content). For the rainfall radar data, we find no significant effect. This finding could relate to the lower accuracy of radar measurements with respect to rain gauges, the use of hourly instead of daily accumulations and the higher spatial variability of hourly rainfall accumulations derived from the radar compared to the smoother daily rain gauge accumulations. The overall lower correlation between the rainfall and the damage figures may also relate to the fact that a significant number of pluvial floods may be caused by blockages of inlets and sewer pipes. In these cases the observed damage is not necessarily related to heavy rainfall. Finally, for both the 'depth-damage' and the 'rainfall-damage' models we find that also non-hazard indicators are important for explaining pluvial flood damage and are therefore interesting to be included in future flood damage research. The information obtained from these models can also help to design early-warning systems and target the early-warning systems to specific groups in the society.

## Acknowledgements

## References

Apel H., Aronica G.T., Kreibich H. & Thieken A.H. Flood risk analyses – how detailed do we need to be? *Nat Hazards* 2009, **49**, (1), 79–98.

Bartus T. Estimation of marginal effects using margeff. *Stata J* 2005, **5**, (3), 309–329.

Blanc J., Hall J., Roche N., Dawson R., Cesses Y., Burton A. & Kilsby C. Enhanced efficiency of pluvial flood risk estimation in urban areas using spatial-temporal rainfall simulations. *J Flood Risk Manage* 2012, **5**, (2), 143–152.

Collis J., Grayson A. & Johal S. *Econometric analysis of alcohol consumption in the UK*. London: HM Revenue Customs, 2010.

Elmer F., Thieken A.H., Pech I. & Kreibich H. Influence of flood frequency on residential building losses. *Nat Hazards Earth Syst Sci* 2010, **10**, 2145–2159.

Ernst J., Dewals B., Archambeau P., Detrembleur S., Erpicum S. & Pirotton M. Integration of accurate 2D inundation modelling, vector land use database and economic damage evaluation. In: *Flood Risk Management: Research and Practice, Proceedings of the European Conference on Flood Risk Management – Flood Risk 2008*. Rotterdam: Balkema, 2008, 1643–1653.

European Environment Agency *Urban adaptation to climate change in Europe*. Copenhagen: European Environment Agency, 2012, 143.

Freni G., La Loggia G. & Notaro V. Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Sci Technol* 2010, **61**, (12), 2979–2993.

Goudenhoofdt E. & Delobbe L. Generation and verification of rainfall estimates from 10-year volumetric weather radar measurements. *J Hydrometeorol* 2016, **17**, (4), 1223–1242. doi: 10.1175/JHM-D-15-0166.1.

Hurford A.P., Priest S.J., Parker D.J. & Lumbroso D.M. The effectiveness of extreme rainfall alerts in predicting surface water flooding in England and Wales. *Int J Climatol* 2012, **32**, (11), 1768–1774.

IPCC Managing the risks of extreme events and disasters to advance climate change adaptation. In: *A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge and New York: Cambridge University Press, 2012, 582.

Jonkman S.N., Bockarjova M., Kok M. & Bernardini P. Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecol Econ* 2008, **66**, (1), 77–90.

Kellens W., Vanneuville W., Verfaillie E., Meire E., Deckers P. & De Maeyer P. Flood risk management in Flanders: past developments and future challenges. *Water Resour Manage* 2013, **27**, (10), 3585–3606.

Kreibich H. & Thieken A.H. Assessment of damage caused by high groundwater inundation. *Water Resour Res* 2008. doi: 10.1029/2007WR006621.

Kreibich H., Seifert I., Merz B. & Thieken A.H. Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector. *Hydrol Sci J* 2010, **55**, (8), 1302–1314.

McDonald J.F. & Moffitt R.A. The uses of Tobit analysis. *Rev Econ Stat* 1980, **62**, (2), 318–321.

Merz B., Kreibich H. & Lall U. Multi-variate flood damage assessment: a tree-based data-mining approach. *Nat Hazards Earth Syst Sci* 2013, **13**, 53–64.

Messner F. & Meyer V. Flood damage, vulnerability and risk perception – challenges for flood damage Research. In: J. Schanze, E. Zeman, & J. Marsalek, eds. *Flood risk management: hazards, vulnerability and mitigation measures*. NATO science series. Dordrecht, The Netherlands: Springer, 2006, 149–167.

de Moel H. & Aerts J.C.J.H. Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. *Nat Hazards* 2011, **58**, (1), 407–425.

National Climate Commission. Belgian national climate change adaptation strategy. Report prepared by the National Climate Commission, December 2010.

Pistrika A.K. & Jonkman S.N. Damage to residential buildings due to flooding of New Orleans after hurricane Katrina. *Nat Hazards* 2009, **54**, (2), 413–434.

Ririassa H.A. & Hoen A.R. Neerslag en Schade – Onderzoek naar het verband tussen neerslag en de schadelast voor brandverzekeraars met het oog op de klimaatverandering. Den Haag, 2010, p. 20.

Spekkers M.H., Kok M., Clemens F.H.L.R. & ten Veldhuis J.A.E. A statistical analysis of insurance damage claims related to rainfall extremes. *Hydrol Earth Syst Sci* 2013, **17**, 913–922.

Tobin J. Estimation of relationships for limited dependent variables. *Econometrica* 1958, **26**, (1), 24–36.

Van Dyck J. & Willems P. Probabilistic flood risk assessment over large geographical regions. *Water Resour Res* 2013, **49**, 3330–3344.

Van Ootegem L. & Verhofstadt E. Well-being, life satisfaction and capabilities of flood disaster victims. *Environ Impact Assess Rev* forthcoming 2016, **57**, 134–138. doi: 10.1016/j.eiar.2015.12.001.

Van Ootegem L., Verhofstadt E., Van Herck K. & Creten T. Multivariate pluvial flood damage models. *Environ Impact Assess Rev* 2015, **54**, 91–100.

ten Veldhuis J.A.E. How the choice of flood damage metrics influences urban flood risk assessment. *J Flood Risk Manage* 2011, **4**, (4), 281–287.

Villarini G. & Krajewski W.F. Review of the different sources of uncertainty in single-polarization radar-based estimates of rainfall. *Surv Geophys* 2010, **31**, 107–129.

Willems P. Revision of urban drainage design rules after assessment of climate change impacts on precipitation extremes at Uccle, Belgium. *J Hydrol* 2013, **496**, 166–177.

Willems P., Olsson J., Arnbjerg-Nielsen K., Beecham S., Pathirana A., Bülow Gregersen I., Madsen H. & Nguyen V.T.V. *Impacts of climate change on rainfall extremes and urban drainage*. London, UK: IWA Publishing, 2012, 252.

Wooldridge J.M. *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT Press, 2010.

Zhou Q., Mikkelsen P.S., Halsnaes K. & Arnbjerg-Nielsen K. Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. *J Hydrol* 2012, **414**, 539–549.

# Appendix

In a first step, we estimated a simple Ordinary Least Squares (OLS) model, considering a number of socio-economic characteristics. The OLS model analysed the impact of the main variable of interest ($\text{MainVar}_i$), which is the depth of the flood for 'depth-damage' models and the rainfall at the time of the flood for the 'rainfall-damage' models, on the reported damage expressed as a natural logarithm of the monetary damage (to the building, content, or total damage) for a household $i$. This variable linearly depends on depth and a vector of $n$ control variables $X_i$. Hence, we estimated the following model:

$$\text{Damage}_i = \alpha_0 + \alpha_1 \text{MainVar}_i + \sum_{j=2}^{n} \alpha_j X_{i,j} + \varepsilon_i \qquad (A1)$$

where $\text{MainVar}_i = \{\text{Depth}_i$ – depth of the flood (in the basement or at the ground floor) for 'depth-damage' models and

$\text{Rainfall}_i$ – rainfall intensity (gauge or radar data) for 'rainfall-damage' models$\}$; $X_{i,j}$ represents a vector of $n$ control variables discussed in Table 2 and $\varepsilon_i$ the error term. Table A1 presents the marginal effects of the estimated OLS estimation (Eqn (A1) with depth as the main variable of interest). Table A2 marginal effects of the estimated OLS model (Eqn (A1) with rainfall as the main variable of interest). The reported standard errors are bootstrapped.

The results show that there are differences in the results of the OLS and Tobit models, in particular with respect to the magnitude of the effect for some of the control variables. However, overall the main findings remain consistent across the different econometrical specifications.

**Table A1** Ordinary Least Squares estimations for the 'depth-damage' model, based on average marginal effects, using depth data as main variable

| Variables | Damage to building | | Damage to content | | Total damage | |
|---|---|---|---|---|---|---|
| *Flood characteristics* | | | | | | |
| Depth basement | 0.012*** | | 0.011*** | | 0.010*** | |
| | (0.00260) | | (0.00336) | | (0.00283) | |
| Depth ground floor | | 0.030*** | | 0.052*** | | 0.036*** |
| | | (0.00797) | | (0.0111) | | (0.00797) |
| Part: basement | | −0.189 | | 1.318** | | 0.588 |
| | | (0.374) | | (0.594) | | (0.429) |
| Part: ground floor | 0.393 | | 0.303 | | 0.196 | |
| | (0.378) | | (0.510) | | (0.417) | |
| Part: garage | 0.372 | 0.366 | 1.074** | 0.0225 | 0.387 | −0.0430 |
| | (0.363) | (0.382) | (0.471) | (0.539) | (0.353) | (0.420) |
| *Characteristics of the house* | | | | | | |
| Type: semi-detached | 0.603 | 0.671* | | | 0.542 | 0.631 |
| | (0.456) | (0.356) | | | (0.446) | (0.408) |
| Type: terraced | −0.379 | −1.039* | | | −0.442 | −1.396*** |
| | (0.494) | (0.537) | | | (0.445) | (0.511) |
| Size of the dwelling | 0.002 | 0.001 | 0.002 | 0.003** | 0.002 | 0.002 |
| | (0.00164) | (0.000972) | (0.00178) | (0.00128) | (0.00150) | (0.00108) |
| *Behavioural and socio-economic characteristics* | | | | | | |
| Recurrence | 0.024 | −0.078 | −0.130*** | −0.071 | −0.029 | −0.099 |
| | (0.0505) | (0.0734) | (0.0479) | (0.0872) | (0.0548) | (0.0784) |
| Risk awareness | −0.383 | −1.021* | −2.211*** | −1.222 | −0.554 | −0.495 |
| | (0.611) | (0.590) | (0.774) | (0.752) | (0.601) | (0.613) |
| Precautionary measures | 0.278 | 0.349** | 0.567** | 0.249 | 0.210 | −0.0605 |
| | (0.185) | (0.178) | (0.289) | (0.313) | (0.195) | (0.260) |
| Income: 1000–1499 | −1.097* | −0.627 | 0.977 | 0.431 | −0.436 | 0.191 |
| | (0.648) | (0.610) | (1.125) | (1.470) | (0.772) | (1.113) |
| Income: 1500–1999 | −0.884 | −0.468 | 1.232 | 2.190 | −0.352 | 0.634 |
| | (0.541) | (0.519) | (1.102) | (1.351) | (0.685) | (1.038) |
| Income: 2000–2499 | −0.690 | −0.152 | 2.479** | 1.968 | 0.358 | 1.161 |
| | (0.572) | (0.596) | (1.150) | (1.423) | (0.724) | (1.113) |
| Income: 2500+ | −1.031* | −0.210 | 0.506 | 0.991 | −0.443 | 0.614 |
| | (0.568) | (0.587) | (1.166) | (1.417) | (0.695) | (1.090) |
| Constant | 5.982*** | 7.459*** | 2.276** | 1.858 | 6.371*** | 6.620*** |
| | (0.729) | (0.648) | (1.096) | (1.374) | (0.857) | (1.163) |
| Observations | 220 | 199 | 247 | 221 | 242 | 214 |
| $R^2$ | 0.162 | 0.211 | 0.162 | 0.171 | 0.128 | 0.184 |

Bootstrapped robust standard errors in parentheses.

***$P < 0.01$.

**$P < 0.05$.

*$P < 0.1$.

**Table A2** Ordinary Least Squares estimations for the 'depth-damage' model, based on average marginal effects, using rainfall data as main variable

| Variables | Damage to building | | Damage to content | | Total damage | |
|---|---|---|---|---|---|---|
| *Rainfall, topography, and flood characteristics* | | | | | | |
| Rainfall: gauge data | 0.013*** | | 0.017** | | 0.008* | |
| | (0.00502) | | (0.00685) | | (0.00492) | |
| Rainfall: radar data | | 0.024 | | 0.001 | | −0.005 |
| | | (0.0181) | | (0.000850) | | (0.0206) |
| Topography | −0.029** | −0.029** | 0.004 | 0.004 | −0.015 | −0.015 |
| | (0.0139) | (0.0144) | (0.0211) | (0.0202) | (0.0154) | (0.0144) |
| Part: basement | −0.858** | −0.816** | 0.515 | 0.931 | 0.0357 | 0.0823 |
| | (0.339) | (0.353) | (0.697) | (0.569) | (0.481) | (0.481) |
| Part: ground floor | 0.356 | 0.343 | 0.268 | 0.824 | 0.299 | 0.350 |
| | (0.473) | (0.443) | (0.593) | (0.519) | (0.440) | (0.433) |
| Part: garage | 0.281 | 0.278 | 1.470*** | 0.882* | 0.469 | 0.439 |
| | (0.357) | (0.357) | (0.523) | (0.505) | (0.398) | (0.412) |
| | | | | | | |
| *Characteristics of the house* | | | | | | |
| Type: semi-detached | 0.347 | 0.326 | | | 0.446 | 0.401 |
| | (0.376) | (0.370) | | | (0.408) | (0.429) |
| Type: terraced | −1.005** | −1.115** | | | −0.983** | −1.108 |
| | (0.448) | (0.433) | | | (0.468) | (0.497) |
| Size of the dwelling | 0.00103 | 0.000861 | 0.00301** | 0.00207** | 0.00171** | 0.00157 |
| | (0.00102) | (0.000845) | (0.00146) | (0.000854) | (0.000867) | (0.000996) |
| | | | | | | |
| *Behavioural and socio-economic characteristics* | | | | | | |
| Reoccurrence | 0.045 | 0.039 | −0.089 | −0.096* | 0.022 | 0.015 |
| | (0.0501) | (0.0474) | (0.0640) | (0.0531) | (0.0464) | (0.0540) |
| Risk awareness | −0.393 | −0.383 | −1.682** | −1.379** | −0.367 | −0.349 |
| | (0.474) | (0.498) | (0.795) | (0.690) | (0.529) | (0.540) |
| Precautionary measures | 0.139 | 0.130 | 0.285 | 0.376 | −0.170 | −0.153 |
| | (0.149) | (0.153) | (0.317) | (0.289) | (0.230) | (0.222) |
| Income 1000–1499 | −1.464*** | −1.172** | 0.488 | 0.470 | −0.0345 | 0.0991 |
| | (0.542) | (0.530) | (1.342) | (1.232) | (0.877) | (0.929) |
| Income 1500–1999 | −1.578*** | −1.455*** | 0.913 | 0.949 | −0.229 | −0.154 |
| | (0.500) | (0.513) | (1.339) | (1.233) | (0.868) | (0.929) |
| Income 2000–2499 | −1.228** | −1.111** | 1.282 | 1.234 | −0.00279 | 0.0994 |
| | (0.542) | (0.520) | (1.343) | (1.261) | (0.936) | (0.971) |
| Income 2500+ | −1.131** | −1.114** | −0.291 | −0.149 | −0.178 | −0.0942 |
| | (0.499) | (0.511) | (1.361) | (1.277) | (0.872) | (0.882) |
| Constant | 8.040*** | 8.226*** | 2.345 | 2.693* | 6.737*** | 7.188 |
| | (0.682) | (0.653) | (1.557) | (1.383) | (1.104) | (1.155) |
| Observations | 218 | 218 | 236 | 278 | 253 | 253 |
| $R^2$ | 0.187 | 0.162 | 0.113 | 0.086 | 0.096 | 0.086 |

Bootstrapped robust standard errors in parentheses.

***$P < 0.01$.

**$P < 0.05$.

*$P < 0.1$.

# Development and evaluation of a coupled 1D-2D urban pluvial flood model

Victor NTEGEKA[a], Damian MURLA[a], *Patrick WILLEMS[a,b]

KU Leuven, Department of Civil Engineering, Hydraulics Section, Kasteelpark Arenberg, 40, BE-3001 Leuven, Belgium

Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Belgium

*Corresponding author

E-mail address: patrick.willems@kuleuven.be

**Abstract**

Due to urbanization trends and climate change effects, there is an increasing need for urban flood simulation, forecasting, warning and control systems. One- or two-dimensional (1D-2D) hydraulic models form the core of such systems. Their success strongly depends on their detail and accuracy, and - for real time forecasting and control applications - also on the computational time. Because these needs are conflicting, in this study an efficient coupled 1D-2D hydraulic modelling approach has been tested, linking the underground sewer system with the overland surface for an urban catchment in Gent (Belgium). For the surface flow modelling, different approaches with various complexities were compared, while searching for a compromise between precision and simulation time. Results emphasize the importance of finding an optimal mesh resolution for urban inundation modelling, taking infiltration in green areas into account, and the importance of the dynamic interaction between the sewer network and the river system.

**Keywords:** urban drainage, urban flooding, flood modelling, pluvial flooding, surface flooding

## 1    Introduction

It has become more apparent that increased urbanization and upward climate trends are exacerbating pressure on stormwater management systems leading to negative social-economic and environmental effects (Lehmann et al., 2015; Willems et al., 2012). There is indeed an increased risk of urban flooding due to a convergence of factors related to increased percentage of impervious surfaces due to urbanization (Chen et al., 2015; Marshall, 2007), an increased frequency of intense rainfall events (Arnbjerg-Nielsen, 2013; Kunkel et al., 2013; Mallakpour and Villarini, 2015; Ntegeka and Willems, 2008) and aging drainage systems. The increased concentration of population and property explains why urban regions tend to have more pronounced damages under flooding conditions (Crompton and McAneney, 2008; Priest et al., 2011). Enhanced flood damages in cities have raised concerns about the effectiveness of flood proof measures. There hence is a need to better apply the insights gained from urban drainage models (Dawson et al., 2008; Wright, 2014). In this context, the use of dynamic models is increasingly tenable because of faster numerical algorithms (Casulli and Stelling, 2013; Ghimire et al., 2013; Smith et al., 2015), access to Geographical Information System (GIS) tools, availability of high-resolution terrain data, availability (Thorndahl et al., 2014) of new data collection sensors and advancements in computing technologies (Glenis et al., 2013). Dynamic models, solving the appropriate approximation of mass and momentum conservation shallow water equations, integrate detailed process descriptions covering a wide range of flow patterns (Hunter et al., 2008; Néelz and Pender, 2013). The application of dynamic models for urban flooding, however, remains a challenging problem in practice. Computational times remain prohibitive and the availability of data to calibrate the models is a consistent problem. Nonetheless, recent studies have shown that optimal approaches that account for modelling objectives, different sources of in-situ data, catchment characteristics and computational demands are complementary guides for solving urban drainage challenges.

25    Traditionally, urban drainage models only describe the underground system, by solving the one-

26    dimensional (1D) full St Venant equations. Urban flood modelling, however, requires such models to

27    be extended or linked to a surface inundation model. This allows the storm water drainage to be

28    described both in the underground system and the surface system. That is why this modelling

29    approach is also called dual drainage method. The surface system model in this method can be zero-

30    dimensional (0D), 1D or two-dimensional (2D). This leads to several concepts ranging from 1D-0D,

31    coupled 1D-1D, full 2D and the more sophisticated coupled 1D-2D approaches. 1D-0D modelling was

32    the first concept adopted for modelling sewer floods. In this concept, the sewer pipe system is

33    discretized as a set of nodes (0D) connected by links (1D) without interaction between the sewer

34    underground system and the overland flow. Overflows that overwhelm the sewer system are either

35    assumed lost or stored in virtual reservoirs above the nodes. These virtual reservoirs are also called

36    flood cones. Because there is no direct interaction considered between these flood cones, no surface

37    water movement is considered. The surcharged water that leads to surface inundations remains in the

38    localized flood cones. This approach can be improved by transforming the 0D surface inundation

39    model to a 1D model. The 1D-1D approach involves the coupling of the 1D underground network with

40    the surface system, schematized by a 1D network. This implies that urban overland flow processes are

41    conceptualized as a set of storage areas connected by links which are then connected to nodes of the

42    1D sewer system (Djordjević et al., 2005; Lhomme et al., 2004). The storage areas represent the lower

43    lying areas at the surface, where water can be ponded, and the links are the higher elevations

44    pathways, schematized by overflow units, in between the storage areas. The predefinition of overland

45    flow paths in such 1D surface model poses a problem for cases where flood flow paths could vary such

46    as in areas of large flood depth and in flow separation and recirculation areas (Ghostine et al., 2014).

47    Moreover, the simplification of the overland surface by a 1D model introduces additional uncertainties

48    (Leandro et al., 2009) and potentially increases the model setup time. A further advancement of the

49    surface model into a 2D model is necessary for a better representation of surface geometries. 1D-2D

50    approaches involve the coupling of the 1D underground models with a full 2D model for the overland

51  flow (Seyoum et al., 2012; Simões et al., 2011; Pina et al., 2016; Yu et al., 2016). They are considered

52  more physically based especially regarding the way flow conditions on the surface are conceptualized.

53  This concept is widely applied as evidenced from various software packages that have incorporated

54  the methodology, e.g. InfoWorks ICM (Innovyze, 2014), SOBEK, MIKE URBAN. However, the

55  computational demands of 1D-2D models are still prohibitive for some applications such as

56  operational forecasting of floods where 1D approaches are preferred.

57  The purpose of this study was to evaluate different options for the 1D-2D dual drainage approach

58  based on the two criteria that relate to the needs discussed above, but which are often conflicting:

59  accuracy and computational time. This involved setting up and testing the model performance based

60  on available data from diverse sources. Different resolutions were tested for the surface system

61  model, from 0D to 2D, and different mesh sizes for the 2D approach. The optimal resolution was

62  determined with the goal of achieving sufficient accuracy while maintaining reasonable runtimes. To

63  reach this goal, the mesh sizes were not kept constant, but were varied depending on the area (2D

64  meshes with different resolutions for different types of areas). Evaluation of the accuracy of the

65  models in simulating urban floods is, however, very challenging. This is because of the typical lack of

66  observations on such floods, mainly due to their short duration and local nature. This study was

67  applied to an urban catchment in Belgium, where data typically available from diverse sources had to

68  be combined. It includes rainfall data from rain gauges and radar, water level observations from

69  monitoring points in the sewer network, information on reported floods by the city authorities and

70  from social media. The different modelling approaches were implemented using the Info Works ICM

71  software (Innovyze, 2014).

72  To achieve these objectives, the following steps were carried out: (1) identification of the flood risk

73  areas; (2) 2D surface mesh sensitivity analysis and identification of the optimal resolution for the 2D

74  surface flow model; (3) performance testing of the final coupled 1D-2D model. The few records of

75  urban floods necessitated the need for sensitivity analysis to evaluate the importance of the model

76    structure and parameters. Additionally, measures were taken to constrain data input uncertainty to

77    allow for a more clear identification of model imperfections. Rainfall data at the small spatial and

78    temporal resolutions required for urban drainage modelling indeed is often more lacking and/or

79    subjected to higher uncertainty (Berne et al., 2004; Bruni et al., 2015; Ochoa-Rodriguez et al., 2015)

80    than is the case for river flood applications.

81

82    **2      Materials and methods**

83    **2.1     Case study**

84    The selected study area is the urban drainage system of the villages of Oostakker and Sint

85    Amandsberg, two districts of the city of Gent, Belgium (Fig. 1). This area is mainly constituted by a

86    residential domain covering about 2747ha, and a population density of 15.8 inhabitants/ha. Its sewer

87    system design is essentially combined, meaning that waste and storm water are transported through

88    the same pipe network. The ground is predominantly pervious and its surface topography flat with an

89    average slope of 0.0035m/m.

90    **2.2     Preliminary set up**

91    Prior to the model development, a preliminary setup and standardization of the required data was

92    performed. At first, a set of topographical data was required for the development of the urban surface

93    approach including a Digital Terrain Model (DTM) of 0.5m horizontal resolution and a vertical accuracy

94    of 0.05m (AGIV, 2015), which was provided by the Flemish Agency for Geographical Information (AGIV)

95    as well as building features, streets and land use polygons. Building data were used to accurately

96    represent no flow zones onto the surface. In order to reduce unnecessary detail, a two-step

97    simplification and improvement was applied. First, adjacent building polygons with overlapping sides

98    were joined to reduce the amount of information to be processed. Then, complex shapes were

99    simplified with a minimum area of 10m$^2$ and a simplification tolerance of 2m. The simplification was

100   attained by joining adjacent street polygons, filling in holes smaller than 10m$^2$ size and closing residual

101   gaps between street and building polygons. Land use mapping was needed to help identifying the

102   proper surface cover types. Its properties were obtained from two sources: AGIV provided surface

103   boundaries covered by vegetation within the study area, Open Street Maps (Haklay and Weber, 2008;

104   Leandro et al., 2016) provided a variety of further reclassified information of impervious, pervious

105   areas, infiltration and roughness zones.

106   **Fig. 1. Here**

107   **2.3   Precipitation and pumping chamber data**

108   Four synthetic rainfall design storms, the so-called composite storms developed following the Chicago

109   storm design concept (Keifer and Chu, 1957), with return periods of 2, 5, 10 and 20 years were used

110   for the analysis (Fig. 2). These storms are based on rainfall intensity-duration-frequency (IDF) curves

111   for Belgium and are commonly applied for the design of urban drainage systems in the region

112   (Willems, 2013). Each composite storm had a time step of five minutes and a duration of two days.

113   The four events were used for generating flood hazard zones through simulation in the 1D sewer

114   network and for comparison of model simulations results.

115

116   **Fig. 2. Here**

117   In addition, local rainfall data were considered for three historical events. They were selected from

118   five rain gauges, installed since March 2016. The most severe event, on 30 May 2016, had rainfall

119   return periods in the range that is typical for sewer design: from about 2 years to about 10 years,

120   depending on the location. The other two selected events, on 15 June 2016, and on 22 May 2016,

121   have rainfall return periods less than 2 years. The event on 22 May 2016 was selected for evaluating

122   the urban inundation results as data on that that were provided by the city authorities regarding

123   locations of reported flooding. The model rainfall input was based on 5-minutes C-band radar data

124     with a resolution of 539m from the Royal Meteorological Institute of Belgium (Goudenhoofdt and

125     Delobbe, 2016) merged with the 1-minute rainfall intensity observations from the five rain gauges

126     using a simple mean field bias correction.

127     Water level observations in pumping chambers, routinely recorded for monitoring purposes, were

128     considered for the pumping locations at Bredestraat (P1) and Koutergoedstraat (P2), of which the

129     locations are shown in Fig. 1. They have a combined upstream drainage area of 416 and 16 ha

130     respectively.

131     **2.4    1D-0D model**

132     A 1D-0D full hydrodynamic model of the sewer network was implemented in Info Works ICM

133     (Innovyze, 2014) based on detailed physical information on the location of the sewer network, the

134     geometrical and material properties of the sewer pipes and the location and regulation of pumps and

135     structures. The model, governed by the full St Venant equations, is composed of approximately 7400

136     nodes, 4500 subcatchments (38% impervious areas) and 13700 pipes spread over a total length of

137     290km. Around 84% of the system was configured as a combined system, 15% was storm type and the

138     rest conveyed sanitary water type.

139     In this 1D-0D approach, flood storage above each manhole was initially defined in the form of a double

140     flood cone. The first flood compartment starts from the manhole to a height of 1m above the ground

141     level with top cross sectional area of 10% of sub catchment area. The second compartment follows to

142     a height of 5m above the ground level with top cross sectional area of 100% of the sub catchment

143     area. The heights of the cones can be modified to better represent the volume area relationship based

144     on the DTM for each potentially flooded zone. While such adjustments can be made for the cone

145     dimensions, instabilities can be introduced; for instance, for flat cones, a small height difference can

146     change the flood volume significantly. Besides, flood cones do not allow for routing of surface water

147     because connectivity between cones is nonexistent. Because of such limitations, flood volumes and

148 depths simulated by the flood cone approach are often biased. For this reason, flood cones are used

149 in this study for identifying flood prone locations and for comparison the results of the 1D-2D

150 approach.

151 The flood prone locations were identified by simulating the composite storms for the different return

152 periods in the 1D-0D model. This was done through analysis of the flooded nodes and their

153 corresponding flood volumes. After comparing with historically flooded streets, a threshold of 10 m$^3$

154 was selected as the volume in the flood cone above which flooding typically starts. Flood sensitive

155 areas were then selected whose 20-year storm flood peak volume exceeds the flooding threshold (Fig.

156 3). The aim was to generate a primary map of the high flood hazard areas along the sewer system

157 network. In tandem with land use information on the location of the buildings, these areas were

158 classified as high risk zones.

159 **Fig. 3. Here**

160 **2.5    1D-2D model**

161 After the identification of the high flood risk areas, the 1D-2D model was setup. The link between the

162 1D conduits and the 2D surface was through 2D manholes. A 2D manhole is represented as a weir with

163 a crest level taken as the ground level and a crest length equal to the shaft circumference of the node.

164 A discharge coefficient of 0.5 was applied to all manholes of type 2D. In this way, water is exchanged

165 to the 2D surface when the pressure head at the manhole exceeds the ground level.

166 After surcharge flooding, the 2D surface model is activated. The 2D calculations are based on a

167 procedure described by Alcrudo and Mulet (2005). The shallow water equations (SWE) for the depth

168 averaged version of the Navier-Stokes equation are used for the mathematical representation of the

169 2D flow. Roughness zones are defined to represent varying roughness for streets, green zones and

170 other surfaces. The 2D surface is represented as a network of unstructured grids that allow for a better

171 flow path connectivity especially around builds which are considered as void regions. However, the

172  generation of the 2D grid is not trivial. It is important that the generated mesh is hydraulically

173  appropriate with acceptable calculation times and accuracy. A very fine grid discretization may

174  theoretically be more accurate but calculation times are burdensome. It is not clear how mesh sizes

175  can influence the flooding extent as the flow paths can change based on site conditions. Thus based

176  on a sensitivity approach, various mesh resolutions were tested with the purpose of establishing a

177  compromise between resolution and calculation time.

### 2.5.1  Meshing for 2D surface flood modelling

179  As recommended by Innovyze (2014), a minimum and maximum element size ratio of 1:4 was applied

180  in all areas in order to ensure good consistency of the mesh. Next, in areas with large height variation,

181  terrain sensitive meshing was implemented to increase the resolution. This consisted in a more

182  detailed triangle mesh creation in areas where the terrain changes the most. To study the model

183  stability that is linked to the mesh resolution and the calculation time step, the mass balance error

184  was considered. This error was computed by subtracting from the net runoff inflow in the 1D sewer

185  network, the combined surface runoff and the losses that include evaporation and infiltration. For

186  accuracy, the mass balance error (%) should be close to 0.

187  The aim of the meshing analysis was to set lower mesh resolutions for low flood risk areas, which do

188  not need as much detail as the other areas, in order to increase the simulation speed without

189  adversely affecting the accuracy of the results. Three approaches were tested: (1) 2D mesh without

190  mesh zone delineation; (2) 2D mesh with mesh zone delineation for streets and flood risk zones; (3)

191  2D mesh with mesh zone delineation of streets and infiltration 2D zones (Fig. 4 a, b, c). For the low

192  detail 2D mesh approach (Figure 5a), there was no infiltration considered thus assuming the whole

193  surface area as impervious with buildings used as no mesh zones (voids). In the second approach (Fig.

194  4b), preference was given to the streets with detailed mesh zones of $3.75m^2$ - $15m^2$ element size and

195  flood risk areas with $12.5m^2$ - $50m^2$ resolution and varying mesh over the remaining 2D surface, which

196  was kept to a coarser resolution. This street range was chosen based on the typical maximum street

197   width within the study area. As in the first approach, no infiltration was considered here. The third

198   approach (Fig. 4c) was similar to the second but with the addition of infiltration for the mesh elements

199   covering green areas. These could overlap any portion of the surface with the exception of streets,

200   considered fully impervious, and building polygons that were considered as no flow zones. For each

201   of the three approaches, eight mesh resolutions were implemented for the low flood risk areas

202   ($125m^2$-$500$ $m^2$; $100m^2$-$400m^2$; $75m^2$-$300m^2$; $50m^2$-$200m^2$; $37.5m^2$-$150m^2$; $25m^2$-$100m^2$; $18.75m^2$-

203   $75m^2$ and $12.5m^2$-$50m^2$) and further analyzed. For approaches 2 and 3, mesh resolutions were fixed

204   for the high flood risk and street areas to preserve accuracy of flood results for these areas. The aim

205   of the meshing analysis was to set lower mesh resolutions for the low flood risk areas, which do not

206   need as much detail as the other areas, in order to increase the simulation speed without adversely

207   affecting the accuracy of the results.

208

209   **Fig. 4. Here**

210   **2.5.2    Infiltration zones**

211   Infiltration zones were defined for regions where infiltration effects are significant such as green

212   zones. It is worth noting that infiltration is implicitly applied in the runoff generation process for the

213   urban runoff subcatchments albeit not explicitly as the infiltration on the 2D surface. As the runoff

214   volume generation is based on a lumped approach per subcatchment (see section 2.7), infiltration is

215   considered as part of the losses implicit in the runoff coefficients applied for the subcatchments.

216   However, a 2D infiltration model is implemented when excess volumes are released from the 1D

217   sewer nodes onto the 2D mesh after which infiltration is estimated for the mesh elements covering

218   the green zones. In other words, infiltration during the runoff process is based on a lumped approach

219   while infiltration during flooding is based on a distributed approach.

220   Infiltration along the 2D mesh is modelled using a widely applied variant of the Horton infiltration

221   equation (Akan, 1992). In this variant, the actual amount of water infiltrated from the start of the

222   infiltration is incorporated in the equation, which implies that the infiltration at a given time is not just

223   a variable with time as is assumed in the original equation. Moreover, it allows for infiltration to occur

224   without rainfall. A sensitivity analysis was carried out where the infiltration conditions were tested for

225   three soil groups (US Soil Conservation Service (SCS) groups A, B and C) to study the impact that

226   infiltration can have on surface flood modelling results.


227   **2.6   Runoff generation model**


228   Prior to both types of hydraulic urban inundation models considered (1D-0D and 1D-2D), a runoff

229   generation model transforms rainfall intensities in runoff volumes that enter the sewer manholes. This

230   model first accounts for the initial losses from small depressions on the land surface by means of a

231   depression storage model. The depression storage capacity is estimated using subcatchment slope,

232   surface perviousness and porosity (Innovyze, 2014).


233   After removing the depression storage from the rainfall, a percentage of the net rainfall was

234   transformed to runoff. Runoff volumes from subcatchments were generated using runoff coefficients

235   for different land use surfaces. Four types of surfaces were assumed to represent the overall land use

236   types within each subcatchment unit. They consisted of paved surfaces and roof areas, unpaved

237   surfaces with negligible runoff and paved surfaces with minimal runoff. Suggested first guess

238   coefficients for impermeable surfaces are 80% for roads and roofs and between 0% and 10% for

239   permeable areas (Innovyze, 2014). For this study, permeable surfaces were assigned a runoff

240   coefficient of 8%. However, runoff coefficients for the impermeable surfaces were calibrated based

241   on collected rainfall and runoff-related data.


242   Following the Wallingford routing model (Innovyze, 2014), the resulting runoff volume is routed to the

243   sewer manholes from the centroid of each sub catchment. This volume is the total runoff after

244     accumulating the runoff volumes from the pervious and impervious surfaces. The time delay is

245     accounted for by a routing model based on two linear reservoirs in series. The reservoir coefficients

246     are derived from catchment characteristics of slope, area and rainfall intensity. All catchment runoff

247     is transferred to the sewer manholes. As there was not sufficient gully inlet data, this approach was

248     adopted. As such, the flooding mechanism is mainly surcharge based. Surface ponding generated from

249     the limited capacity of the inlets is not modelled, albeit the latter may have an important effect (Pina

250     et al., 2016).

251     **3     Results and discussions**

252     **3.1     Sensitivity analysis**

253     Fig. 5 shows the obtained mesh sensitivity results. Approach 2 generally shows the largest flooding

254     volumes and extents. Approach 1 also exhibits large flood areas although lower when compared with

255     approach 2. The differences between approaches 1 and 2 are attributed to the discretization of the

256     streets and flood risk zones. As approach 1 does not apply a specific meshing zone for these areas, the

257     resulting mesh configuration discretizes less flow pathways. When using meshing zones for approach

258     2, the mesh elements are higher in number and usually decrease in size during meshing as algorithm

259     uses smaller elements to accurately map the polygon boundary. In addition, as streets are also meshed

260     at a finer resolution, the number of potential flow path ways captured increases. Moreover, the

261     streets contain nodes where surcharged excess volumes are discharged. In flat areas, high resolution

262     street discretization (more flow paths) leads to higher flood extents due to more spreading. At a

263     coarser resolution, there is a loss of flow paths, which means that the diffusion of water is limited.

264     However, in Fig. 5, the changes of the flood volume with mesh resolution for approach 1 and 2 do not

265     follow a clear pattern. For instance, while volumes converge at coarser resolutions for the 5-year

266     return period, the volumes diverge for a 20-year return period for coarser resolutions. For the 2- and

267     10-year return periods, the relationship between mesh size and flood volume does not follow a clear

268     pattern. As volumes are aggregated for the entire region, it is not straightforward to determine the

269    reasons for the variation of volumes with mesh sizes. For a more in-depth analysis, three regions were

270    investigated for maximum flood volume differences. The first region (region a) is characterized by the

271    distribution of flood volumes along the straight street with a slope of 2.38%. The second region (area

272    b) is characterized with flood volumes distributed in a branched street network with average slope of

273    1.92%. In the third region (area c), the flood volume is distributed in a cross shaped street network

274    with an average slope of 2.05%. For all the areas, the proximity of the buildings to the streets restricts

275    the flood extent along the streets until an opening is created for another street. This is especially

276    important for high return periods where new paths potentially extend the flood volumes. A

277    comparison was made between the highest ($12.5m^2$-$50m^2$) and lowest ($125m^2$-$500\ m^2$) resolutions for

278    both approach 1 (mesh) and approach 2 (mesh zones). The ratio of flood volumes for approach 2 to

279    approach 1 was used to quantify the differences. Results in Fig. 6 show that for return periods less

280    than 5 years, the steeper areas (regions a and c) show higher differences for high resolutions. For the

281    higher return periods the differences are again higher for high resolution settings compared to coarse

282    resolution settings. This is especially apparent for regions with more complex street geometry (b and

283    c). It is notable that region c has more pronounced differences between approach 1 and approach 2.

284    This is observed for all return periods indicating that the dynamics of flooding involve more complex

285    flow pathways that are better represented using higher resolutions. The inherent assumption is that

286    higher resolutions are better suited for discretizing flow paths in complex geometry. The importance

287    of defining higher resolutions for high flood risk areas is demonstrated based on a comparison of the

288    high resolution settings for approach 1 and 2. Since approach 1 does not involve a specific delineation

289    of flood prone areas, it is likely that some of these flood prone regions are not discretized using finer

290    resolutions which leads to lower flood volumes compared to approach 1. This leads to

291    underestimation of flood volumes. In other words, approach 1 is influenced by topography while

292    approach 2 is influenced by the sewer system drainage network as well as the topography.

293    Approach 3, which considers infiltration, shows the lowest flood volumes and extents. It is apparent

294    that infiltration reduces calculation times as less mesh elements are wet. Due to the non-linear sewer

13

295   system response to rainfall, the reduction in volumes is more important for the lower return periods.

296   Comparing approach 2 and 3, it is noticeable that simulation times are reduced when infiltration is

297   incorporated. This is explained by the fewer calculations for the 2D surface inundation model due to

298   the presence of fewer wet mesh elements. Thus, the inclusion of infiltration will not only make the

299   simulation more representative; it also reduces the calculation time. Approach 1 shows the lowest

300   calculation times as the number of 2D elements are lower and it does not include infiltration

301   calculations. Thus, approach 1 would be useful as a first attempt for identifying flooded areas. Results

302   shown in Fig. 7 indicate that infiltration has an important impact on flood volumes and extents for the

303   study area. The highest impact of infiltration was found for the 2-year return period event for the

304   flood volumes and extents.

305

306   **Fig. 5. Here**

307   **Fig. 6. Here**

308   **Fig. 7. Here**

309

310

311   **3.2    Identification of optimal model resolution and configuration**

312   With the aim to develop an optimal model setup, a reasonable compromise between simulation time

313   and accuracy was established. Mesh sizes in the range $75m^2 - 300m^2$ (average size of around $80m^2$ in

314   Fig. 5 from approach 3 results) were chosen for low flood risk areas. Considering approach 3, this range

315   of mesh sizes, is optimal considering calculation times as other factors such as flood volumes tend to

316   be similar for other mesh size ranges. Compared to the highest resolution range ($12.5m^2 - 50m^2$), the

317   overall flood extents using the selected range would be underestimated by 7% for the 20 year return

318   periods.

319    Hence, the final 2D surface model consists of mesh zones with sizes in the range of $3.75m^2 – 15m^2$ for

320    the streets, $12.5m^2 – 50m^2$ for the high flood risk areas, and $75m^2 – 300m^2$ for the low risk areas. The

321    final model also includes the Horton infiltration model for permeable 2D areas. It is worth pointing

322    out the importance of the time step used for linking the 1D hydraulic model with the 2D surface model.

323    The timestep for 1D-2D linking is not the same timestep for stability in either 1D or 2D systems. The

324    timestep for each of the systems is adjusted for accuracy and efficiency during simulation based on

325    stability criteria. The timestep for linking is applied as an initial estimate of the calculations for the 1D

326    system after which tolerance criteria are used to estimate appropriate timesteps. But the 2D engine

327    timestep for stability is calculated based on the Courant-Friedrichs-Lewy condition (Innovyze, 2014)

328    related to the 2D element mesh size. It is recommended to select a linking timestep for which the

329    mass balance errors are low. A linking timestep of 5 seconds was sufficient for keeping the mass

330    balance errors below 5 percent.

331

332    **3.3    Comparing 1D-0D to 1D-2D**

333    Flood characteristics between the 1D-0D sewer model with flood cones and the 1D-2D model were

334    compared. Table 1 shows the magnitude factor difference between the peak volumes, at the time

335    instant of the maximum flood extent, for the two approaches. The factor is based on peaks whose

336    time of occurrence is different for the two models. For comparison, three locations (Fig. 8) were

337    selected. The flood volumes for the flood cone approach were calculated from the cones within the

338    flooded street sections. It is apparent that the 1D-0D approach has lower flood volumes. This is most

339    significant for location c. It is discernable from Fig. 8 that the distance to the sewer outfalls for location

340    c is the shortest out of the three locations. This means that for this location the flood cones contain

341    less flood volumes because  the sewer system has higher capacity as a result of being emptied at a

342    faster rate. This is consistent with the 1D-0D approach where the stored flood volume depends on the

343    available underground storage capacity. Infiltration losses during flooding are less significant for

344 locations b and c as green zone areas are not present within the vicinity of the flooded streets. The

345 fact that flood cones are leading to lower flood volumes for all three locations indicates that infiltration

346 losses have a low impact on the surface volumes in the 1D-2D surface model hence leading to higher

347 surface volumes. Additionally, the volume in the overland surface of the 1D-2D model is increased by

348 the presence of surface depressions which store more flood volumes that do not immediately return

349 to the sewer network. However, this does not mean that the 1D-0D model always leads to lower flood

350 volumes. For instance, for steeper catchments the 1D-0D model could have higher flood volumes in

351 the upstream locations. The shape of the flood cones also has an important impact on the stored flood

352 volumes with flatter cones storing more flood volumes. As such, the shape of the flood cones can be

353 adjusted to give a better representation of the flood volumes.

354 **Table 1. Here**

355 **Fig. 8. Here**

356 **3.4    Estimation of runoff coefficients**

357 The urban runoff model requires an estimate of the runoff coefficient based on the types of runoff

358 surfaces. This estimation was combined with the model performance evaluation based on field

359 observations of water levels. Of particular interest is the runoff coefficient from impervious surfaces,

360 which were classified as paved surfaces and roof surfaces. The recommended range for poor to high

361 quality paved roads is 0.8-0.9. Because flow meters were not present, a comparison of observed with

362 simulated water levels was made assuming that range of coefficients. Fig. 9 shows the water levels for

363 two events and two locations. These are water levels in pumping chambers. Because pump operations

364 in the field are unpredictable, hence may differ from the implementation in the model, it is not logical

365 to compare the entire time series of levels. Instead, peak water levels are of interest. The average

366 absolute peak differences were 0.16m and 0.20m for 0.8 and 0.9 respectively. The runoff coefficient

367 was henceforth set to 0.8.

---

368

369 **Fig. 9. Here**

370 **3.5    Evaluation of 2D flooding**

371 To better understand the performance of the model in simulating surface flooding, a model simulation

372 with an extreme event was made. Detailed surface flooding evidence is often unavailable for such

373 evaluation. In this study, flood event measurements on inundation depths and extent were not

374 available. Instead, photographic and video data from different sources were collected. Information

375 regarding flooding locations was provided by the city authorities. It is reasonable to assume that some

376 locations were reported as flooded because the depth was high enough to cause some significant

377 interruptions. Nonetheless, important insights can be gleaned from this type of information regarding

378 the consistency with the flooded streets in the model. They were available for the recent flood event

379 of 30 May 2016. In addition, water level measurements available for the same event at two monitoring

380 locations show that the model results are close to the observations. Fig. 10 shows flooding at

381 Bredestraat location and high but non flooding conditions at Koutergoedstraat. This finding was

382 consistent with reported flooded locations in Fig. 11.

383 **Fig. 10. Here**

384 Furthermore, results are examined for three locations identified as flood prone based on previous

385 composite storm analysis. Fig. 11 shows the simulated flood extents for the areas around these three

386 locations along with the flooded locations as simulated in the models. Reported flood locations appear

387 to mostly match the flooded streets (depth greater than 10 cm) in the model. The northern part of the

388 area experienced the highest flood impact with most of the flooding resulting from overflowing of the

389 channel banks. It is interesting to note the consistency between the evidence for the inundations at

390 the locations (a) and (c). Flooding at location (b) was minor and this could explain why no reported

391 floods were recorded at this location. The extent of the flooding at location (c) is found

392 underestimated by the model (Fig. 11), given the flood evidence in Fig. 12 in images P1 and P2.

17

393 Investigation on the possible explanation for this discrepancy revealed two potential explanations.

394 First, backwater effects resulting from high river water levels at the sewer outfalls were found to have

395 important effects on the street flooding. Although water level measurements were not available for

396 the sewer outfall near location (c), reports by the Flemish river authorities provide evidence that

397 before the event, water levels in rivers were higher than normal.  These normal levels are defined as

398 the long term average water levels. To investigate the influence of these river levels, the level at the

399 outfall near location (c) was raised with 20, 50 and 80 mm. Fig. 13 shows that this increase has an

400 important impact on the urban flooding at location (c). This indicates the importance of a dynamic

401 interaction between the river and sewer systems. By coupling a hydrodynamic model for the river with

402 the sewer model, the effects of backwater interaction would be better simulated (Kandori and

403 Willems, 2008). The other possible explanation considered, is the urban sewer flood model

404 conceptualization which assumes that all runoff volumes are transferred directly to the sewer system

405 without any surface ponding resulting from limited capacity of inlets. Therefore flooding only occurs

406 from sewer surcharge which implies that surface flooding is limited for an area with sufficient

407 underground storage and/or drainage. This is consistent with location (c) where the outfall location is

408 close to the flooded street and the drainage network allows for a larger sewer system storage and

409 downstream throughflow, which reduces the potential for surcharged flooding.

410 For location (a), evaluation of the model result for the flooded street was done based on the

411 timestamp the image (Fig. 12 P3) was taken. Based on visual inspection in this image of the inundation

412 depth and the approximate inundation area, they are found consistent with the average inundation

413 depth of 0.2 m as simulated by the model.

414 For the same event and based on the same flood information, sensitivity analysis was conducted on

415 the soil infiltration rates. These rates were varied using two different soil types representing higher

416 and lower infiltration rates than the assumed soil type. Fig. 14 shows that infiltration has an important

417 influence on the extent of the inundation. The magnitude of the reduction in flood extent hence

418 largely depends on the soil type. Compared to a case of no infiltration, a higher infiltration rate (soil

419 A) leads to a reduction of 12 percent, medium infiltration rate (soil B) leads to a 10 percent reduction

420 and a low infiltration rate (soil C) leads to a reduction of 7 percent in flood extent. This demonstrates

421 the benefits of increased infiltration and the important flood mitigation effect if infiltration is impeded.

422 However, it is hard to conclude which soil type leads to the most accurate flood extent as the

423 information to be obtained from the available image is not that precise. Nonetheless, it is reasonable

424 to assume that soil type B would be representative of the flood extent and volume based on the

425 existing soil maps. The predominant type of soil in the region is sandy loam.

426 **Fig. 11. Here**

427 **Fig. 12. Here**

428 **Fig. 13. Here**

429 **Fig. 14. Here**

430 **4    Conclusion**

431 This study has explored the process of developing a 1D-2D model for urban flood modelling. Through

432 a three step process, an optimal model setup has been determined. The first step involved the use of

433 1D-0D model setup to identify flood prone zones that are consequently given higher resolutions in the

434 second step. The second step was concerned with the sensitivity analysis of 2D meshing which is a

435 major concern for 2D models especially regarding model run time. A fine resolution mesh offers

436 benefits of accuracy at the expense of calculation time while simulation times are less prohibitive with

437 coarse resolutions but with a loss of accuracy. Sensitivity analysis allows for establishing optimal model

438 settings. The third step focused on validation and evaluation of the 1D-2D model.

439 Using synthetic design composite storms, different mesh size resolutions were tested. An optimal

440 resolution with mesh sizes in the range of $75m^2$—$300m^2$ was selected considering mass balance errors

441 (below 5%), flood extent and volume variability as well as calculation time. Infiltration areas were

442 included using a Horton infiltration model, because sensitivity analysis showed that infiltration has a

443    large impact on the flood characteristics. Moreover, it was found that flood extents increase with finer

444    mesh sizes. However, this finding cannot be generalized for other study areas since flow pathways

445    may change after meshing based on site conditions such as topography and street geometry. For this

446    study area, which is characterized by flat topography, coarse mesh resolutions led to a loss of flow

447    paths during averaging; finer mesh sizes were more effective in capturing the diverging flow paths

448    which tend to occur in flat topography. Furthermore, street boundaries defined by building voids,

449    confine the flood extent that would otherwise be higher for coarser mesh sizes. Furthermore, model

450    runs with delineated street and flood risk zones showed higher flood extents compared to model runs

451    without delineated mesh zones. This was consistent with increased flow pathways in the delineated

452    zones. The reduction of flood extent and flood volumes with infiltration was more pronounced for low

453    return periods indicating the importance of infiltration for low intensity events.

454    With the identified optimal resolution, flood 1D-2D flood characteristics were compared with the 1D-

455    0D model based on flood cones. This comparison helped to better understand the effect of model

456    conceptualization on flood volumes. By analyzing three areas, it was established that the flood cone

457    approach leads to lower flood volumes. This was more apparent at locations close to outfalls where

458    the differences are more pronounced. At these locations, the downstream throughflow to receiving

459    rivers allows for a larger sewer system storage capacity which reduces the flood volumes.

460    Due to the lack of extensive measurements, literature based parameters as suggested by Innovyze

461    (2014) were implemented for some model parameters such as roughness coefficients. However,

462    coefficients for the fixed runoff model were verified using recorded water levels at two pump well

463    locations using rainfall events with similar patterns to water level measurements. A coefficient of 0.8

464    was found to be reasonable.

465    From photographic evidence, water depths and extents were investigated for one extreme event.

466    Because both rain gauge and radar data were available for this event, and at high temporal and spatial

467    resolutions, it was an important event to validate the flood modelling results. The model showed

468     consistent results with the flooding locations reported by the regional authorities. However, results

469     from three regions revealed the importance of the dynamic interaction between the sewer network

470     and the receiving river system. Higher river water levels were found to explain the flood extent in the

471     region close to an outfall to the river. Higher river levels were reported by the water authorities but

472     the model simulation assumed normal river boundary conditions at the outfalls. For locations less

473     affected by the downstream river conditions, the model performance in terms of flood characteristics

474     was found to be consistent with the observations.

475     It is clear that the increased use of 1D-2D models will be fostered by the availability of data that is

476     often limited. This study relied on different data sources from meteorological services, water

477     authorities, social media and monitoring campaigns. Data availability at high temporal and spatial

478     resolution made it possible to identify improvements in the dynamics of the model. The consistency

479     between the 1D-2D approach and observations increased confidence in the use of such models for

480     urban flood management.

487 **References**

488 AGIV, 2015. Digitaal Hoogtemodel Vlaanderen II (2013-2015).

489 <https://www.agiv.be/producten/digitaal-hoogtemodel-vlaanderen/meer-over-dhm-

490 v/producten-onderdelen/digitaal-hoogte-model-vlaanderen-ii >

491 Akan, A.O., 1992. Horton infiltration equation revisited. Journal of Irrigation and Drainage Engineering,

492 118(5): 828-830.

493 Alcrudo, F., Mulet, J., 2005. Urban inundation models based on the shallow water equations.

494 Numerical and practical issues. Finite Volumes for Complex Applications IV: Problems and

495 Perspectives, Hermes Science Publishing, pp. 3-12

496 Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen,

497 H., Nguyen, V-T-V., 2013. Impacts of climate change on rainfall extremes and urban drainage

498 systems: a review. Water Science and Technology, 68(1): 16-28.

499 Berne, A., Delrieu, G., Creutin, J.-D., Obled, C., 2004. Temporal and spatial resolution of rainfall

500 measurements required for urban hydrology. Journal of Hydrology, 299(3): 166-179.

501 Bruni, G., Reinoso, R., van de Giesen, N.C., Clemens, F.H.L.R., ten Veldhuis, J.A.E., 2015. On the

502 sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution.

503 Hydrology and Earth System Sciences, 19(2): 691-709.

504 Casulli, V., Stelling, G.S., 2013. A semi-implicit numerical model for urban drainage systems.

505 International Journal for Numerical Methods in Fluids, 73(6): 600-614.

506 Chen, Y., Zhou, H., Zhang, H., Du, G., Zhou, J., 2015. Urban flood risk warning under rapid urbanization.

507 Environmental Research, 139: 3-10.

508 Crompton, R.P., McAneney, K.J., 2008. Normalised Australian insured losses from meteorological

509 hazards: 1967–2006. Environmental Science & Policy, 11(5): 371-378.

510 Dawson, R., Speight, L., Hall, J.W., Djordjevic, S., Savic, D., Leandro, J., 2008. Attribution of flood risk

511 in urban areas. Journal of Hydroinformatics, 10(4): 275-288.

512     Djordjević, S., Prodanović, D., Maksimović, Č., Ivetić, M., Savić, D., 2005. SIPSON–Simulation of

513                 Interaction between Pipe flow and Surface Overland flow in Networks. Water Science and

514                 Technology, 52(5): 275-283.

515     Ghimire, B., Chen, A.S., Guidolin, M., Keedwell, E.C., Djordjević, S., Savić, D.A., 2013. Formulation of a

516                 fast 2D urban pluvial flood model using a cellular automata approach. Journal of

517                 Hydroinformatics, 15(3): 676-686.

518     Ghostine, R., Hoteit, I., Vazquez, J., Terfous, A., Ghenaim, A., Mose, R., 2014. Comparison between a

519                 coupled 1D-2D model and a fully 2D model for supercritical flow simulation in crossroads.

520                 Journal of Hydraulic Research, 53(2): 274-281.

521     Glenis, V., McGough, A., Kutija, V., Kilsby, C., Woodman, S., 2013. Flood modelling for cities using Cloud

522                 Computing. Journal of Cloud Computing, 2(1): 1-14.

523     Goudenhoofdt, E., Delobbe, L., 2016. Generation and Verification of Rainfall Estimates from 10-Yr

524                 Volumetric Weather Radar Measurements. Journal of Hydrometeorology, 17(4): 1223-1242.

525     Haklay, M., Weber, P., 2008. Openstreetmap: User-generated street maps. IEEE Pervasive Computing,

526                 7(4): 12-18.

527     Hunter, N., Bates, P.D., Neelz, S., et al., 2008. Benchmarking 2D hydraulic models for urban flood

528                 simulations. Proceedings of the Institution of Civil Engineers: Water Management, 161 (1). 13-

529                 30.

530     Innovyze, 2014. InfoWorks ICM software v.5.5.4, Monrovia, CA, USA.

531     Kandori, C., Willems, P., 2008. Impact of the two-directional interaction of sewer and river systems on

532                 the flood risk, 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK.

533     Keifer, C.J., Chu, H.H., 1957. Synthetic storm pattern for drainage design. Journal of the Hydraulics

534                 Division, 83(4): 1-25.

535     Kunkel, K.E. et al., 2013. Monitoring and understanding trends in extreme storms: State of knowledge.

536                 Bulletin of the American Meteorological Society, 94(4): 499-514.

537  Leandro, J., Chen, A.S., Djordjević, S., Savić, D.A., 2009. Comparison of 1D/1D and 1D/2D coupled

538      (sewer/surface) hydraulic models for urban flood simulation. Journal of Hydraulic Engineering,

539      135(6): 495.

540  Leandro, J., Schumann, A., Pfister, A., 2016. A step towards considering the spatial heterogeneity of

541      urban key features in urban hydrology flood modelling. Journal of Hydrology, 535: 356-365.

542  Lehmann, J., Coumou, D., Frieler, K., 2015. Increased record-breaking precipitation events under

543      global warming. Climatic Change, 132(4): 501-515.

544  Lhomme, J., Bouvier, C., Perrin, J.-L., 2004. Applying a GIS-based geomorphological routing model in

545      urban catchments. Journal of Hydrology, 299(3): 203-216.

546  Mallakpour, I., Villarini, G., 2015. The changing nature of flooding across the central United States.

547      Nature Climate Change, 5(3): 250-254.

548  Marshall, J.D., 2007. Urban land area and population growth: a new scaling relationship for

549      metropolitan expansion. Urban Studies, 44(10): 1889-1904.

550  Néelz, S., Pender, G., 2013. Benchmarking the latest generation of 2D hydraulic modelling packages.

551      Environment Agency, Horison House, Deanery Road, Bristol, UK.

552  Ntegeka, V., Willems, P., 2008. Trends and multidecadal oscillations in rainfall extremes, based on a

553      more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. Water

554      Resources Research, 44(7).

555  Ochoa-Rodriguez, S. et al., 2015. Impact of spatial and temporal resolution of rainfall inputs on urban

556      hydrodynamic modelling outputs: A multi-catchment investigation. Journal of Hydrology, 531:

557      389-407.

558  Pina, R.D., Ochoa-Rodriguez, S., Simões, N.E., Mijic, A., Marques, S.A., Maksimović, Č., 2016. Semi- vs.

559      fully-distributed urban stormwater models: model set up and comparison with two real case

560      studies. Water, 8(2), 58.

561    Priest, S.J., Parker, D.J., Hurford, A., Walker, J., Evans, K., 2011. Assessing options for the development

562        of surface water flood warning in England and Wales. Journal of Environmental Management,

563        92(12): 3038-3048.

564    Seyoum, S., Vojinovic, Z., Price, R., Weesakul, S., 2012. Coupled 1D and Noninertia 2D Flood Inundation

565        Model for Simulation of Urban Flooding. Journal of Hydraulic Engineering, 138(1): 23-24.

566    Simões, N. et al., 2011. Urban drainage models for flood forecasting: 1D/1D, 1D/2D and hybrid models,

567        Porto Alegre, pp. 11-16.

568    Smith, L.S., Liang, Q., Quinn, P.F., 2015. Towards a hydrodynamic modelling framework appropriate

569        for applications in urban flood assessment and mitigation using heterogeneous computing.

570        Urban Water Journal, 12(1): 67-78.

571    Thorndahl, S., Nielsen, J.E., Rasmussen, M.R., 2014. Bias adjustment and advection interpolation of

572        long-term high resolution radar rainfall series. Journal of Hydrology, 508: 214-226.

573    Willems, P., 2013. Adjustment of extreme rainfall statistics accounting for multidecadal climate

574        oscillations. Journal of Hydrology, 490: 126-133.

575    Willems, P., Molnar, P., Einfalt, T., Arnbjerg-Nielsen, K., Onof, C., Nguyen, V-T-V., Burlando, P., 2012.

576        Rainfall in the urban context: Forecasting, risk and climate change. Atmospheric Research,

577        103(0): 1-3.

578    Wright, N., 2014. Advances in flood modelling helping to reduce flood risk, Proceedings of the

579        Institution of Civil Engineers - Civil Engineering. Thomas Telford Ltd, pp. 52-52.

580    Yu, D., Yin, J., Liu, M., 2016. Validating city-scale surface water flood modelling using crowdsourced

581        data. Environmental Research Letters, 11, 124011.

582

583

584    Table 1. Peak volume ratio (1D-2D vs 1D-0D)

585

586    Table 1. Peak volume ratio (1D-2D vs 1D-0D)

| Location: | Return period [years]: | | | |
|:---:|:---:|:---:|:---:|:---:|
| | 2 | 5 | 10 | 20 |
| a | 1.69 | 1.45 | 1.46 | 1.56 |
| b | 1.92 | 1.69 | 1.73 | 1.71 |
| c | - | 7.82 | 3.54 | 2.99 |

587

588

589    List of Figures

590    Fig. 1. Study area water level monitoring pumping stations P1 and P2 and rain gauge locations at

591    Wittwalle (R1), Eksaarderijweg (R2), Louise Derachestraat (R3), Grondwetlaan (R4), and Adolf

592    Baeyensstraat (R5) (left). Elevation map (right).

593    Fig. 2. Composite storms for return periods of 2, 5, 10 and 20 years. Only 90 minutes centered at the

594    peak of the two day duration storm are shown to emphasize the differences.

595    Fig. 3. Flood risk zones identified based on the composite storm event for return period of 20 years.

596    Fig. 4. Meshing approaches. Approach 1 without street and flood risk mesh zoning. Approach 2 with

597    streets and flood risk mesh zoning. Approach 3 with streets and flood risk mesh zoning including 2D

598    infiltration zoning.

599    Fig. 5. Mesh zone sensitivity analysis for composite storms with return periods of 2, 5, 10 and 20 years.

600    Approach 1 (Mesh), Approach 2 (MeshZones) and Approach 3 (Infil-MeshZones). Simulations were run

601    on a computer with a 64-bit intel core i7-2600 at 3.40GHz and 16 gb of RAM.

602    Fig. 6. Comparing approaches 1 and 2 at locations a,b,c(left). Volume ratio of approach 2 to approach

603    1 for high resolutions (grey) and low resolutions (white).

604    Fig. 7. Flooded area (left) and flooded volume (right) factor difference for model without infiltration

605    (InfilOff) against model with infiltration (Infl).

606    Fig. 8. Flood depth and extent at three locations a,b,c for return period of 20 years.

607    Fig. 9. Simulated and observed water levels at Bredestraat and Koutergoedstraat locations. Pump

608    chamber roof, switch on and switch off levels are shown as grey horizontal lines.

609    Fig. 10. Water levels simulated and observed for 30 May 2016 at Bredestraat and Koutergoedstraat.

610    Pump chamber roof, switch on and switch off levels are shown as grey horizontal lines.

611    Fig. 11. Flood depth and extent at locations a,b,c for 30 May 2016 . P1 , P2 and P3 are view points from

612    which images shown in Fig. 12 were taken. For visibility of flooded streets (right side), a buffer of 10m

613    was added to flooded streets.

614    Fig. 12. Observed flood evidence for two flooded locations a and c for 30 May 2016. Sint-

615    Bernadettestraat street closed to traffic (P2).

616    Fig. 13. Flood depth and extent for normal river levels (a) and normal river levels increased by 20mm

617    (b), 50mm (c) and 80mm (d) at the sewer outfall for 30 May 2016.

618    Fig. 14. Flood extent at flood peak time for soil types A, B and C for 30 May 2016 event at location a.

619
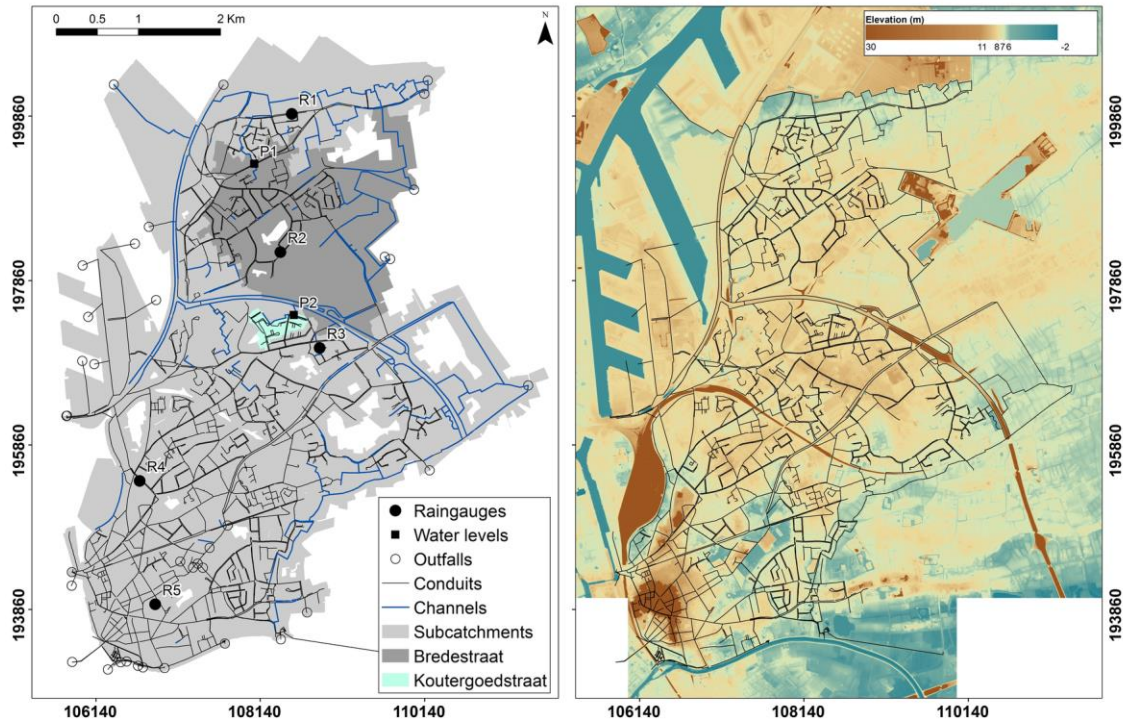
Fig. 1. Study area water level monitoring pumping stations P1 and P2 and rain gauge locations at Wittwalle (R1), Eksaarderijweg (R2), Louise Derachestraat (R3), Grondwetlaan (R4), and Adolf Baeyensstraat (R5) (left). Elevation map (right).
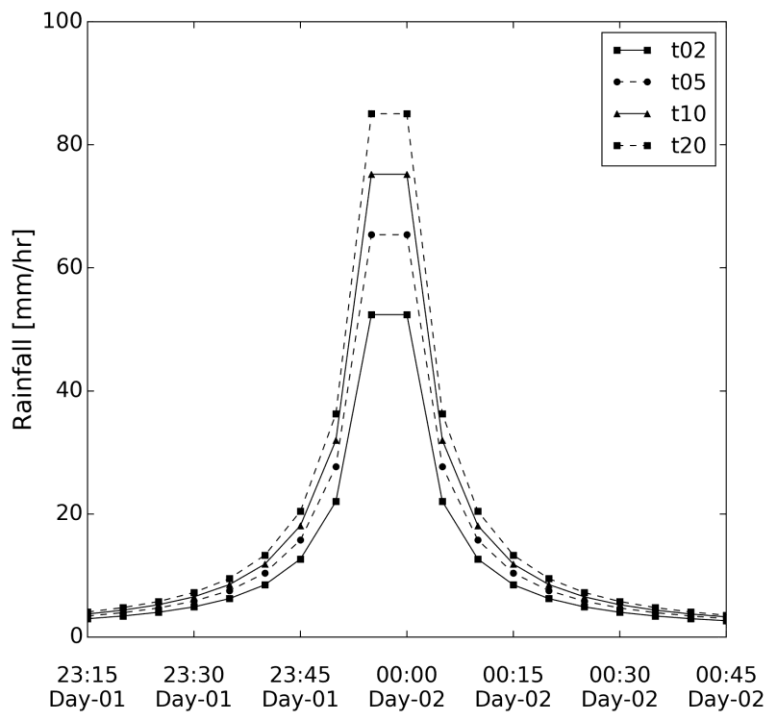
625

626

627   Fig. 2. Composite storms for return periods of 2, 5, 10 and 20 years. Only 90 minutes centered at the

628   peak of the two day duration storm are shown to emphasize the differences.

629
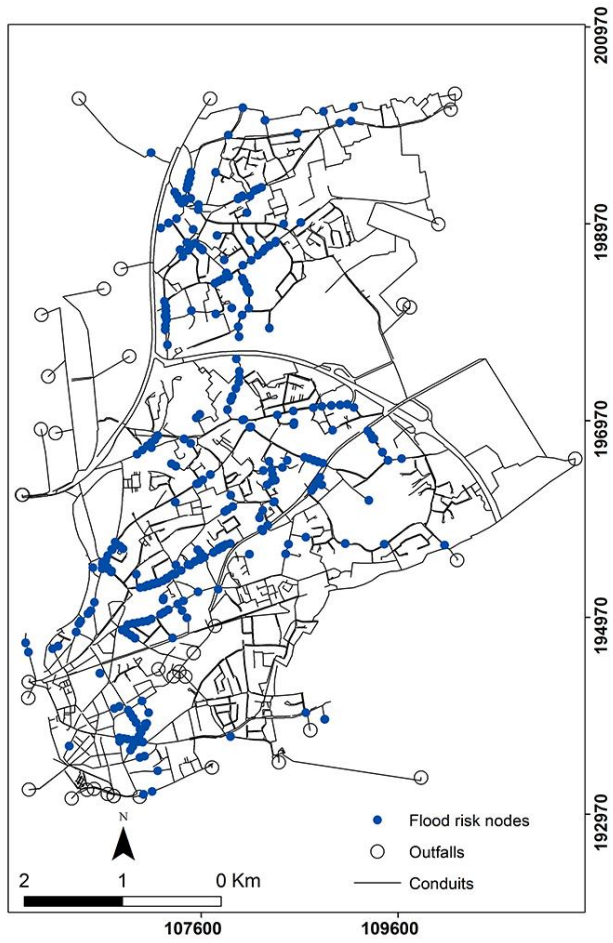
630

631    Fig. 3. Flood risk zones identified based on the composite storm event for return period of 20 years.

632
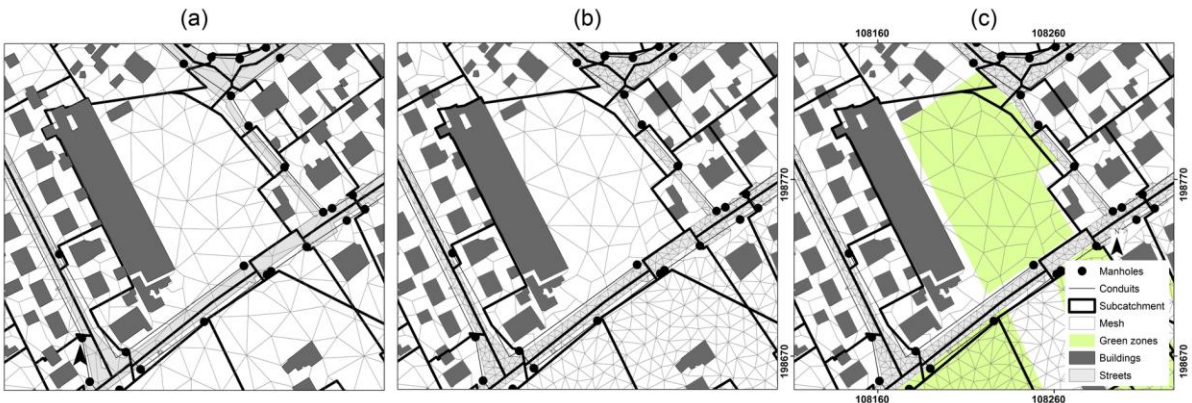
633

634

Fig. 4. Meshing approaches. Approach 1 without street and flood risk mesh zoning. Approach 2 with streets and flood risk mesh zoning. Approach 3 with streets and flood risk mesh zoning including 2D infiltration zoning.

638
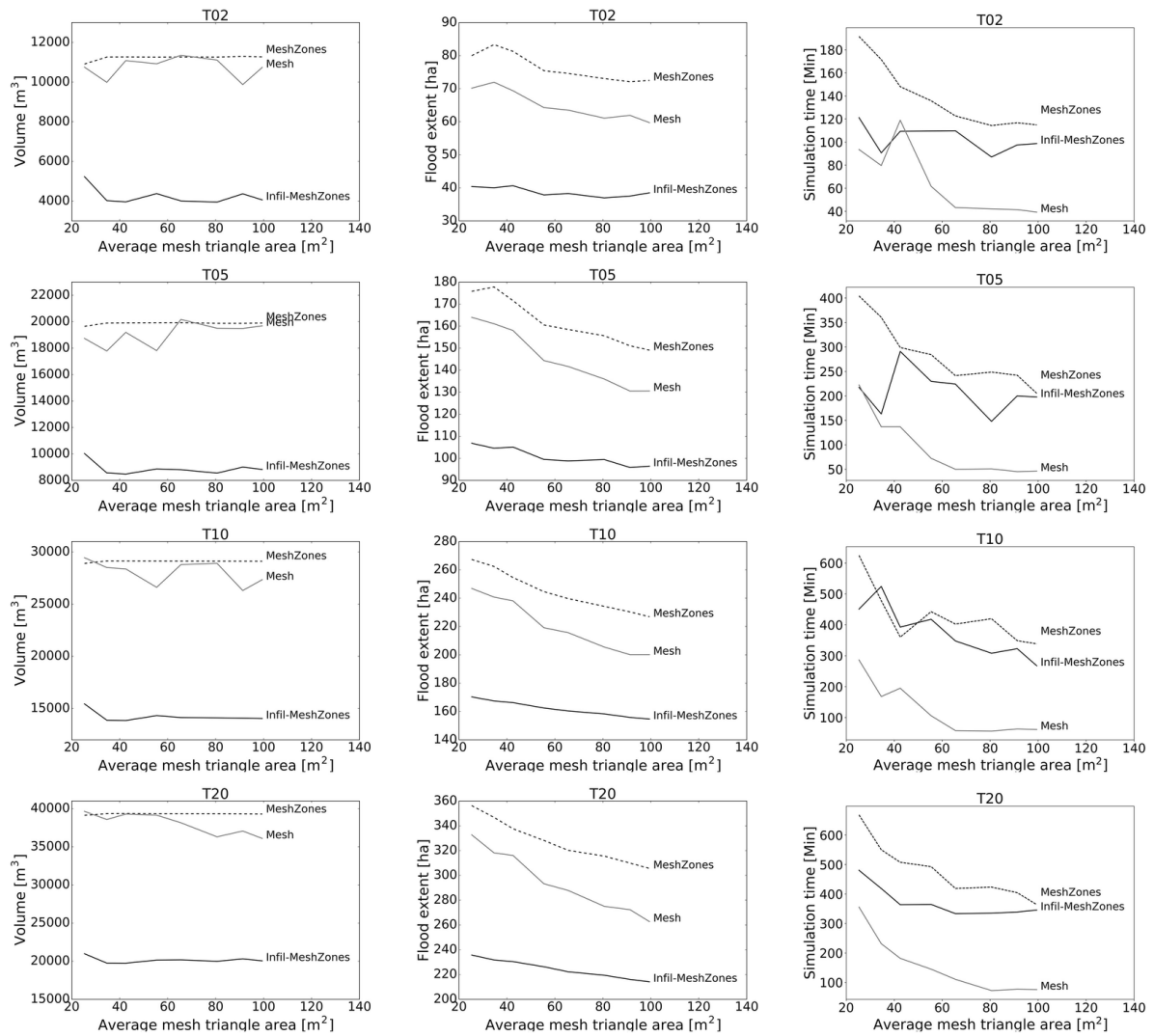
639

640

Fig. 5. Mesh zone sensitivity analysis for composite storms with return periods of 2, 5, 10 and 20 years.

Approach 1 (Mesh), Approach 2 (MeshZones) and Approach 3 (Infil-MeshZones). Simulations were run

on a computer with a 64-bit intel core i7-2600 at 3.40GHz and 16 gb of RAM.

644
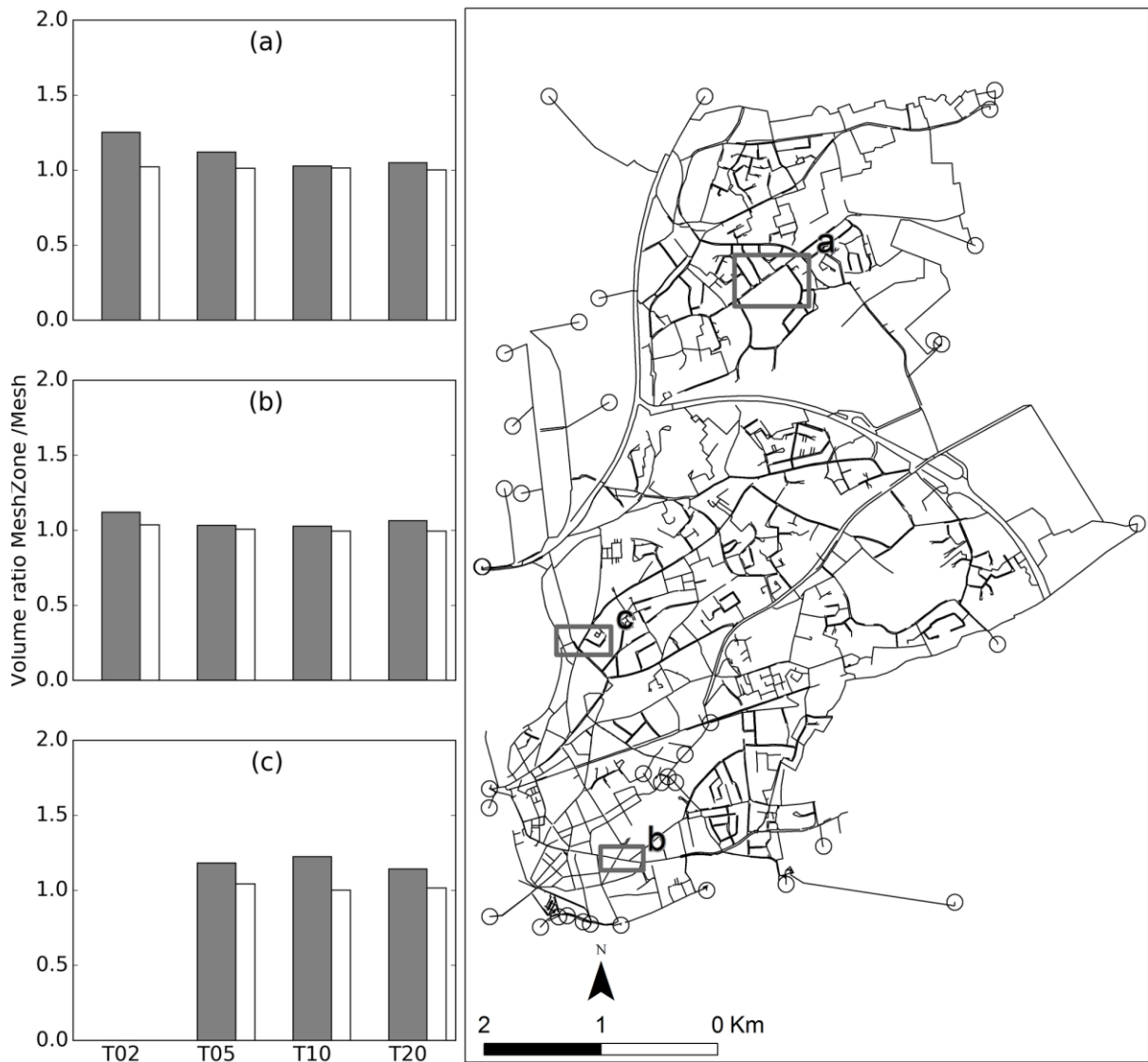
645

646

647 Fig. 6. Comparing approaches 1 and 2 at locations a,b,c(left). Volume ratio of approach 2 to approach

648 1 for high resolutions (grey) and low resolutions (white).

649

650
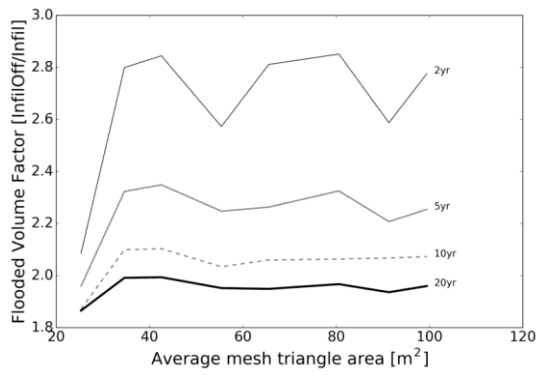


651
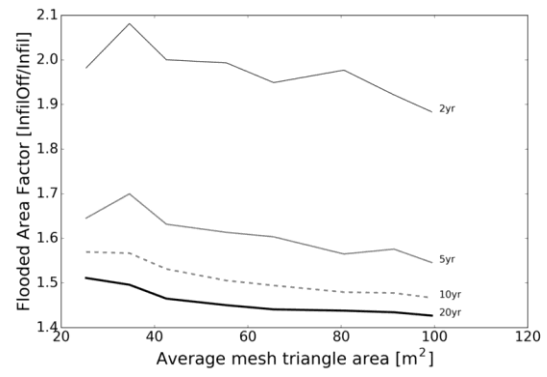
Fig. 7. Flooded area (left) and flooded volume (right) factor difference for model without infiltration

(InfilOff) against model with infiltration (Infl).

654

655
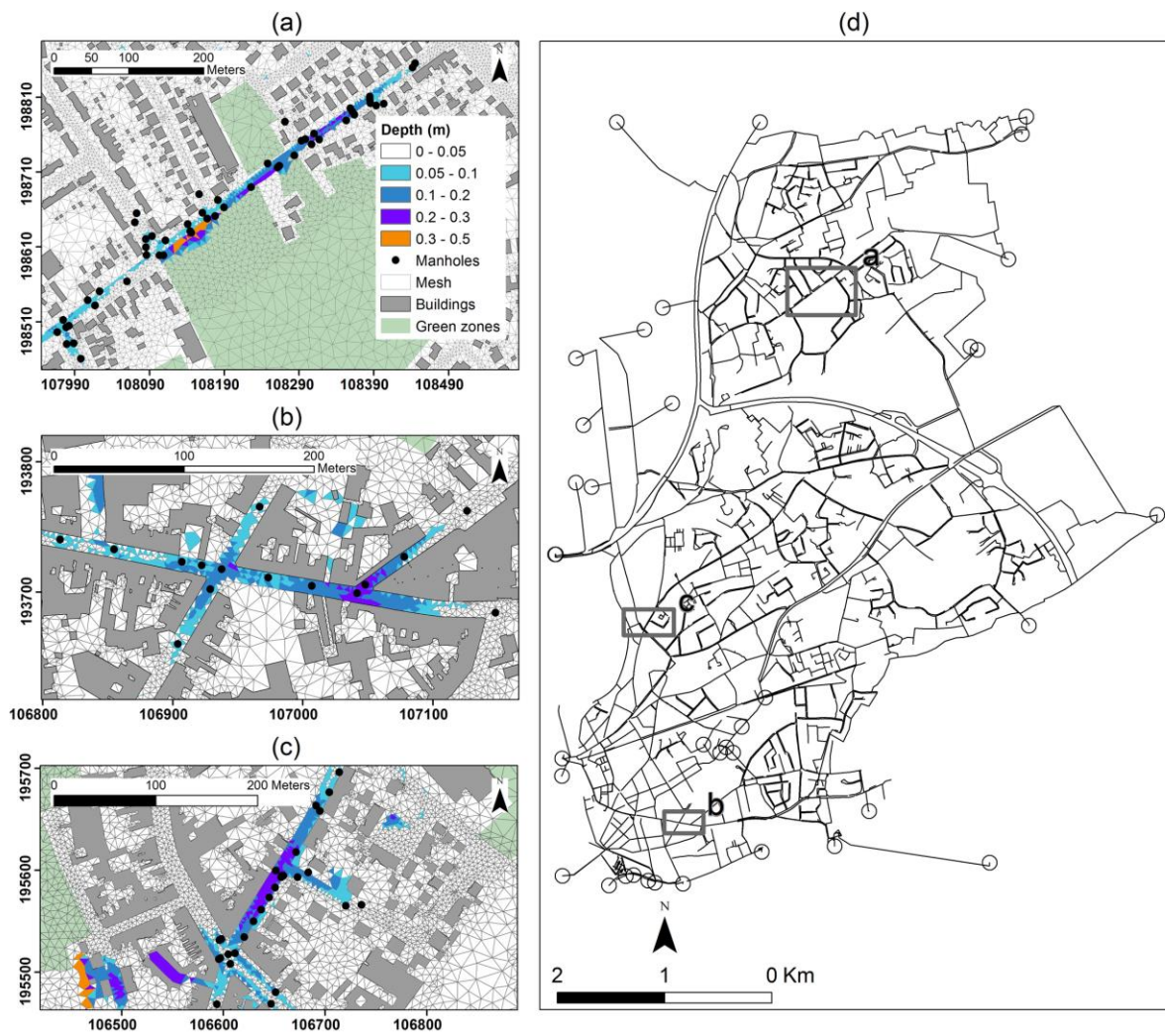


Fig. 8. Flood depth and extent at three locations a,b,c for return period of 20 years.

660
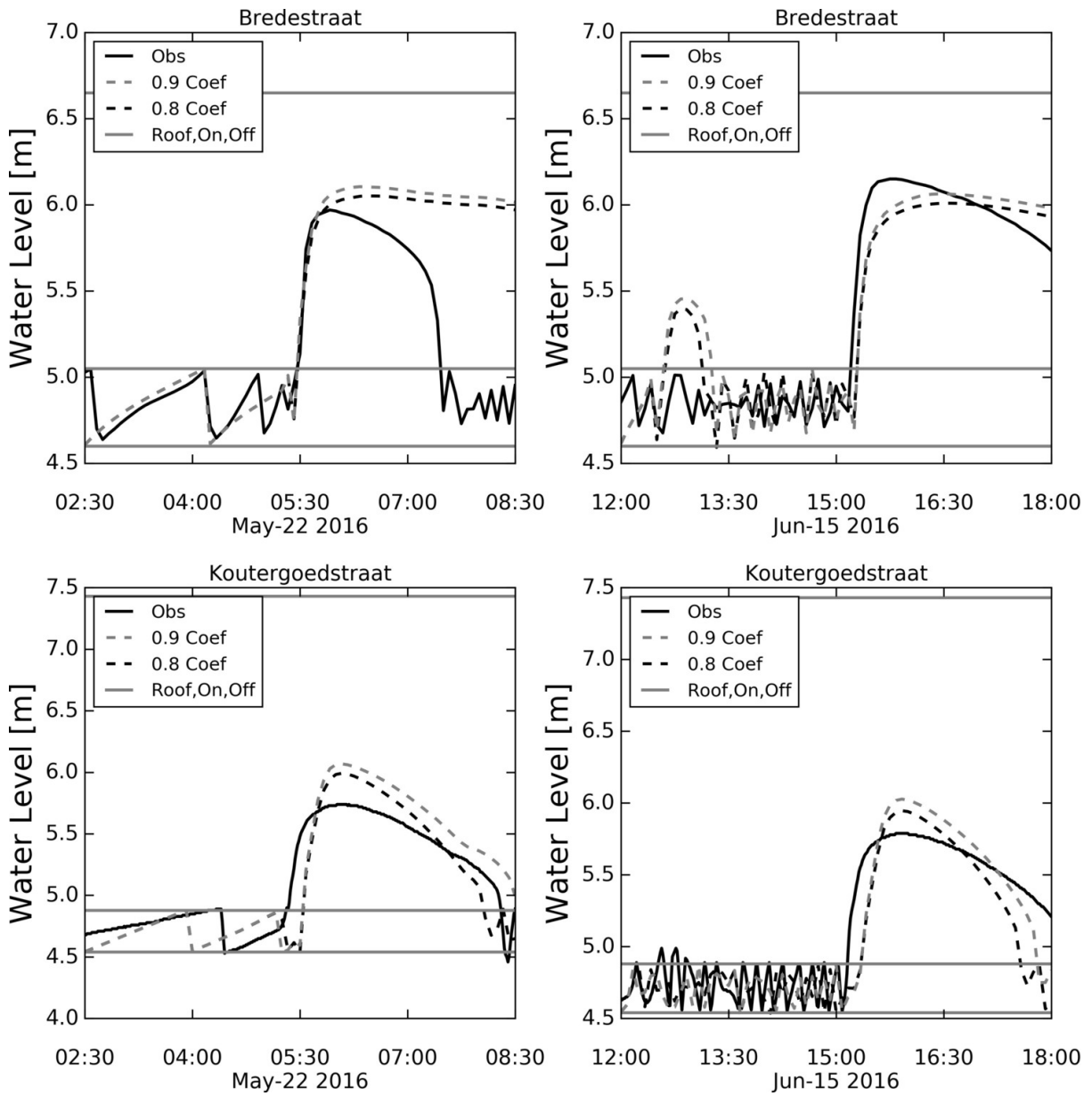


661

662

663    Fig. 9. Simulated and observed water levels at Bredestraat and Koutergoedstraat locations. Pump

664    chamber roof, switch on and switch off levels are shown as grey horizontal lines.

665
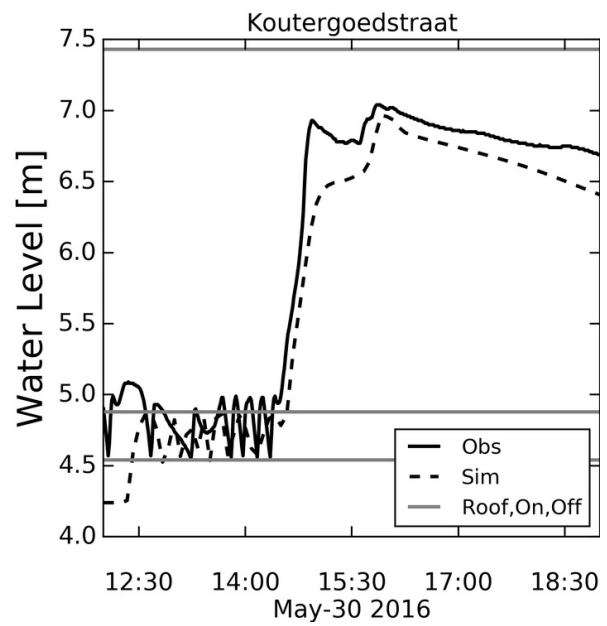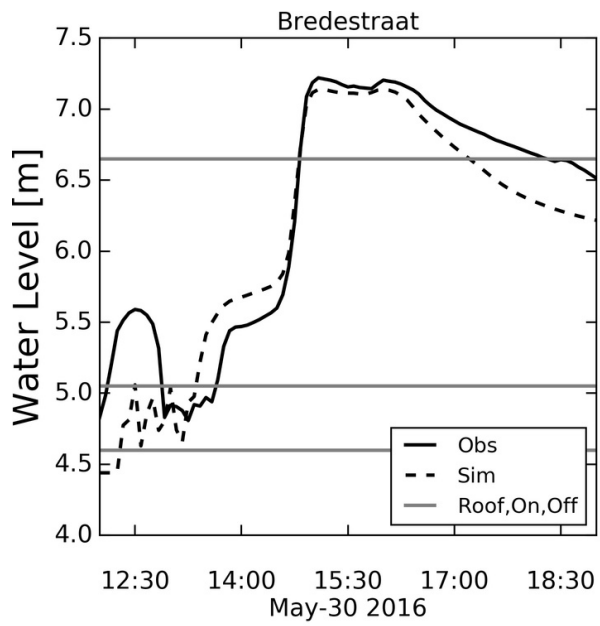
666

667

668 Fig. 10. Water levels simulated and observed for 30 May 2016 at Bredestraat and Koutergoedstraat.

669 Pump chamber roof, switch on and switch off levels are shown as grey horizontal lines.

670
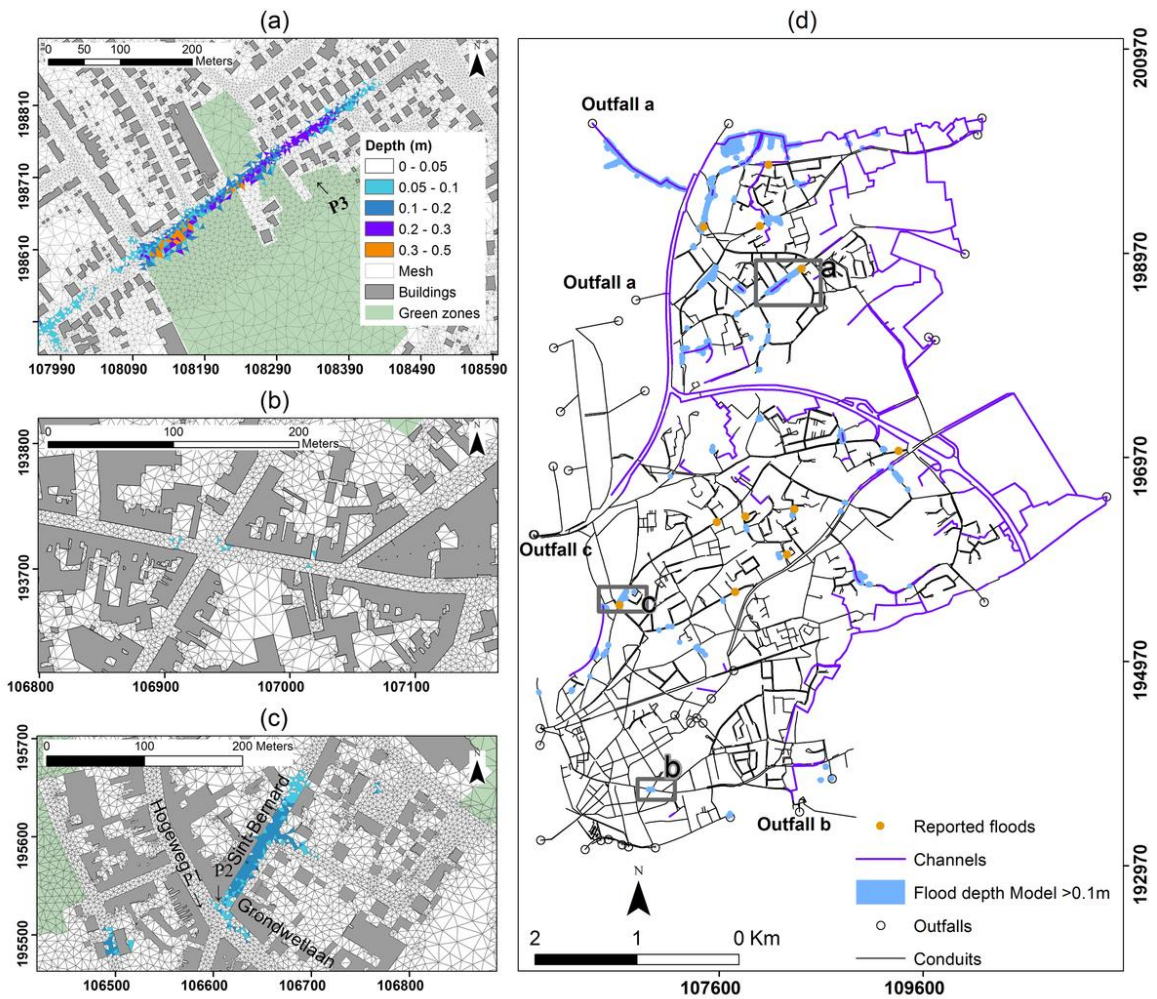
Fig. 11. Flood depth and extent at locations a,b,c for 30 May 2016 . P1 , P2 and P3 are view points from which images shown in Fig. 12 were taken. For visibility of flooded streets (right side), a buffer of 10m was added to flooded streets.

676

677

678    Fig. 12. Observed flood evidence for two flooded locations a and c for 30 May 2016. Sint-

679    Bernadettestraat street closed to traffic (P2).

680
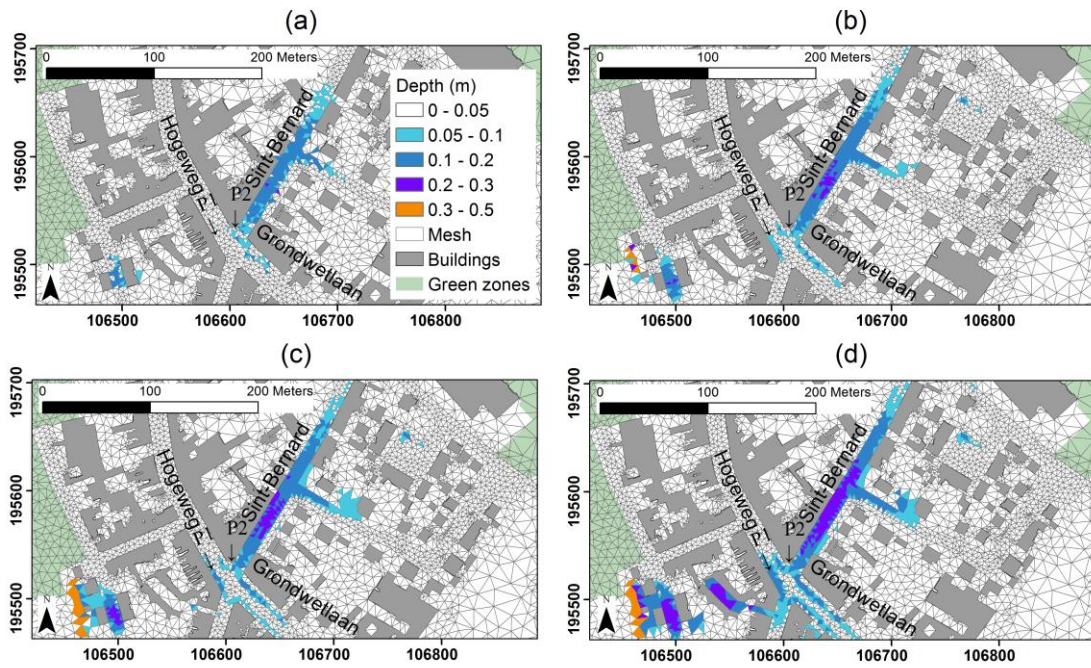
681

682

683 Fig. 13. Flood depth and extent for normal river levels (a) and normal river levels increased by 20mm

684 (b), 50mm (c) and 80mm (d) at the sewer outfall for 30 May 2016.

685
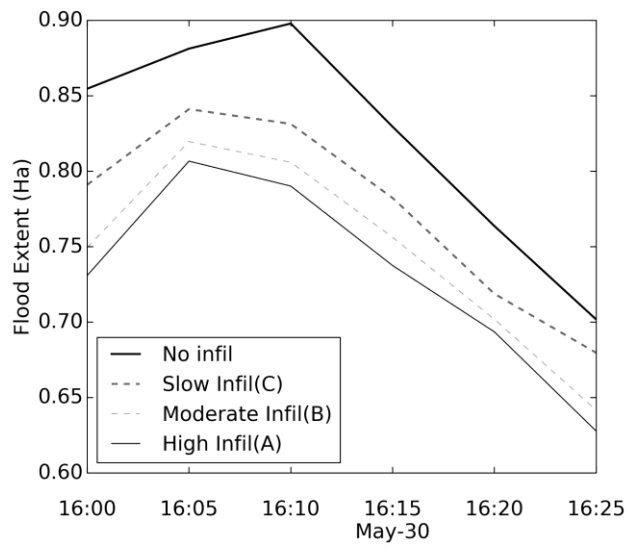
686

687

688　Fig. 14. Flood extent at flood peak time for soil types A, B and C for 30 May 2016 event at location a.

Title: **Development and comparison of two fast surrogate models for urban pluvial flood simulations**

María Bermúdez[1,2], Victor Ntegeka[1], Vincent Wolfs[1], Patrick Willems[1]

1 Hydraulics Section, Department of Civil Engineering, KU Leuven, Belgium
2 Water and Environmental Engineering Group, Universidade da Coruña, Spain

**Summary**

Detailed full hydrodynamic 1D-2D dual drainage models are a well-established approach to simulate urban pluvial floods. However, despite modelling advances and increasing computational power, this approach remains unsuitable for many real time applications. We propose and test two computationally efficient surrogate models. The first approach links a detailed 1D sewer model to a GIS-based overland flood network. For the second approach, we developed a conceptual sewer and flood model using data-driven and physically based structures, and coupled the model to pre-simulated flood maps. The city of Ghent (Belgium) is used as a test case. Both surrogate models can provide comparable results to the original model in terms of peak surface flood volumes and maximum flood extent and depth maps, with a significant reduction in computing time.

## 1. Introduction

Full hydrodynamic 1D-2D dual drainage models are a well-established approach to simulate urban pluvial flooding, which can provide a realistic description of flood conditions. They simulate the coupling between the sewer network, represented by a detailed 1D model, and the surface flow, computed using a 2D hydrodynamic model. This approach is now available in commercial software packages such as SOBEK [*Deltares*, 2017], XP-SWMM 2D [*XP Solutions*, 2014], MIKE FLOOD [*DHI*, 2012] or InfoWorks ICM [*Innovyze*, 2015], and is commonly applied in urban flood studies [*Schmitt et al.*, 2004; *Carr and Smith*, 2006; *Jahanbazi and Egger*, 2014]. However, such set up is generally too computationally intensive for real-time applications such as urban pluvial flood forecasting or for probabilistic approaches that require multiple simulations, even with the aid of high performance computing [*Lacasta et al.*, 2015].

Therefore, there remains a need for developing computationally more efficient surrogates of these models. Two broad families of surrogate models can be considered: lower-fidelity models, which are simplified physically based models preserving the main body of processes modeled in the original system, and data-driven models which emulate the original model responses without analyzing the physical processes involved [*Razavi et al.*, 2012]. In the context of 1D-2D dual drainage models, lower fidelity approaches can be developed for the underground sewer model, the surface model, or both. However, given that the 2D surface flow model is generally more computationally expensive than the 1D sewer model, the approaches that rely solely on simplified representations of the sewer system do not necessarily result in a reduction of computation times. Therefore, most often lower fidelity approaches apply simplifications in both the sewer and the surface system. Different simplified representations can be used, but only the most popular ones, widely used in practice and research, are mentioned here.

The sewer system can be represented in the simplest approaches by subtracting a constant rate of rainfall [*Environment Agency*, 2013; *van Dijk et al.*, 2014] or by using a constant infiltration rate [*Chen et al.*, 2009a] that reflects the draining and storage capacity of the sewer. These approaches can be interesting if the underground drainage does not have much influence on the studied flood phenomenon or if network data are

missing [*Henonin et al.*, 2013]. More complicated yet popular approach is skeletonisation [*Leitão et al.*, 2010; *Simões et al.*, 2010], in which secondary elements of the network model are removed or changed.

On the other hand, the urban surface system can be represented in the simplest approaches by virtual reservoirs, located on top of the manholes, which store the overflow volumes. This approach alone cannot be used for a realistic assessment of flood dynamics, but can give an estimation of potential overflow locations [*Maksimovic and Prodanovic*, 2001]. On the other hand, the urban surface can be represented by a network of open channels and ponds, forming a so-called 1D-1D dual drainage model when combined with a 1D sewer model [*Maksimović et al.*, 2009]. This approach can provide accurate results if the surface flow is well channeled and the lateral dispersion of the flood wave is not significant [*Mark et al.*, 2004; *Lhomme et al.*, 2006; *Leandro et al.*, 2009]. If this is not the case (e.g. flat areas with multidirectional flow paths), a two-dimensional treatment of the surface flow hydraulics is required to resolve the complex flow paths. Model complexity can then be reduced by solving simplified forms of the 2D shallow water equations, typically the local inertial approximation or the diffusive wave approximation [*Neal et al.*, 2012], applying grid coarsening methods with a subgrid treatment to account for the topographic variability that is too small to resolve with the computational mesh [*Schubert and Sanders*, 2012], or using nested grids in studies that require a large computational domain [*Bermudez et al.*, 2017]. However, these models may still not fully satisfy the needs of city emergency management due to model complexity, setup data requirements and computing times. Further simplifications have thus been proposed in recent years, such as urban inundation models based on geographic information systems [*Chen et al.*, 2009b; *Jahanbazi and Egger*, 2014; *Zhang and Pan*, 2014] or on a cellular automata approach [*Ghimire et al.*, 2013; *Liu et al.*, 2015].

The physically-based models described above are being more and more complemented by data-driven models [*Solomatine and Ostfeld*, 2008], which use machine learning methods to approximate the response of the original model given explanatory variables. In the water resources field, artificial neural networks (ANN) are a popular function approximation technique [*Razavi et al.*, 2012] which can capture complex input/output relationships. They have been widely and successfully applied for modelling rainfall-runoff processes, forecasting streamflow [*Yaseen et al.*, 2015] and approximating rating curves [*Wolfs and Willems*, 2014], but applications to urban hydraulics and hydrology are still scarce [*Li et al.*, 2010]. ANN models have been applied to predict flooding at the nodes of a sewer network based on rainfall input [*Duncan et al.*, 2011, 2013], or to emulate Combined Sewer Overflow (CSO) dynamics and water quality variables [*Keupers and Willems*, 2015], relating rainfall and depth of flow in the CSO chamber [*Mounce et al.*, 2014]. [*Wolfs and Willems*, 2016] recently developed a conceptual sewer modelling approach which divides the sewer system into interconnected cells and uses ANN models to estimate flows between them. All the above studies have reported significant speed gains over conventional hydrodynamic models with an acceptable loss of accuracy for most intended applications, which highlights the potential of this approach for predicting urban flooding in real-time. The generalization ability beyond the training data is nevertheless a concern when applying this type of models.

In this study we propose and test two surrogate models of a 1D-2D dual drainage model. The first model relies on lower fidelity simplifications to represent the surface system. The second model combines both physically-based and data driven modelling approaches to develop a conceptual sewer and flood model. The capabilities of these two models for pluvial flood inundation mapping are evaluated and compared.

## 2. Methodology

## 2.1. Case study

The city of Ghent in Belgium was used as a test case. The studied area covers 27.5 km$^2$ of flat terrain and is highly urbanized, with a population equivalent of 43,626. Four regions of the study area, with sizes around 0.3 km$^2$, were defined to develop the surrogate models (Figure 1). Model performance was evaluated considering 4 rainfall events: 2 historical events, occurred on 28th July 2013 and 30th May 2016, and 2 events obtained with a stochastic rainfall generator [*Muñoz et al.*, 2015]. To set up surrogate model 2 (see §2.4), a total of 120 storms based on the rainfall data of the meteorological station of the Royal Meteorological Institute of Belgium at Uccle were created and employed. This additional data enhances the generalization capabilities of the data-driven modelling approach, and can increase its accuracy. The 120 storms are based on different historical storms with large rainfall intensities and/or cumulative volumes. Out of the 120 storms, 15 were created by merging 2 separate storms, leading to events with two consecutive peaks. Such succession of peaks can have a major impact of urban flooding, as storages (in the sewer and on the surface) can already be filled after the first peak. Also, the rainfall intensities of most storms were artificially increased to cause flooding, and thus create additional flood calibration data for surrogate model 2. Additionally, 8 synthetic storms with return periods from 2 to 100 years are used to derive pre-simulated flood maps.
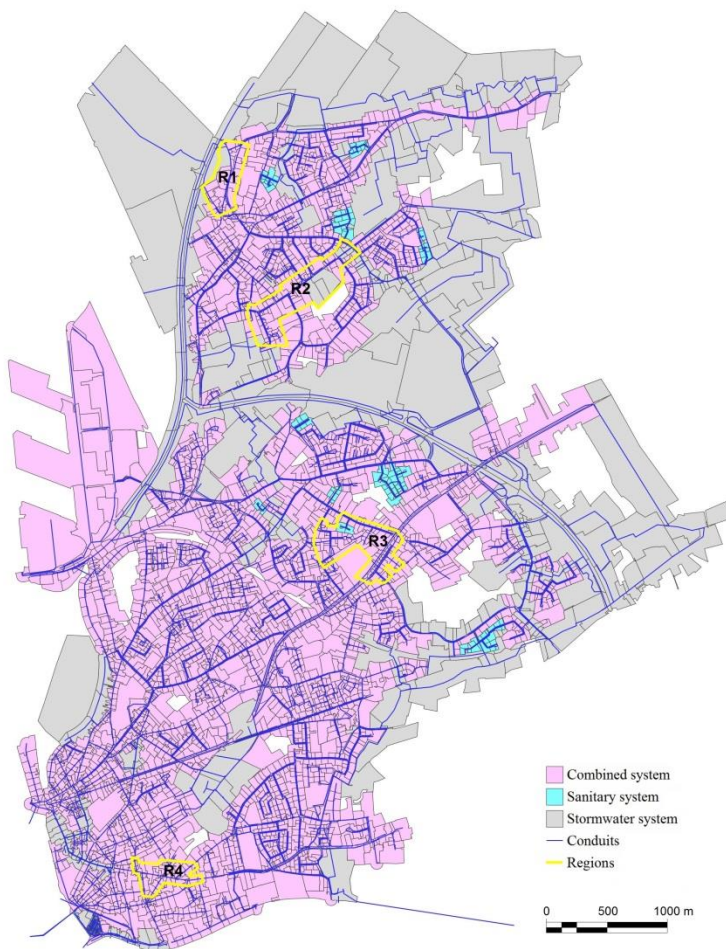


*Figure 1. Definition of the sewer network in the 1D-2D original model. The four regions analyzed are marked in yellow.*

## 2.2. 1D-2D dual drainage model

A detailed 1D-2D full hydrodynamic model of the sewer network and the surface of the city was implemented in InfoWorks ICM (Figure 1). The model covers an area of 27.5 km$^2$ and is comprised of 6025 conduits, 182 hydraulic structures and 5855 manholes. The resolution of the surface triangular mesh ranged from 3.75 m$^2$ up to 50 m$^2$ in the flood prone areas. The interaction between the 1D underground sewer conduits and the 2D surface was through the manholes, conceptualized as weirs. The double linear reservoir model (or Wallingford model) was used to route the flow from the surface to the manholes based on the input rainfall.

2.3. Surrogate model 1

Two different surrogate models of the above model were developed in this work. A schematic diagram of the two models is shown in Figure 2. The first surrogate model consists of a 1D representation of the sewer network in which the flood volumes are stored in virtual reservoirs on top of the manholes. A conical flood storage volume was defined for each manhole, specifying a cross-sectional area versus height relation. Two alternative definitions were implemented and tested: a default definition based on the size of the contributing areas draining to each node and an arbitrary total height of 100 m (named surrogate model 1A), and an enhanced definition based on the floodable areas estimated from the surface topography (named surrogate model 1B).
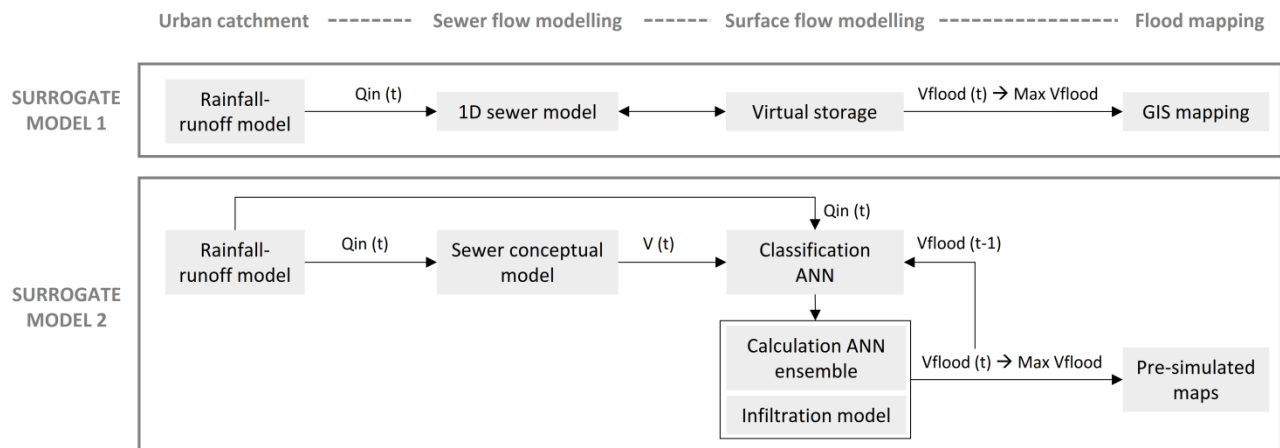


Figure 2. Schematic overview of the different sub-models that comprise the two surrogate models and their main input and output data: the runoff inflows Qin, the volumes in the sewer V and the flooded volumes Vflood.

The GIS mapping (Figure 2, Surrogate model 1) consists of a flood volume spreading algorithm that translates the maximum flood volume from surcharged manholes into a flood depth map. The method is based on an iterative algorithm that moves flood volumes over the terrain from grid cell to grid cell of the raster digital elevation model (DEM), and is similar to the procedure applied in [*Shapiro and Westervelt*, 1992; *Shook et al.*, 2013]. This approach does not require topographical pre-processing to define or remove depressions. It aims to move water over the landscape. When flood volumes are spread in this way, water moves into and out of depressions dynamically. The approach assumes that the drainage connectivity depends on the DEM. Thus, efforts must be made to minimize the potential connectivity inaccuracies such as areas with culvert or bridge crossings.

The flood volume spreading algorithm operates as follows. At each time step, a fraction of the depth in a grid cell is drained to the eight neighbouring grid cells. Each iteration therefore requires selecting an individual grid cell and a distribution of the volume for the neighbouring grid cells, after which the calculation is repeated for all grid cells within the flood zone. Specifically, a height difference is calculated between the selected grid cell

and its neighbouring grid cells. The height is calculated as the summation of the water depth and the ground elevation. If the neighbouring grid cell's height is greater than the selected grid cell's height, the selected grid cell receives a portion of the height difference, otherwise the neighbouring grid cell receives a fraction of the difference. Since such calculations are computationally expensive, the distribution of the volumes can be implemented using parallel computing. To allow for this, grid cells are grouped to avoid overlapping calculations. A total of 9 group simulations can be run concurrently to allow for a faster calculation of the flood extent. Figure 3 shows a group of grid cells that meet the non-overlapping criteria.

After each iteration, it is important to verify that the change in water depth to the previous iteration is within a tolerance limit. After thousands of iterations, the water surface depth approaches a realistic representation of the final water depth. When the maximum change in water depth at any location within the study area is smaller than the tolerance limit, the final water depth profile is achieved. It is worth noting that a 100-year flood plain was used for selecting the grid cells over which calculations were made. This was necessary to reduce limit the iterative procedure to those grids that would potentially be flooded.

Figure 4 illustrates the flood spreading procedure. For this study, an initial water depth map is required as a starting point for the flood spreading. This initial water depth map was calculated by estimating the flood depth at the manhole locations. Given the grid size, we can estimate the flood depth for a grid area. For manhole locations that are close to each other (within 2 meters), the volumes are combined and then spread on the grid cells close to the flooded grid cells (as shown in area 3 in Figure 3). Initial depths are usually high but after several iterations, the volume is spread over the terrain and realistic flood depths are achieved. The threshold for convergence was set at 10mm.
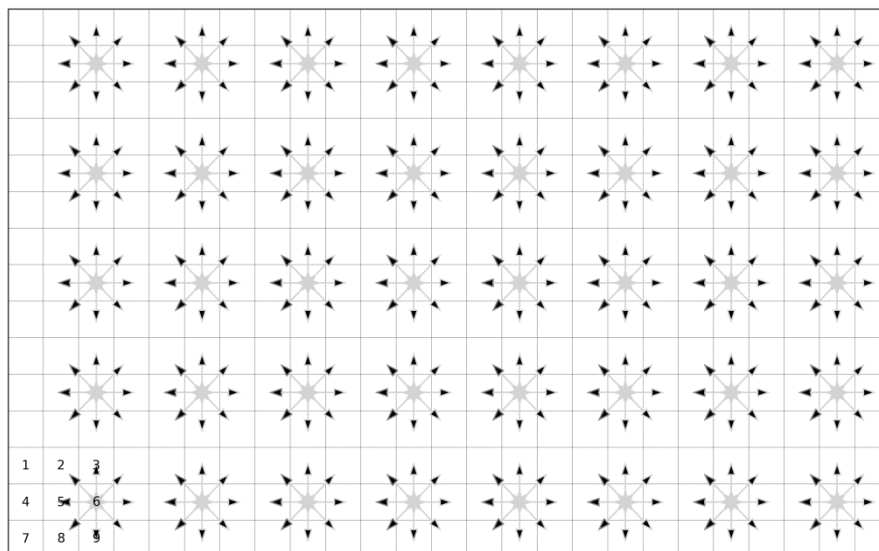


*Figure 3. Flood volume spreading for selected non overlapping grid cells of group 6 (centre cell in grey), and neighbouring cells (shown with black arrows).*
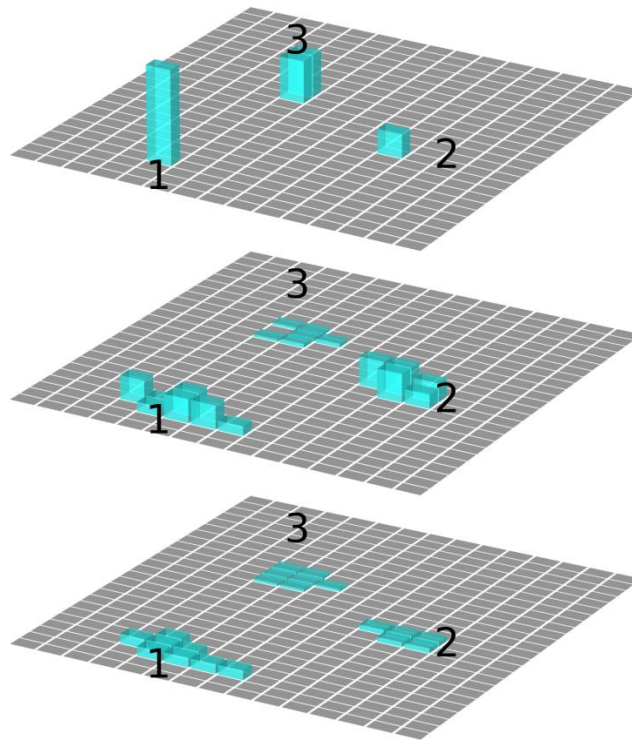
*Figure 4. Flood volume spreading at three time moments. Initial flood depth (top panel), intermediate flood depth (middle panel) and final depth (bottom panel). For simplicity topography is shown as a flat surface. In practice the grey cells have different elevations based on the DEM.*

## 2.4. Surrogate model 2

The second surrogate model combines a conceptual lumped hydraulic sewer model and a simplified flood model, using both data-driven and physically based model structures. Given rainfall and nearby river level series, the model yields volumes and maps at different locations in the city. The surrogate model was calibrated using simulation results of the detailed full hydrodynamic 1D-2D InfoWorks ICM model. The same rainfall runoff routing was used as in surrogate model 1. If such detailed 1D-2D model is unavailable, one could also calibrate this model to measurements. However, the level of model detail and its accuracy will depend on the availability and reliability of such measurements.

The underground system is emulated using the conceptual modelling approach described in [*Wolfs and Willems*, 2016]. First, the sewer network is divided into interconnected storage cells, each representing different parts of the sewer system. The conceptual model topology was chosen which minimizes flows between cells, which finally resulted in a model consisting of 13 storage cells (Figure 5). Experiments also showed that lumping processes further and creating fewer cells could lead to inaccurate model predictions. Next, the flows between the storage cells were emulated using a variety of model structures, including neural networks, transfer functions and piecewise linear relationships (see [*Wolfs and Willems*, 2016] for further information on the model structures and calibration algorithms). For each flow path, the most appropriate model structure was chosen based on the dynamics of the system. The model outcomes are volumes in each storage cell, and discharges at several locations.
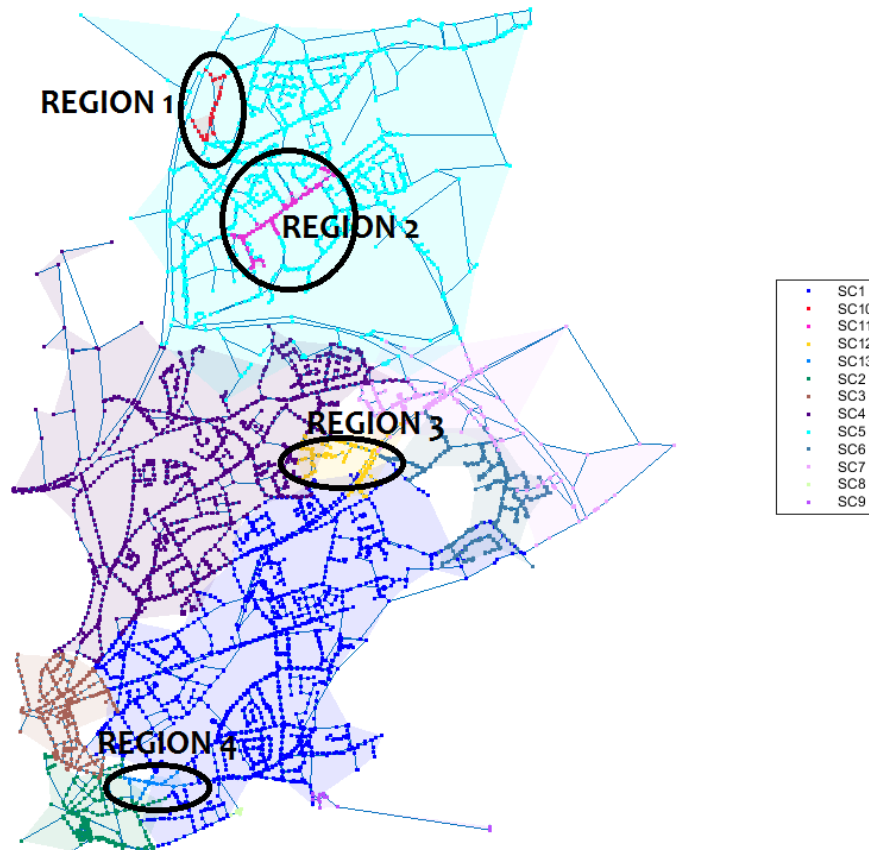
*Figure 5. Division of the entire area in storage cells (SCs) for the underground sewer hydraulic model.*

The flood model estimates surface flood volumes, aggregated in the pre-defined regions of the study area (Figure 1), using a serial connection of two artificial neural networks (ANNs). The first is a neural classification network that can identify when flooding from the underground system emerges. This network holds 10 hidden neurons arranged in one layer, and was trained using the scaled conjugate gradient approach minimizing cross-entropy. Next, an ensemble of five feedforward neural networks quantifies the magnitude of the flood. Using such ensemble reduces the risk of overfitting and increases the generalization capabilities. These ANNs have between 15 and 25 hidden neurons configured in one hidden layer. The flood model distinguishes flood volumes which are connected directly to the sewer system and thus can drain quickly, and flood volumes stored in local depressions which can only be emptied slowly via infiltration and evaporation. Only the former volumes are predicted directly via the ANN ensemble, which can be used to deduce the other volumes via GIS processing. By using such division, the slow decrease of the flood volume due ponding and infiltration, and events with two consecutive peaks, can also be simulated accurately by using a separate infiltration/evaporation module. Both ANNs use the same three inputs, which are the rainfall runoff volumes aggregated over 10 and 30 minutes windows, and the volume in the underground system of the closest storage cell.

In the final step of the modelling cascade, the predicted flood volumes are translated into flood depth maps (Figure 2). Pre-simulated scenarios, corresponding to synthetic storms with return periods from 2 to 100 years, are used to predict the spatial distribution of flooding within each region. The method uses the flood volume predicted by the surrogate model as input for scenario selection.

## 3. Results

The performance of the surrogate models is compared to that of the original 1D-2D model, first in terms of computational cost. The simulation times of the three models differ significantly. The 1D-2D dual drainage model takes approximately 408 CPU minutes to simulate a 6-hour event on an i7 processor at 3.40 GHz and 16 GB RAM. Surrogate model 1 takes around 91 CPU minutes, whereas the surrogate model 2 requires less than a second. This vast speed gains would enable numerous applications that require a large number of real time simulations, provided that the loss of accuracy is acceptable. This aspect is evaluated in the following.

The results of the surrogate models are compared with those obtained with the 1D-2D dual drainage model. Given the limited field data available, the 1D-2D dual drainage model is considered to provide the more accurate representation of the system behavior. The comparison is performed for each pre-defined region in terms of: (1) evolution of total flood volume in the surface, (2) flood inundation and depth maps corresponding to the maximum surface flood volume.

### 3.1. Flood volume prediction

The results show that the first surrogate model is sensitive to the definition of the virtual storage reservoirs in the sewer model (Figure 6). With the default definition (model 1A), the surface flood volumes are significantly underestimated, being the peak flood volumes between 30% and 67% lower than those obtained with the 1D-2D dual drainage model (Table 1). When the definition of the virtual storage is based on the floodable areas estimated from the surface topography (model 1B), the maximum flood volumes are better estimated. The maximum underestimation of peak flood volumes is 27 % (Table 1).

*Table 1. Nash-Sutcliffe efficiencies (NSE) indicating the fit of the surface volumes calculated with the surrogate models to the results of the 1D-2D dual drainage model. Differences in peak surface volumes (Δvol) in absolute and relative terms. Surface volumes are spatially aggregated in regions (R1 to R4 as indicated in Figure 1).*

| Event | Metrics | Model 1A | | | | Model 1B | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | R1 | R2 | R3 | R4 | R1 | R2 | R3 | R4 | R1 | R2 | R3 | R4 |
| May 2016 | NSE | -0.52 | 0.67 | -2.47 | -1.33 | 0.66 | 0.80 | -2.06 | -0.63 | 0.98 | 0.97 | -1.17 | 0.80 |
| | Δvol (m3) | -539 | -597 | -57 | -46 | -239 | -32 | -31 | -15 | 29 | -157 | 136 | 12 |
| | Δvol (%) | -60 | -38 | -30 | -65 | -27 | -2 | -16 | -21 | 3 | -10 | 71 | 17 |
| July 2013 | NSE | -0.16 | 0.54 | -0.57 | -0.25 | 0.61 | 0.97 | -0.37 | 0.42 | 0.96 | 0.94 | 0.61 | 0.94 |
| | Δvol (m3) | -198 | -234 | -45 | -80 | -80 | -82 | -23 | -26 | -39 | -87 | 39 | -1 |
| | Δvol (%) | -52 | -40 | -43 | -67 | -21 | -14 | -23 | -22 | -10 | -15 | 37 | -1 |
| Event A | NSE | -0.13 | 0.52 | 0.32 | - | 0.42 | 0.96 | 0.68 | - | 0.89 | 0.94 | 0.88 | - |
| | Δvol (m3) | -197 | -576 | -627 | 0 | -107 | -175 | 72 | 0 | -82 | -116 | -91 | 0 |
| | Δvol (%) | -48 | -40 | -28 | - | -26 | -12 | 3 | - | -20 | -8 | -4 | - |
| Event B | NSE | -0.26 | 0.29 | -2.72 | 0.10 | 0.55 | 0.92 | -2.66 | 0.30 | 0.94 | 0.90 | -0.32 | 0.79 |
| | Δvol (m3) | -406 | -264 | -2 | -135 | -152 | -70 | 0 | 33 | -140 | -17 | 14 | -136 |
| | Δvol (%) | -43 | -35 | -18 | -19 | -16 | -9 | -3 | 5 | -15 | -2 | 104 | -19 |

Surrogate model 1 can predict the start of the flood event, but is however unable to emulate the behavior of the surface flow, as shown by the low values of the Nash-Sutcliffe efficiencies in Table 1. The surcharged volumes return to the sewer system more quickly than with the 1D-2D dual drainage model, resulting in lower flood durations (as in R4 in Figure 6) and underestimation of peak flood volumes in multiple peak events (as

in R3 in Figure 6). The straightforward explanation is that this model does not replicate storage in local depressions that can only slowly infiltrate or evaporate, and thus cannot be used to predict the evolution of the volumes in the surface.

Surrogate model 2 emulates the results of the 1D-2D dual drainage model accurately, in terms of flood volumes, in regions 1, 2 and 4. The differences in peak flood volumes with respect to the predictions of the 1D-2D model are below 20 % in these regions. The model can provide accurate estimations not only of the maximum volume, but also of its evolution during the flood event, as reflected in the Nash-Sutcliffe efficiency coefficients obtained (above 0.79 in these regions, Table 1). In region 3, although the model can correctly identify the occurrence of a flood event, the flood volume predictions are poor. The surrogate model succeeds in predicting the moment of inundation correctly, but overestimates the maximum flood volume for events May 2016, July 2013 and event B. Performance can be increased by using more advanced neural networks (e.g. a higher number of neurons) and including more training data. In any case, the flood duration is emulated more accurately by the surrogate model 2 than with the surrogate model 1. It is clear that the simulation of the receding flood volume is more realistic with the surrogate model 2.
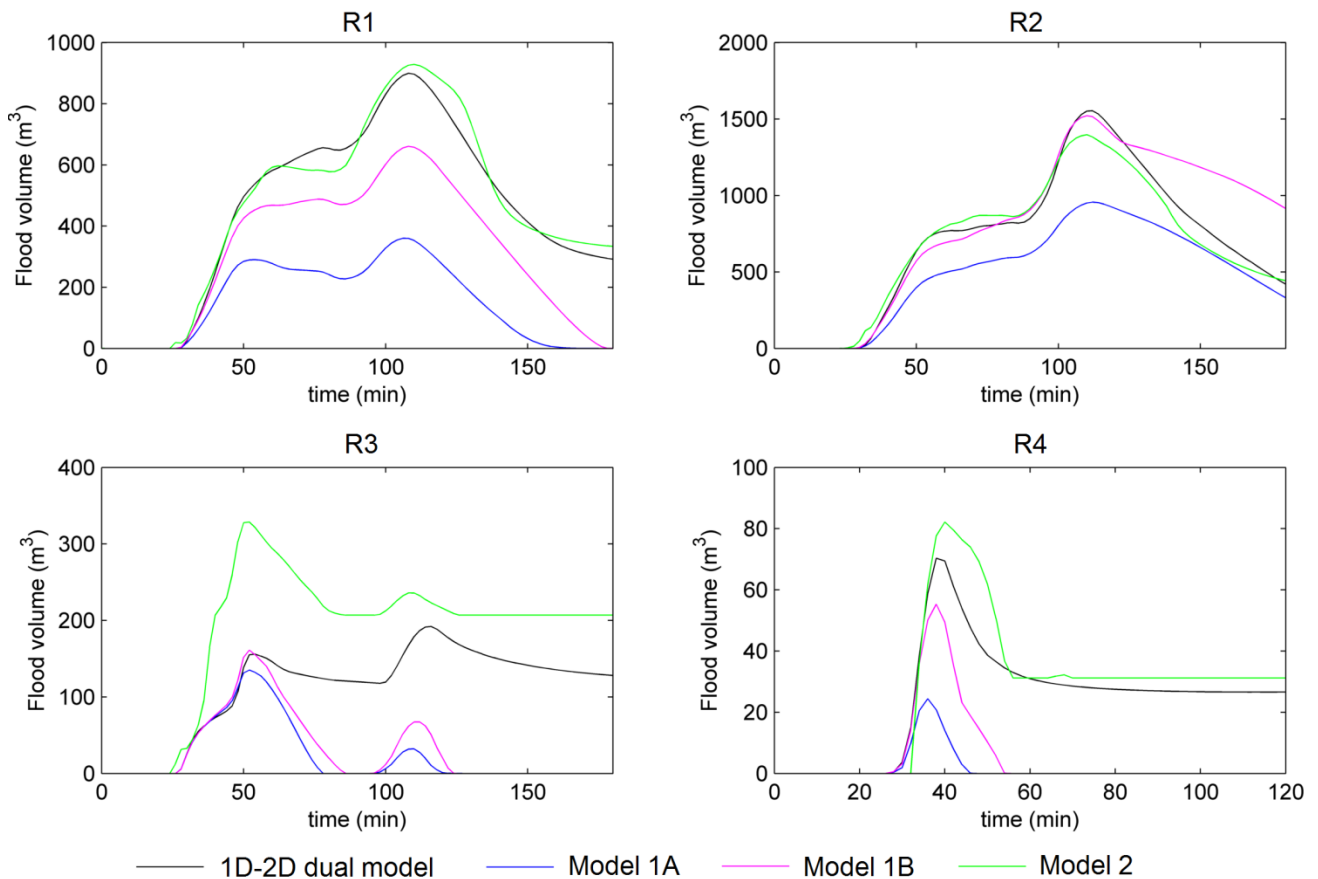


*Figure 6. Surface flood volumes predicted by the 1D-2D dual drainage model and the surrogate models for the May-2016 rainfall event. Surface volumes are spatially aggregated in regions (R1 to R4 as indicated in Figure 1).*

3.2. Flood inundation mapping

In order to compare the predictions of inundation extent obtained with the surrogate models and the 1D-2D dual drainage model, we employ the precision (p), the recall (r) and the F1 score as performance measures, defined as follows:

$$p = \frac{A}{B}$$

$$r = \frac{A}{C} \qquad\qquad (1)$$

$$Fl = 2\frac{p \cdot r}{p + r}$$

where A is the area correctly predicted as flooded by the surrogate model, B is the total area predicted as flooded by the surrogate model and C is total area predicted as flooded by the 1D-2D dual drainage model. A wet-dry threshold of 0.01 m is considered to delineate the flooded areas. A high precision $p$ means that most of the area that the surrogate model predicts as flooded is also predicted as such by the 1D-2D dual drainage model. A high recall $r$ means that most of the flooded area predicted by the 1D-2D-dual drainage model is well captured by the surrogate model. The lower and upper bounds of precision and recall are 0 and 1 respectively. In the context of early warning flood forecasting systems, a high recall ensures that the authorities are alerted to take action in most of the actual flooded areas, whereas a high precision avoids unnecessary actions and preparations in non-flooded areas. F1 is the harmonic mean of precision and recall, and is equal to 1 when the flooded areas predicted by the surrogate and the 1D-2D dual drainage model coincide exactly.

In order to evaluate the quality of water depth predictions of the surrogate models, the mean absolute error (MAE) of the flood depth predictions was computed. With this metric, the flood depth predictions obtained using the surrogate models were compared with the results of the original 1D-2D dual drainage model. The lower the MAE value, the better the agreement between the predictions of the two models. All the above metrics are commonly used for evaluating flood inundation models [*Aronica et al.*, 2002; *Pappenberger et al.*, 2007; *Liu and Pender*, 2013; *Stephens et al.*, 2014].

The surrogate models developed in this work use the peak flood volume as input for flood mapping (Figure 2 and Figure 7). Given the poor predictions of peak volumes obtained with model 1A, only model 1B and model 2 are compared at this stage. Table 2 shows the performance metrics obtained with both models. The surrogate models' predictions show a good overall agreement with the maps calculated with the 1D-2D dual drainage model. It should be noted that for certain events, surface flood volumes can be very low in some regions, and hence the derivation of a flood map is irrelevant (these cases are marked in italics in table 2). For event A and region 4, the 1D-2D model predicts no flooding for that location, and both surrogate models correctly produce a zero flood depth map. In the remaining cases, surrogate model 2 shows a slightly better performance than surrogate model 1, due to the underestimation of peak flood volumes in the latter. Most of the area that the surrogate model 1 predicts as flooded is also predicted as such by the 1D-2D dual drainage model, but the 1D-2D dual drainage model is mapping a larger flood extent (as in R1 in Figure 7). This results in higher precision than recall (Table 2). Surrogate model 2 presents high precision and recall, with F1 scores above 0.81 and MAE of around 0.02 m. The use of pre-simulated maps based on surface flood volumes can thus provide a reasonable characterization of the flood depth field at the street level.

*Table 2. Performance of the surrogate models based on the predicted flood inundation and depth maps: precision (p), recall (r), F1 score and mean absolute depth error (MAE). Results corresponding to maps with surface flood volumes < 200 m³ are marked with an asterisk (\*).*

| Event | Metrics | Model 1B | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | R1 | R2 | R3 | R4 | R1 | R2 | R3 | R4 |
| May 2016 | p | 0.75 | 0.82 | 0.38* | 0.57* | 0.93 | 0.78 | 0.47* | 0.93* |
| | r | 0.47 | 0.78 | 0.29* | 0.45* | 0.92 | 0.90 | 0.67* | 0.87* |
| | F1 | 0.58 | 0.80 | 0.33* | 0.50* | 0.92 | 0.84 | 0.55* | 0.90* |
| | MAE (m) | 0.06 | 0.04 | 0.04* | 0.03* | 0.02 | 0.03 | 0.03* | 0.01* |
| July 2013 | p | 0.73 | 0.80 | 0.57* | 0.71* | 0.91 | 0.91 | 0.57* | 0.96* |
| | r | 0.52 | 0.59 | 0.41* | 0.53* | 0.89 | 0.91 | 0.40* | 0.58* |
| | F1 | 0.61 | 0.68 | 0.48* | 0.60* | 0.90 | 0.91 | 0.47* | 0.72* |
| | MAE (m) | 0.05 | 0.05 | 0.03* | 0.04* | 0.02 | 0.01 | 0.02* | 0.02* |
| Event A | p | 0.72 | 0.77 | 0.78 | - | 0.93 | 0.94 | 0.83 | - |
| | r | 0.54 | 0.66 | 0.63 | - | 0.87 | 0.86 | 0.67 | - |
| | F1 | 0.62 | 0.71 | 0.70 | - | 0.90 | 0.90 | 0.74 | - |
| | MAE (m) | 0.05 | 0.04 | 0.06 | - | 0.02 | 0.02 | 0.04 | - |
| Event B | p | 0.77 | 0.70 | 0.45* | 0.77 | 0.92 | 0.94 | 0.19* | 0.97 |
| | r | 0.60 | 0.61 | 0.36* | 0.59 | 0.83 | 0.71 | 0.99* | 0.71 |
| | F1 | 0.67 | 0.65 | 0.40* | 0.67 | 0.87 | 0.81 | 0.32* | 0.82 |
| | MAE (m) | 0.06 | 0.04 | 0.03* | 0.04 | 0.02 | 0.02 | 0.02* | 0.02 |
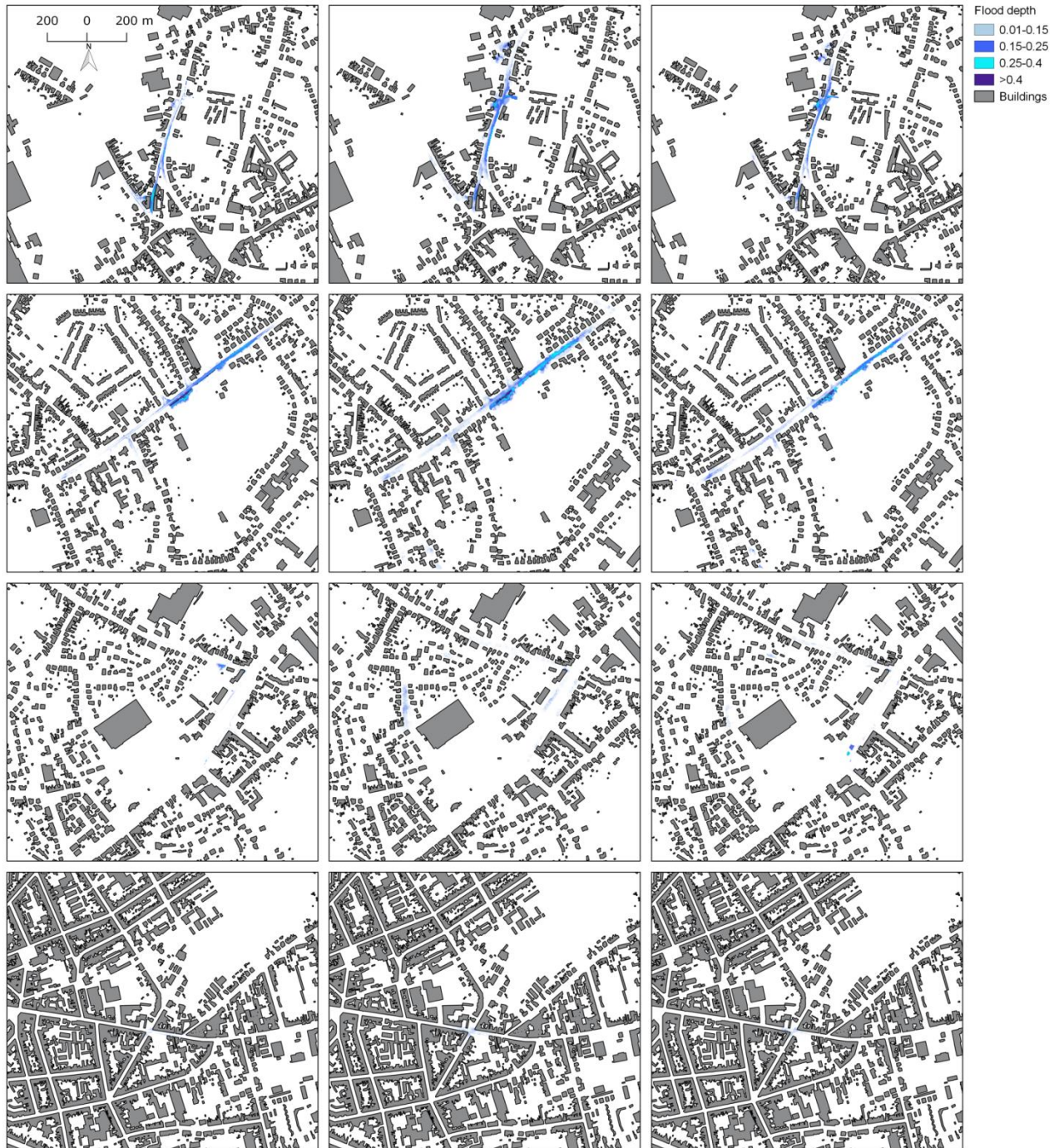
*Figure 7. Flood depth maps obtained with the surrogate model 1 (left), the 1D-2D dual drainage model (middle) and the surrogate model 2 (right) for May-2016 flood event. Regions R1 to R4 (up to down).*

## 4. Discussion

In the context of urban flood applications, the choice of a modeling strategy is mainly influenced by data availability and flood context (flood type, aimed application, etc.), and involves balancing accuracy, computation time, data needs and communication possibilities [*Henonin et al.*, 2013; *van Dijk et al.*, 2014]. Therefore, the proposed modelling approaches are compared on these topics.

Table 3 summarizes the capabilities and applicability of the two surrogate modelling approaches proposed and tested in this paper. Both models can identify the occurrence of a flood and predict the moment of inundation correctly as evidenced by Table 1 and Figure 6. Surrogate model 1 is not well suited to simulate flood evolution and durations accurately, whereas surrogate model 2 can reproduce flood dynamics more precisely (see Figure 6). Regarding the spatial flood prediction, surrogate model 1 can identify which manholes are surcharged under the existing flow conditions. On the contrary, surrogate model 2 predicts the surface flood volumes lumped in pre-defined regions, so it is not possible to identify the precise overflow locations.

*Table 3. Model capabilities and applicability. If both approaches are suitable but their performance differs, the preferred option in terms of accuracy is marked in italics.*

| | | Model 1 | Model 2 |
|---|---|---|---|
| Flood identification (yes/no) | | Yes | Yes |
| Flood magnitude (peak flood volume) | | Yes | *Yes* |
| Spatial flood prediction | Overflow location (manholes) | Yes | No |
| | Maximum flood extent | Yes | *Yes* |
| | Maximum flood depth | Yes | *Yes* |
| Flood dynamics | Flood start time | Yes | Yes |
| | Flood duration | No | Yes |
| | Flood evolution | No | Yes |
| Computational time (speed gain) | | ~5 | ~$10^4$ |

The flood maps predicted by the 1D-2D dual drainage model and the surrogate models compare well in terms of flood extent and flood depth. Both surrogate models use the peak flood volume as input for flood mapping. The flood maps thus reflect the errors in the prediction of peak flood volumes, which are higher with surrogate model 1 than with surrogate model 2. In surrogate model 1, the parameterization of the virtual reservoirs for each manhole has a significant impact on the simulated flood volumes, and the default parameterization underestimates flood volumes significantly at all locations. However, reasonably accurate flood maps can be produced with both mapping approaches (i.e., the flood volume spreading algorithm in surrogate model 1 and the pre-simulated maps in surrogate model 2) provided that good estimates of peak flood volumes are used. The mapping approaches proposed in this paper have a very simple setup and require few modelling choices. Almost all of the model configuration steps in both proposed approaches can be fully automated (if a detailed 1D/2D model is available), including the calibration of the flood module in surrogate model 2. The lack of required modelling choices in the proposed approaches constitutes an advantage over other methods such as the 1D representation of the urban surface, which requires pre-defining a network of flow paths and ponds. The latter is not straightforward and is subject to a high degree of uncertainty in very flat areas like the one studied here. However, the flood mapping approaches implemented in this study are not without limitations. The flood spreading approach assumes that flood volumes are distributed over the terrain without explicit consideration of physical processes. Thus unlike the physics-based hydrodynamic models that take into account the detailed processes by solving the Shallow Water Equations, the flood spreading approach uses simple water-elevation rules that rely on the geographic information system of neighbouring cells. Hence, flood velocity and flood volume losses which may occur as flood volume spreads over the surface are not included. Thus, if flooded locations are in a green zone, infiltration has an important impact on the flood volumes and depths. Improvements can be made by incorporating some physical processes within the

methodology. For instance, the effects of infiltration losses can be accounted for by including a loss factor. Nevertheless, the simplified flood spreading approach allows for a diagnostic assessment of flood prone areas as well as identifying flow path connectivity. For cases where flood hazard information is required in a relatively short time, such as emergency management, simplified flood volume spreading is of critical importance. It is important to note, however, that the accuracy of the results critically depends on the resolution and quality of the DEM. A better land surface representation helps to capture relevant features which minimizes anomalies in flood extents.

Table 3 shows the speed gains obtained, which are simply the ratio between the computational time of the 1D-2D dual drainage model and the surrogate models. The speed gains achieved, in particular with the surrogate model 2, would enable numerous applications that require results in a very short time (e.g., real-time applications) or a large number of model simulations (e.g., optimization problems or uncertainty assessments).

Finally, the surrogate models presented here are comprised of several components or modules, as shown in Figure 2. As such, they can be interchanged or combined to define new modelling approaches. For example, the 1D representation of the sewer network with virtual reservoirs (as in surrogate model 1) can be coupled with pre-simulated flood maps (as in surrogate model 2). Similarly, components which are deemed unnecessary for a particular application can be removed. An example of the latter could be the definition of an urban flood warning system based only on the 1D the sewer network model with virtual reservoirs (as in surrogate model 1), without a flood mapping component.

## 5. Conclusions

Two surrogate modelling approaches of a highly detailed 1D-2D dual drainage model were developed and compared to simulate pluvial flooding on an urban area in Belgium. Surrogate model 1 is based on a 1D sewer network model, in which flood volumes are stored in virtual reservoirs on top of the manholes. Flood mapping is performed by means of a GIS volume spreading algorithm. Surrogate 2 combines a conceptual lumped hydraulic sewer model with a simplified flood model, using both data-driven and physically based structures. Flood mapping is based on pre-simulated scenarios.

The first surrogate model can identify overflow locations and give an estimation of the maximum flood volume, provided that the parameterization of the virtual reservoirs reflects the urban surface topography. These data can then be used to derive a reasonably accurate maximum flood depth map based on the surface topography. The second surrogate model can emulate the evolution of the flood volumes accurately and can thus be used to predict the dynamics of the flood event. Pre-simulated maps show a good agreement with the flood maps predicted by the 1D-2D dual drainage model, in terms of flood extent and depth. Both surrogate models require shorter calculation times, surrogate model 2 showing clearly superior performance in this aspect, with speed gains above $10^4$. This vast speed gain and the reduced loss of accuracy demonstrate the great potential of the developed surrogate modelling approach for real-time use.

# References

Aronica, G., P. D. Bates, and M. S. Horritt (2002), Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE, *Hydrol. Process.*, *16*(10), 2001–2016, doi:10.1002/hyp.398.

Bermudez, M., J. C. Neal, P. D. Bates, G. Coxon, J. E. Freer, L. Cea, and J. Puertas (2017), Quantifying local rainfall dynamics and uncertain boundary conditions into a nested regional-local flood modelling system, *Water Resour. Res.*, doi:10.1002/2016WR019903.

Carr, R. S., and G. P. Smith (2006), Linking of 2D and Pipe hydraulic models at fine spatial scales, in *7th International Conference on Urban Drainage Modelling and 4th International Conference on Water Sensitive Urban Design*.

Chen, A. S., S. Djordjevic, H. J. Fowler, A. Burton, C. Walsh, H. Harvey, J. Hall, R. Dawson, and G. Wood (2009a), Pluvial flood modelling of the South East London Resilience Zone in the Community Resilience to Extreme Weather (CREW) Project, in *Flood and Coastal Risk Management Conference*, Telford, UK.

Chen, J., A. A. Hill, and L. D. Urbano (2009b), A GIS-based model for urban flood inundation, *J. Hydrol.*, *373*(1), 184–192, doi:10.1016/j.jhydrol.2009.04.021.

Deltares (2017), *SOBEK. Hydrodynamics, Rainfall Runoff and Real Time Control. User Manual SOBEK 2.15.*

DHI (2012), *MIKE FLOOD User Manual*.

van Dijk, E., J. van der Meulen, J. Kluck, and J. H. M. Straatman (2014), Comparing modelling techniques for analysing urban pluvial flooding, *Water Sci. Technol.*, *69*(2), 305, doi:10.2166/wst.2013.699.

Duncan, A., A. S. Chen, E. Keedwell, S. Djordjevic, and D. Savic (2013), RAPIDS: Early Warning System for Urban Flooding and Water Quality Hazards,

Duncan, A. P., A. S. Chen, E. C. Keedwell, S. Djordjević, and D. A. Savić (2011), Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks, , *351*.

Environment Agency (2013), *What is the updated Flood Map for Surface Water?*, Bristol, UK.

Ghimire, B., A. S. Chen, M. Guidolin, E. C. Keedwell, S. Djordjević, and D. A. Savić (2013), Formulation of a fast 2D urban pluvial flood model using a cellular automata approach, *J. Hydroinformatics*, *15*(3), 676, doi:10.2166/hydro.2012.245.

Henonin, J., B. Russo, O. Mark, and P. Gourbesville (2013), Real-time urban flood forecasting and modelling – a state of the art, *J. Hydroinformatics*, *15*(3).

Innovyze (2015), *InfoWorks ICM Help v5.5*.

Jahanbazi, M., and U. Egger (2014), Application and comparison two different dual drainage models to assess urban flooding, *Urban Water J.*, *11*(7), 584–595, doi:10.1080/1573062X.2013.871041.

Keupers, I., and P. Willems (2015), CSO water quality generator based on calibration to WWTP influent data, in *Proceedings of the 10th International Conference on Urban Drainage Modelling*, pp. 97–104, Québec.

Lacasta, A., M. Morales-Hernández, J. Murillo, and P. García-Navarro (2015), GPU implementation of the 2D shallow water equations for the simulation of rainfall/runoff events, *Environ. Earth Sci.*, *74*(11), 7295–7305, doi:10.1007/s12665-015-4215-z.

Leandro, J., A. S. Chen, S. Djordjević, and D. A. Savić (2009), Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood Simulation, *J. Hydraul. Eng.*, *135*(6), 495–504, doi:10.1061/ASCE HY.1943-7900.0000037.

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

Lhomme, J., C. Bouvier, E. Mignot, and A. Paquier (2006), One-dimensional GIS-based model compared with a two-dimensional model in urban floods simulation., *Water Sci. Technol.*, *54*(6–7), 83–91.

Li, X., F. Zhou, and S. Lodewyk (2010), Applications of Artificial Neural Networks in Urban Water System, in *Watershed Management 2010*, pp. 508–519, American Society of Civil Engineers, Reston, VA.

Liu, L., Y. Liu, X. Wang, D. Yu, K. Liu, H. Huang, and G. Hu (2015), Developing an effective 2-D urban flood inundation model for city emergency management based on cellular automata, *Nat. Hazards Earth Syst. Sci.*, *15*(3), 381–391, doi:10.5194/nhess-15-381-2015.

Liu, Y., and G. Pender (2013), Carlisle 2005 urban flood event simulation using cellular automata-based rapid flood spreading model, *Soft Comput.*, *17*(1), 29–37, doi:10.1007/s00500-012-0898-1.

Maksimovic, C., and D. Prodanovic (2001), Modelling of Urban Flooding—Breakthrough or Recycling of Outdated

Concepts, in *Urban Drainage Modeling*, pp. 1–9, American Society of Civil Engineers, Reston, VA.

Maksimović, Č., D. Prodanović, S. Boonya-Aroonnet, J. P. Leitão, S. Djordjević, and R. Allitt (2009), Overland flow and pathway analysis for modelling of urban pluvial flooding, *J. Hydraul. Res.*, *47*(4), 512–523, doi:10.1080/00221686.2009.9522027.

Mark, O., S. Weesakul, C. Apirumanekul, S. B. Aroonnet, and S. Djordjevi (2004), Potential and limitations of 1D modelling of urban flooding, *J. Hydrol.*, *299*, 284–299, doi:10.1016/j.jhydrol.2004.08.014.

Mounce, S. R., W. Shepherd, G. Sailor, J. Shucksmith, and A. J. Saul (2014), Predicting combined sewer overflows chamber depth using artificial neural networks with rainfall radar data, *Water Sci. Technol.*, *69*(6), 1326, doi:10.2166/wst.2014.024.

Muñoz, C., L. P. Wang, and P. Willems (2015), Towards a high resolution stochastic rainfall generator for urban applications, in *Proceedings of the 10th International Workshop on Precipitation in Urban Areas*, edited by P. Molnar and N. Peleg.

Neal, J., I. Villanueva, N. Wright, T. Willis, T. Fewtrell, and P. Bates (2012), How much physical complexity is needed to model flood inundation?, *Hydrol. Process.*, *26*(15), 2264–2282, doi:10.1002/hyp.8339.

Pappenberger, F., K. Frodsham, K. Beven, R. Romanowicz, and P. Matgen (2007), Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations, *Hydrol. Earth Syst. Sci.*, *11*(2), 739–752, doi:10.5194/hess-11-739-2007.

Razavi, S., B. A. Tolson, and D. H. Burn (2012), Review of surrogate modeling in water resources, *Water Resour. Res.*, *48*(7), n/a-n/a, doi:10.1029/2011WR011527.

Schmitt, T. G., M. Thomas, and N. Ettrich (2004), Analysis and modeling of flooding in urban drainage systems, *J. Hydrol.*, *299*(3), 300–311, doi:10.1016/j.jhydrol.2004.08.012.

Schubert, J. E., and B. F. Sanders (2012), Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency, *Adv. Water Resour.*, *41*, 49–64, doi:10.1016/j.advwatres.2012.02.012.

Shapiro, M., and J. Westervelt (1992), R.MAPCALC An Algebra for GIS and Image Processing,

Shook, K., J. W. Pomeroy, C. Spence, and L. Boychuk (2013), Storage dynamics simulations in prairie wetland hydrology models: evaluation and parameterization, *Hydrol. Process.*, *27*(13), 1875–1889, doi:10.1002/hyp.9867.

Simões, N. E., J. P. Leitão, Č. Maksimović, A. Sá Marques, and R. Pina (2010), Sensitivity analysis of surface runoff generation in urban flood forecasting, *Water Sci. Technol.*, *61*(10), 2595–2601.

Solomatine, D. P., and A. Ostfeld (2008), Data-driven modelling: some past experiences and new approaches, *J. Hydroinformatics*, *10*(1), 3, doi:10.2166/hydro.2008.015.

Stephens, E., G. Schumann, and P. Bates (2014), Problems with binary pattern measures for flood model evaluation, *Hydrol. Process.*, *28*(18), 4928–4937, doi:10.1002/hyp.9979.

Wolfs, V., and P. Willems (2014), Development of discharge-stage curves affected by hysteresis using time varying models, model trees and neural networks, *Environ. Model. Softw.*, *55*, 107–119, doi:10.1016/j.envsoft.2014.01.021.

Wolfs, V., and P. Willems (2016), Modular Conceptual Modelling Approach and Software for Sewer Hydraulic Computations, *Water Resour. Manag.*, 1–16, doi:10.1007/s11269-016-1524-2.

XP Solutions (2014), *XP SWMM. Stormwater &amp; Wastewater Management Model. Getting started manual.*

Yaseen, Z. M., A. El-Shafie, O. Jaafar, H. A. Afan, and K. N. Sayl (2015), Artificial intelligence based models for stream-flow forecasting: 2000–2015, , doi:10.1016/j.jhydrol.2015.10.038.

Zhang, S., and B. Pan (2014), An urban storm-inundation simulation method based on GIS, *J. Hydrol.*, *517*, 260–268, doi:10.1016/j.jhydrol.2014.05.044.

# Implementation of the Short-Term Ensemble Prediction System (STEPS) in Belgium and verification of case studies

Loris Foresti, Maarten Reyniers, and Laurent Delobbe

Radar and Lightning Detection group, Royal Meteorological Institute of Belgium, Brussels, Belgium
(loris.foresti@gmail.com)

The Short-Term Ensemble Prediction System (STEPS) is a probabilistic precipitation nowcasting scheme developed at the Australian Bureau of Meteorology in collaboration with the UK Met Office. In order to account for the multiscaling nature of rainfall structures, the radar field is decomposed into an 8 levels multiplicative cascade using a Fast Fourier Transform. The cascade is advected using the velocity field estimated with optical flow and evolves stochastically according to a hierarchy of auto-regressive processes. This allows reproducing the empirical observation that the rate of temporal evolution of the small scales is faster than the large scales. The uncertainty in radar rainfall measurement and the unknown future development of the velocity field are also considered by stochastic modelling in order to reflect their typical spatial and temporal variability.

Recently, a 4 years national research program has been initiated by the University of Leuven, the Royal Meteorological Institute (RMI) of Belgium and 3 other partners: PLURISK ("forecasting and management of extreme rainfall induced risks in the urban environment"). The project deals with the nowcasting of rainfall and subsequent urban inundations, as well as socio-economic risk quantification, communication, warning and prevention. At the urban scale it is widely recognized that the uncertainty of hydrological and hydraulic models is largely driven by the input rainfall estimation and forecast uncertainty. In support to the PLURISK project the RMI aims at integrating STEPS in the current operational deterministic precipitation nowcasting system INCA-BE (Integrated Nowcasting through Comprehensive Analysis).

This contribution will illustrate examples of STEPS ensemble and probabilistic nowcasts for a few selected case studies of stratiform and convective rain in Belgium. The paper focuses on the development of STEPS products for potential hydrological users and a preliminary verification of the nowcasts, especially to analyze the spatial distribution of forecast errors. The analysis of nowcast biases reveals the locations where the convective initiation, rainfall growth and decay processes significantly reduce the forecast accuracy, but also points out the need for improving the radar-based quantitative precipitation estimation product that is used both to generate and verify the nowcasts. The collection of fields of verification statistics is implemented using an online update strategy, which potentially enables the system to learn from forecast errors as the archive of nowcasts grows. The study of the spatial or temporal distribution of nowcast errors is a key step to convey to the users an overall estimation of the nowcast accuracy and to drive future model developments.

# Improving the applicability of gauge-based radar rainfall adjustment methods to urban pluvial flood modelling and forecasting using local singularity analysis

## L.-P. WANG[1], S. OCHOA-RODRIGUEZ[2], P. WILLEMS[1], C. ONOF[2]

[1]*Hydraulics Laboratory, KU Leuven, 3001, Heverlee (Leuven), Belgium (lipen.wang@bwk.kuleuven.be)*
[2]*Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK*

**Abstract** Gauge-based radar rainfall adjustment techniques have been largely-used to improve the applicability of radar rainfall estimates to large-scale hydrological modelling. Their applicability to urban hydrology is however insufficient since these techniques were mostly developed based upon the Gaussian approximations and therefore smoothed off the so-called 'singularity' (or non-normality) that can be observed in the fine-scale rainfall structure. Overlooking the singularities could be critical because their distribution is highly consistent with that of local extreme magnitudes. This deficiency may cause tremendous errors in the subsequent urban hydrological modelling. In this paper, a methodology is proposed to incorporate an existing gauge-based radar rainfall adjustment technique with the local singularity analysis, aiming for improving the applicability of existing adjustment techniques at urban scales. Three historical storm events recorded by a flow survey campaign in 2011 in Edinburgh (UK) were selected as case study to evaluate the proposed methodology. The result suggests that the proposed 'singularity-sensitive' methodology can in general better re-construct the non-normality in local rainfall structure and at the same time preserve the advantage of the original adjustment techniques of generating unbiased estimates.

**Key words** Gauge-based adjustment; urban rain; singularity; fractals

## INTRODUCTION

Traditionally, urban hydrological applications relied mainly upon rain gauge data as input as these provide accurate point rainfall estimates near the ground surface. However, they cannot capture the spatial variability of rainfall, which has a significant impact on the urban hydrological system and thus on the modelling of urban pluvial flooding. Thanks to the development of radar technology, weather radar has been playing an increasingly important role in urban hydrology. Radars can survey large areas and better capture the spatial variability of the rainfall, thus improving the short term predictability of rainfall and flooding. However, the accuracy of radar measurements is in general insufficient, particularly in the case of extreme rainfall magnitudes. This has a tremendous effect on the subsequent hydraulic model outputs.

In order to improve the accuracy of radar rainfall estimates while preserving their spatial description of rainfall fields, it is possible to dynamically adjust them based on rain gauge measurements. Studies on this subject have been carried out over the last few years, though most of them focus on the hydrological applications at large scales. A couple of recent research works have examined the applicability of these adjustment techniques to urban-scale hydrological applications and concluded that these techniques can effectively reduce rainfall bias, thus leading to improvements in the reproduction of hydraulic outputs (Wang et al., 2013). However, underestimation of storm peaks can still be seen after adjustment and this is particularly significant in the case of small drainage areas and for extreme rainfall magnitudes. This may be due to the fact that the underlying adjustment techniques, mainly based upon 1st or 2nd order (statistical-) moment approximations, cannot properly cope with the non-normality observed in urban scale applications. In fact, it is often the case that the radar image captures striking local extremes (albeit the actual rainfall depths may be inaccurate), but these structures are lost or smoothened through the merging process. These striking local extremes correspond to singularity points within the rainfall field and can be identified through a local singularity analysis (Cheng et al., 1994; Schertzer and Lovejoy, 1987).

With the purpose of improving this aspect, a methodology has been developed which identifies the local extremes or 'singularities' of radar rainfall fields and preserves them throughout the merging

process. A preliminary test of this methodology in an urban area in London (Wang and Onof, 2013a, 2013b) has demonstrated that the original Bayesian data merging technique (Todini, 2001) could be effectively improved by incorporating this singularity analysis. In this work, this incorporation has been further used to reconstruct a number of storm events observed in an urban catchment in Edinburgh during the Summer of 2011 and for which high density rainfall and flow data are available.

**EXPERIMENTAL SITE AND DATA SET**

As aforementioned, the proposed methodology was originally developed using the radar and raingauge data over the Maida Vale catchment (London) in June 2009. However, due to the confidential reason and lack of flow measurements, its impact on urban hydrological modelling could not be evaluated in this catchment. Therefore, in the context of this paper, the dataset of the Maida Vale catchment will be used merely for demonstrating the intermediate results in the development of the methodology, and the description of the catchment and the dataset used will not be given in this paper. For readers who are interested in the details, please find the link in (Wang and Onof, 2013b).

An alternative catchment in Portobello (Edinburgh area) was used in this paper as case study due to the completeness of rainfall and flow data. A full-scale test of rainfall estimation and the subsequent hydrological modelling was carried out in this catchment. A description of the catchment and the local monitoring data (including raingauge, flow and depth data) available and used in this study is next provided.

In addition to the local monitoring data, the experimental catchment is within the coverage of C-band radars operated by the UK Met Office. Radar rainfall estimates are available through the British Atmospheric Data Centre (BADC) with spatial and temporal resolutions of 1 km and 5 min, respectively. These estimates correspond to a quality controlled and multi-radar composite product generated with the UK Met Office Nimrod system, which includes corrections for the different errors inherent to radar rainfall measurements (Golding, 1998; Harrison et al., 2000).

**Portobello catchment (Edinburgh, UK)**

**Catchment description:** Portobello is a beach town located 5 km to the east of the city centre of Edinburgh, along the coast of the Firth of Forth, in Scotland (Figure 1a). The catchment is predominantly urban and has a drainage area of approximately 53 km$^2$. The storm water drainage system is mainly separate and drains from the south-west to the north-east (towards the sea).

**Hydraulic model:** The model of the sewer system of the Portobello catchment (Figure 1b) is setup in InfoWorks CS and was verified in 2011 based on the medium term flow survey data described below (using solely raingauge data as input). It comprises 2,916 nodes and 2,906 conduits. Rainfall is applied to the model through subcatchments and runoff is estimated using the NewUK model.

**Local monitoring data available for this catchment:** The only local monitoring data available for this catchment is that of the medium term flow survey used for the verification of the model. The flow survey was carried out between April and June 2011 and comprises data from 12 raingauges and 28 flow gauges (Figure 1b). Radar rainfall estimates (at 1 km and 5 min resolution) for the same period of the flow survey were obtained from the BADC.

**Selected storm events**

During the flow survey monitoring period, three relatively large storms were recorded and were used for the verification of the model. The same three storm events were used in this study to test the gauge based adjustment methods. The dates and main characteristics of these events are summarised in Table 1.

Table 1: Rainfall events selected for testing of adjustment methods in the Portobello catchment.

| Event | Date | Duration (hour) | RG Total (mm) | RG Peak Intensity (mm/h) | RD Total (mm) | RD Peak Intensity (mm/h) |
|---|---|---|---|---|---|---|
| Storm 1 | 06-07/05/2011 | 7 | 9.25 | 11.21 | 9.67 | 7.29 |
| Storm 2 | 23/05/2011 | 7 | 7.70 | 5.03 | 10.80 | 4.80 |
| Storm 3 | 21-22/06/2011 | 24 | 32.96 | 8.46 | 25.85 | 5.42 |

RG = Raingauge; RD = Radar. NOTE: The accumulation and peak intensity values shown in this table correspond to areal mean values for the entire domain under consideration.
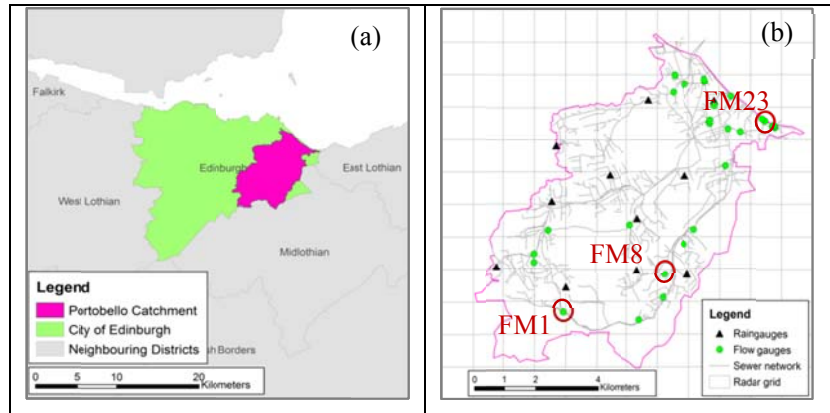


Figure 1 : Portobello catchment (a) general location; (b) sensor location, sewer network and radar grid over the catchment.

## METHODOLOGY

### Bayesian radar-raingauge data merging

The Bayesian data merging (BAY) is a dynamic adjustment method intended for real-time applications (Todini, 2001). It has been proven to outperform many other adjustment techniques in numerical experiments (Mazzetti and Todini, 2004) and in urban-scale hydrological applications (Wang et al., 2013). The underlying idea is to analyse the uncertainty of rainfall estimates from different sources (in this case, radar and raingauge sensors) and combine these estimates in such a way that the overall (estimation) uncertainty is minimised. The key techniques used in this method include the block-kriging interpolation (BK) and the Kalman filter. The principle of the BAY method is summarised as follows.

The first step of the BAY method is, for each time step $t$, to interpolate the raingauge measurements into a synthetic rainfall field using BK interpolation (steps (a) and (c) in Figure 2). This step generates comparable areal raingauge rainfall estimates ($y_t^{RG}$) to the radar estimates ($y_t^{RD}$), based upon which a field of errors (i.e. the bias at each radar grid location: $\varepsilon_t = y_t^{RG} - y_t^{RD}$) can be constructed (steps (d) and (e)). The covariance of this error field can be used to represent the uncertainty of radar estimates ($V_{\varepsilon_t}$) and is further compared and combined with the estimation error covariance of areal raingauge rainfall estimates ($V_{\varepsilon_t^{RG}}$, representing the uncertainty of raingauge estimates) that can be derived from the BK interpolation. The Kalman filter (Kalman, 1960) is employed herein (step (e)) to conduct this combination (where the radar data and the interpolated raingauge estimates respectively act as 'a priori estimate' and 'measurement' in the typical Kalman filter algorithm). The degree of 'certainty' of each type of estimates constitutes a gain value (the so-called Kalman gain, $K_t$) at each radar grid location, which determines the proportions of each type of estimates used to compute the merged output. As mentioned above, this gain value ensures the minimisation of the overall estimation uncertainty and is expressed as

4

$$K_t = V_{\varepsilon_t}(V_{\varepsilon_t} + V_{\varepsilon_t^G})^{-1},$$

and the merged output (i.e. the 'a posteriori' estimates in the Kalman filter) can be obtained from

$$y_t'' = y_t^{RD} + K_t(y_t^{RG} - y_t^{RD}).$$

It can be seen that the Kalman gain is a function of the covariances of radar and raingauge (estimation) errors. When $V_{\varepsilon_t} \gg V_{\varepsilon_t^{RG}}$ (or $K_t \approx 1$, i.e. radar estimates are of much higher uncertainty), the output estimates will be similar to the interpolated raingauge estimates. In contrast, when $V_{\varepsilon_t^{RG}} \gg V_{\varepsilon_t}$ (or $K_t \approx 0$), the output will be similar to the radar estimates.
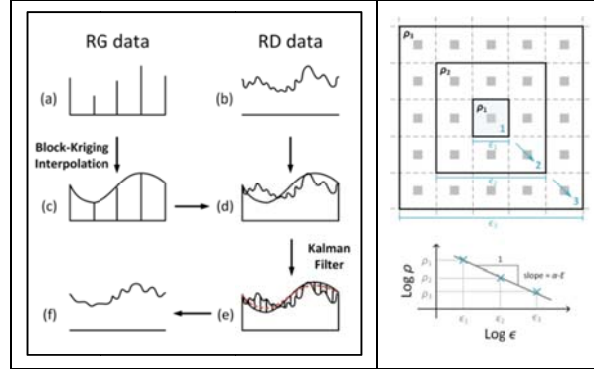


Figure 2: Schematic of the Bayesian radar-raingauge data merging (BAY) technique (left) and the local singularity analysis (right).

**Local singularity analysis**

The local singularity analysis is a simple yet effective method to identify the anomalies from geo-data. This method was proposed in (Cheng et al., 1994), and has been used for the estimation of geo-chemical concentration (Agterberg, 2007; Cheng and Zhao, 2011; Cheng et al., 1994). It employed the definition of the coarse Hölder exponent to identify the local scaling behaviour that follows a power-law relationship (i.e., the areal average measure increases as a power function when the area decreases; see Figure 2 (right)):

$$\rho(\mathbf{x}, \epsilon) \propto \epsilon^{\alpha(\mathbf{x})-E},$$

Where $\propto$ represents proportionality, the term $\rho(\mathbf{x}, \epsilon)$ represents the density of measure (e.g. concentration of geo-data) over a squared area with side-length $\epsilon$ centred at the location $\mathbf{x}$, $\alpha(\mathbf{x})$ is the singularity index (or the coarse Hölder exponent), and $E=2$ is the Euclidean dimension of a plane. By introducing a constant $c(\mathbf{x})$, one can further formulate this power-law relationship as an equation (Cheng et al., 1994):

$$\rho(\mathbf{x}, \epsilon) = c(\mathbf{x})\epsilon^{\alpha(\mathbf{x})-E}.$$

This equation constitutes a useful tool to decompose a rainfall magnitude at a given location $\mathbf{x}$ into two components (Wang et al., 2012): 1) the background (or non-singular, NS) magnitude $c(\mathbf{x})$ that is invariant as measuring scale $\epsilon$ changes and is more approximately normal than the original field, and 2) a local 'scaling' multiplier of which the magnitude changes according to the local singularity index $\alpha(\mathbf{x})$ and measuring scale $\epsilon$. It can be seen that, when $\alpha(\mathbf{x}) < 2$, the rainfall magnitude will strikingly increase as the measuring scale $\epsilon$ decreases (namely local enrichment), so it is a 'peak' singularity. In contrast, when $\alpha(\mathbf{x}) > 2$, the rainfall magnitude decreases as $\epsilon$ decreases (i.e. local depletion), and it is therefore a 'trough' singularity. When $\alpha(\mathbf{x}) = 2$, there is no singularity; the rainfall magnitude within a $\epsilon \times \epsilon$ area retains the same as scale changes (i.e. $\rho(\mathbf{x}, \epsilon) = c(\mathbf{x})$).

An example can be found in Figure 3 a and b of applying this local singularity analysis to the decomposition of a radar image. As compared to the original radar image (a1: RD), the spatial

structure of the non-singular component $c(\mathbf{x})$ (b1: NS-RD) is found to be smoother and more symmetric. In addition, the NS-RD estimates are of better normality than the original RD data (In Figure 3 a2 and b2, it can be seen that the NS-RD estimate quantiles are highly consistent with the Normal theoretical quantiles, but this is not case for the original RD estimates, where a much longer tail at the right end of the data distribution is expected). Therefore, the NS-RD estimates may be a more suitable input than the original RD estimates for many existing data merging techniques under the Gaussian approximation.

**"Singularity-Sensitive" radar-raingauge data merging**

The underlying idea of the proposed methodology is to use the local singularity analysis to decompose each radar snapshot into a non-singular image and a singularity map, where the former's distribution is closer to normality and thus can be better merged with the coincidental raingauge data under the Gaussian assumption. Afterwards, the singularity map is applied back to the merged image for recovering local extreme magnitudes. In implementation, the local singularity analysis is firstly carried out in the step (b) of Figure 2 (left) to decompose the RD image, then the non-singular part (NS-RD) of the original radar image is merged with the BK raingauge estimates (steps (c)-(f)) to obtain the non-singular merged (NS-BAY) estimates, and then the singularity map is multiplied back to the merged output to finally produce the 'singularity-sensitive' merged (SIN) estimates.

An example is shown in Figure 3 to demonstrate the variations in spatial structure of each estimate. It is observed that the structure of the BAY estimates tends to be smoother than that of the SIN estimates, where the latter can better preserve the non-normality of the original RD measurements than the former and thus its pattern is relatively realistic. In addition, due to the lack of raingauge information at the middle-left area of the experimental domain, the BK and BAY techniques failed in reproducing local extreme magnitudes measured by radar at that area. This indicates that the reliability of the BK and BAY estimates is very sensitive to the number and the deployment of raingauges, and the underlying Gaussian approximation causes the BAY technique to give more credit to the 'smooth' estimates generated by the BK technique and subsequently to neglect the local peaks in the RD data. This tendency towards 'smoother' estimates in the original BAY technique can be improved using the proposed methodology and therefore the missing local information at the middle-left area in the BAY can be re-constructed in the SIN estimates.
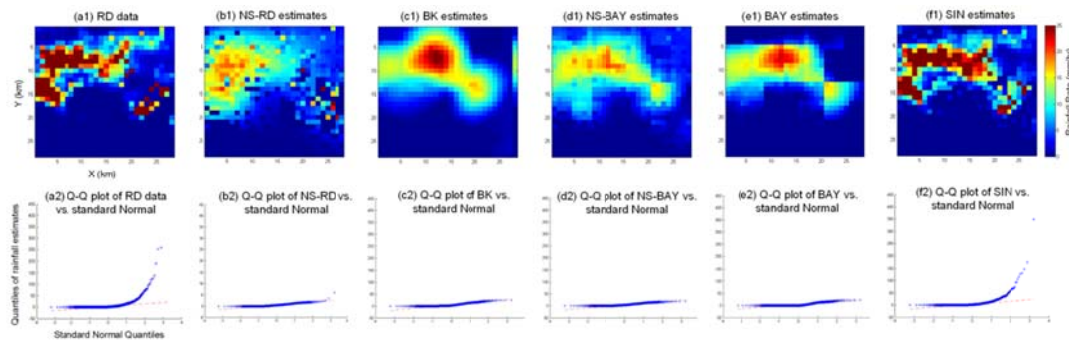


Figure 3: Snapshot images of RD (radar, a1), NS-RD (non-singular radar, b1), BK (block-kriged raingauge, c1) and NS-BAY (NS-RD merged with BK, d1), BAY (RD merged with BK, e1) and SIN (singularity-sensitive merged, f1) rainfall estimates, and the associated Q-Q plots (a2-f2) for demonstrating the degree of normality of each estimates.

**RESULTS AND DISCUSSION**

The proposed SIN methodology, as well as other gauge-based interpolation (BK) and adjustment techniques (BAY) mentioned above, was employed to reconstruct three historical storm events in Portobello (Table 1), and the resulting rainfall estimates were further used as input for hydraulic simulations. In addition, the adjusted estimates generated from a simple yet effective method,

called mean-field bias correction (MFB), were also included in the comparison because it has been a widely-used correction procedure used by many meteorological services (Goudenhoofdt and Delobbe, 2009; Harrison et al., 2000). This adjustment is implemented by comparing the summations of the RG and the co-located RD grid rainfall estimates over a specific area (i.e. the Portobello catchment area in this paper) and duration (i.e. one hour) to obtain a sample bias ratio (i.e. $B = \Sigma RG/\Sigma RD$). This ratio is then multiplied back to each radar grid estimate to ensure that the mean of RD rainfall estimates is the same as (or similar to) that of the RG measurements.

In the following, features of the rainfall estimates resulting from different interpolation and adjustment techniques are firstly presented and discussed. Then, the hydraulic outputs resulting from each rainfall input are presented, inter compared and discussed. Due to space constraints, only the results for Storm 1 are presented and discussed in detail. At the end of this section the results obtained for Storms 2 and 3 are briefly discussed and general conclusions are formulated. Results from Storm 1 were chosen as it is the most intense storm analysed for this catchment and, as such, it is the most relevant from an urban pluvial flood modelling perspective.

**Rainfall estimates**

The features of the rainfall estimates generated by different techniques were characterised by comparing them with the local RG measurements, in terms of areal average and individual-site time series. In Figure 4 (left), the result is presented of a direct comparison of areal average RG intensities versus areal average BK, RD and adjusted estimates' intensities at each time step throughout the whole Storm 1 period. As expected, BK estimates are in good agreement with RG estimates. With regards to RD estimates, it can be seen that they tend to overestimate small rainfall rates and underestimate the peak intensities. This tendency can be explained by the fact that the Z-R conversion that is used to convert radar reflectivity to rainfall rate has to statistically compromise to the range of rainfall rates that frequently occur (whereas the occurrence of very small and large intensities is relatively rare). It can be seen that both sources of error in RD estimates can be largely improved through adjustment techniques. Promising results are obtained from the BAY and, in particular, from the SIN merging methods, which are able to well reproduce low as well as high rainfall rates. As compared to the RD estimates, the MFB method does not seem to provide significant improvements in this respect and its performance is especially poor at higher intensities (which are of outmost importance in the modelling and forecasting of urban pluvial flooding).

Similar comparisons were conducted at each RG location, and the associated statistics are summarised in Figure 4 (middle) and (right). The simple linear regression analysis was applied to each pair of RG measurements and the co-located grid estimates obtained from different gauge-based interpolation and adjustment techniques. The result of these regression analyses can be evaluated in terms of $\beta$ (regression coefficient) an $R^2$ (coefficient of determination). These two statistics provide the measures of how well RG observations are replicated by the RD/BK/merged rainfall estimates at each gauging station. The $R^2$ measure ranges from 0 to 1, describing how much of the observed dispersion is explained by the modelled one. However, the systematic bias (under- or over-estimation) of the modelled estimates cannot be reflected by this measure. The slope of the simple linear regression analysis (i.e. $\beta$) was therefore employed to provide additional information to cope with the drawback of $R^2$ measures.

As expected, the BK estimates in general possess the highest $R^2$ values since the RD information was not taken into account (Figure 4 (right)). However, from the distribution of $\beta$ values of the BK estimates, one can find that the whole box and the whiskers are below the axis of unity (Figure 4 (middle)). A similar result can be found in the BAY estimates, where high $R^2$ values are observed and most of the $\beta$ values are below one. This indicates that both BK and BAY estimates tend to systematically underestimate the RG rainfall intensities at each gauging site. This may be caused by the underlying Gaussian approximation, which tends to smooth off some local extreme magnitudes.

The RD estimates possess the lowest $R^2$ and β values. This is expected because RD data provide rainfall information at a certain elevation above the ground, which is unlikely to be the same as the ground raingauge measurements. Nonetheless, a certain degree of the similarity between RD and RG estimates can be still observed. The MFB adjustment can slightly increase their similarity, but the effect is very limited since this method uses merely the mean-field estimate from the RG data but fully follows the spatial structure of RD estimates.

Although the 'areal average' behaviours of BAY and SIN estimates are similar, the SIN's 'individual-site' behaviour is very different from the BAY's. It can be found that the distribution of the $R^2$ values of the SIN estimates is somewhere between that of the BAY and RD estimates. This difference indicates that, as compared to the original BAY estimates, the SIN estimates inherit more features from the RD estimates. This is consistent with the underlying assumption of the SIN methodology, in which the reliability of the original RD data is improved after singularities are extracted. In addition, it can be found that the distribution of β values of SIN estimates is approximately symmetric to the axis of unity. This means no significant systematic under- or over-estimation is observed in the SIN estimates. This could be due to the process of singularity recovery of the proposed SIN methodology and the re-construction of the local extreme magnitudes (or the local singular quantities) that were smoothed off by the original BAY method.

The feature analysis of different rainfall estimates suggests that the proposed SIN methodology preserves the 'areal average' behaviour of the original BAY, but at the same time introduces more RD information into the data merging, and therefore stronger spatial and temporal variations can be found in the SIN estimates. The impact of these different features on the subsequent hydrological output is further evaluated in the following section.
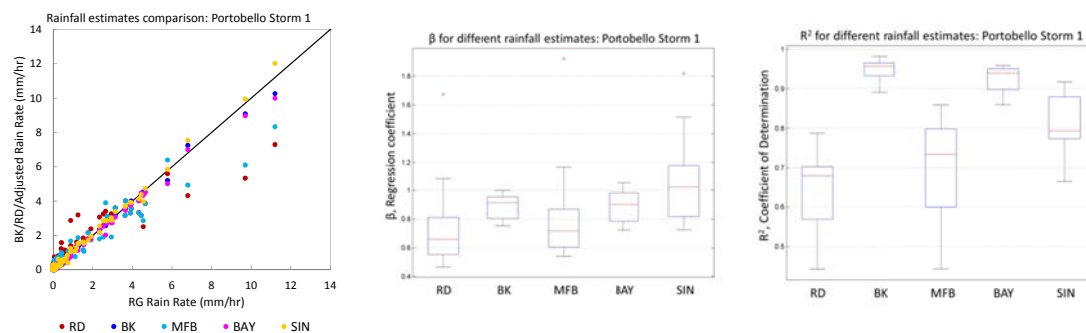


Figure 4: Comparisons of RG data and different rainfall estimates for Portobello's Storm 1: (left) Scatterplot of instantaneous areal RG vs. RD (red markers)/BK (blue)/MFB (light blue)/BAY (pink)/SIN (yellow) estimates; (middle and right) Boxplots of β an $R^2$ for the RG data vs. different rainfall estimates at each RG location.

**Hydraulic outputs**

In Figure 5 (left), a selection is presented of three observed vs. simulated flow and depth hydrographs from different locations within the catchment (respectively in the up-, mid- and downstream parts of the catchment) for Storm 1. In addition, in Figure 5 (right) boxplots are presented which show the distribution of the performance measures, i.e., Nash-Sutcliffe efficiency coefficient (NSE) (Nash and Sutcliffe, 1970) and relative error (RE) in peak flow output, for the simulated depths and flows at the different gauging stations for Storm 1. The RE measure is computed by dividing the difference of the simulated and the observed flow peaks ($S_{peak}$ - $O_{peak}$) by the observed one ($O_{peak}$). This measure gives an estimate of how well, in terms of magnitude, the simulation results can reproduce the true peak flows and depths. Negative RE values indicate that the model underestimates the observed peak flow/depth, while positive values indicate overestimation of the peaks. Moreover, the closer RE is to zero, the better.

From Figure 5 it can be seen that, even though the RG and RD totals are similar (RD is slightly

higher) for Storm 1 (Table 1) , the RD associated hydraulic outputs consistently underestimate flow and depth peaks, with the degree of underestimation changing from location to location and possibly increasing in the direction of flows within the catchment (i.e. larger underestimations are observed in gauging locations further downstream, as compared to upstream locations). The underestimation in hydraulic outputs, in spite of the small difference of the RG and RD totals, can be explained by the fact that the RD estimates cannot well reproduce high rainfall rates (Figure 4). This suggests that not only is it important to get the areal total rainfall accumulations right, but accurately capturing the peak rainfall intensities is also of outmost importance in order to appropriately reproduce the dynamic behaviour of the hydrological system and, in particular, the flow and depth peaks.

The MFB adjustment was found to provide some improvement over the original RD estimates; however, it is still insufficient to effectively reproduce peak rainfall intensities (Figure 4) and the associated flow and depth peaks (Figure 5 (left)). This confirms the fact that more dynamic adjustment radar rainfall adjustment methods which can better account for the spatial variability in the rainfall fields are required for urban-scales applications (rather than simple mean-field bias adjustments).

In general and as would be expected, the hydraulic outputs obtained with the BK estimates are very similar to the RG ones, with BK outputs sometimes performing better than the original RG ones. A striking difference between BK and RG hydraulic outputs and which is worth analysing can be observed in the hydrographs of gauging station 23 (Figure 5 (left, bottom)): it can be seen that the RG outputs largely overestimate the observed peak depth, while the simply interpolated BK rainfall input already leads to much more sound hydraulic results which are in better agreement with the observations. This confirms that accounting for the spatial variability of rainfall fields, even through simple kriging interpolation, could lead to significant benefits in the modelling.

The BAY and SIN outputs appear to be similar to the BK ones (and better than the original RD outputs), with the former (i.e. BAY and SIN) showing slightly more dynamic and realistic flow and depth patterns and with the SIN outputs performing better overall in terms of effectively reproducing peak depths and flows. The better performance of the SIN hydraulic outputs in this respect is clearly illustrated by the RE boxplots (Figure 5 (right, bottom)), where the median of the SIN associated RE for peak depths and flows is closer to zero and the dispersion of the results is smaller as compared to that of other hydraulic outputs, including the RG ones. An interesting example which also illustrates the potential benefits of the SIN method in terms of better capturing storm extremes can be found in gauging station 1: at this location the SIN methodology is the only one capable of generating a higher flow depth peak which is in better agreement with the observations (Figure 5 (left, top)).

From the results of Storm 1 it can be concluded that all adjustment methods can improve the applicability of the original RD rainfall estimates to urban hydrological applications, although the degree of improvement provided by each adjustment method is different. Overall, the BAY and SIN rainfall estimates lead to significantly better simulation results than the MFB adjusted estimates, with the SIN estimates performing particularly well at reproducing peak depths and flows.

In general, the results obtained for Storm 3 are in good agreement with those obtained for Storm 1. However, the results of Storm 2 are somehow different: in this event the RD accumulations were larger than the RG ones (see Table 1) and the RD peak rainfall intensity was very similar to the RG one (though this was a mild storm event with maximum observed rainfall rates in general low). This led to unusual results in which at many gauge stations the RD estimates resulted in better hydraulic outputs (i.e. closer to the observations) than the original RG ones. For this event the benefits of the merged rainfall estimates as compared to the original RD estimates in terms of hydraulic outputs are not evident (some improvements are achieved in NSE, but these are rather minor). Nonetheless, in this as well as in the other storms, there are many sources of uncertainty

affecting hydraulic outputs and it is difficult to separate the effect of rainfall inputs from that of model structure, model parameters and even from errors in flow measurements.
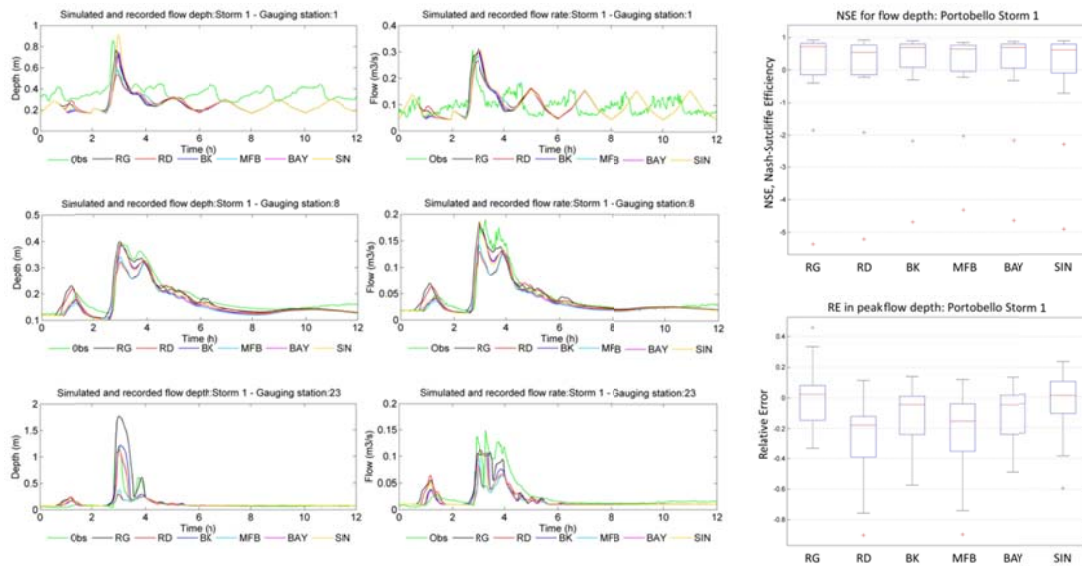


Figure 5: Comparisons of observed and simulated flow outputs for Portobello's Storm 1: (left) Flow rate and depth hydrographs at 3 gauge stations selected from different part of the catchment (from top to bottom, the points FM1, FM8 and FM23 in Figure 1 (b)); (right) Boxplots of NSE (top) and RE (bottom) for flow depths simulated using different rainfall inputs.

**CONCLUSIONS**

In this paper, a new gauge-based radar rainfall adjustment methodology was proposed, aiming at better merging raingauge and radar rainfall data at fine spatial and temporal scales. The proposed methodology incorporates the existing Bayesian data merging technique with the local singularity analysis. This incorporation has proven to be able to better cope with the non-normality (or singularity) in urban-scale rainfall data in this paper.

The applicability of the proposed SIN methodology to urban hydrology was tested and compared with other existing gauge-based interpolation and adjustment techniques (i.e. block-kriging (BK), mean-field bias correction (MFB) and Bayesian merging (BAY)). In terms of rainfall estimates, all adjustment methods led to areal average accumulations close to those recorded by raingauges, but only the BAY and SIN methods were capable of effectively reproducing high rainfall rates. These rates are usually poorly captured by radar, but are of outmost importance in order to properly reproduce flow peaks in the drainage system. Accordingly, in terms of hydraulic outputs, all merged rainfall products in general led to better results than the original radar (Nimrod) estimates. The Bayesian-based methods, in particular the SIN one, led to significantly better reproduction of the systems' dynamics as compared to the MFB adjusted estimates.

While the results are promising and the proposed SIN methodology shows great potential to be used in urban hydrological applications, the real benefits of its products in a verification context are likely to become more evident once the hydraulic model is re-verified. When this is done, the modeller will be able of analysing which rainfall product appears to be more 'logical/consistent' given the recorded depths and flows and the physical characteristics of the catchment and of the sewer system. In addition, the benefits of the SIN method are likely to become more evident in operational conditions, when storms outside the verification period are analysed and when data from fewer raingauge locations are available (when this is the case, radar becomes a necessary source of rainfall data).

10

## REFERENCES

Agterberg, F. P.: Mixtures of multiplicative cascade models in geochemistry, Nonlinear Process Geophys., 14, 201–209, 2007.

Cheng, Q., Agterberg, F. P. and Ballantyne, S. B.: The separation of geochemical anomalies from background by fractal methods, J. Geochemical Explor., 51(2), 109–130, 1994.

Cheng, Q. and Zhao, P.: Singularity theories and methods for characterizing mineralization processes and mapping geo-anomalies for mineral deposit prediction, Geosci. Front., 2(1), 67–79, 2011.

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

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, 2009.

Harrison, D. L., Driscoll, S. J. and Kitchen, M.: Improving precipitation estimates from weather radar using quality control and correction techniques, Meteorol. Appl., 7(2), 135–144, 2000.

Mazzetti, C. and Todini, E.: Combining raingauges and radar precipitation measurements using a Bayesian approach, in geoENV IV – Geostatistics for Environmental Applications, edited by X. Sanchez-Vila, J. Carrera, and J. J. Gómez-Hernández, pp. 401–412, Kluwer Academic Publishers., 2004.

Nash, J. E. and Sutcliffe, J. V: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10(3), 282–290, 1970.

Schertzer, D. and Lovejoy, S.: Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Processes, J. Geophys. Res., 92, 1987.

Todini, E.: A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, Hydrol. Earth Syst. Sci., 5(2), 187–199, 2001.

Wang, L.-P., Ochoa-Rodríguez, S., Simões, N. E., Onof, C. and Maksimović, C.: Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology., Water Sci. Technol., 68(4), 737–47, 2013.

Wang, L.-P. and Onof, C.: High-resolution rainfall field re-construction based upon Kriging and local singularity analysis, in Hydrofractals '13, Kos Island, Greece., 2013a.

Wang, L.-P. and Onof, C.: High-resolution rainfall field re-construction based upon Kriging and local singularity analysis (Presentation in Hydrofractals '13), [online] Available from: http://www.raingain.eu/sites/default/files/hydrofractals13_lwang.pdf, 2013b.

Wang, L.-P., Onof, C., Ochoa-Rodriguez, S. and Simões, N.: Analysis of Kriged rainfields using multifractals, in 9th International Workshop on Precipitation in Urban Areas: Urban Challenges in Rainfall Analysis, St. Moritz, Switzerland., 2012.

# Singularity-sensitive merging of radar and raingauge rainfall data

Li-Pen Wang (1), Patrick Willems (1), Susana Ochoa-Rodriguez (2), and Christian Onof (2)

(1) Hydraulics Laboratory, KU Leuven, 3001 Heverlee (Leuven), Belgium (Patrick.Willems@bwk.kuleuven.be), (2) Civil and environmental Engineering, Imperial College London, London, United Kingdom

Traditionally, urban hydrological applications relied mainly upon rain gauge data as input as these provide accurate point rainfall estimates near the ground surface. However, they cannot capture the spatial variability of rainfall, which has a significant impact on the urban hydrological system and thus on the modelling of urban pluvial flooding. Thanks to the development of radar technology, weather radar has been playing an increasingly important role in urban hydrology. Radars can survey large areas and better capture the spatial variability of the rainfall, thus improving the short term predictability of rainfall and flooding. However, the accuracy of radar measurements is in general insufficient, particularly in the case of extreme rainfall magnitudes. This has a tremendous effect on the subsequent hydraulic model outputs.

In order to improve the accuracy of radar rainfall estimates while preserving their spatial description of rainfall fields, it is possible to dynamically adjust them based on rain gauge measurements. Studies on this subject have been carried out over the last few years, though most of them focus on the hydrological applications at large scales. A couple of recent research works have examined the applicability of these adjustment techniques to urban-scale hydrological applications and concluded that these techniques can effectively reduce rainfall bias, thus leading to improvements in the reproduction of hydraulic outputs (Wang et al., 2013). However, underestimation of storm peaks can still be seen after adjustment and this is particularly significant in the case of small drainage areas and for extreme rainfall magnitudes. This may be due to the fact that the underlying adjustment techniques, mainly based upon Gaussian approximations, cannot properly cope with the non-normality observed in urban scale applications.

With the purpose of improving this aspect, a methodology has been developed which identifies the local extremes or 'singularities' of radar rainfall fields and preserves them throughout the merging process (Wang and Onof, 2013). Singularities are defined through the fact that the areal average rainfall increases as a power function when the area decreases (Cheng et al., 1994). In the proposed methodology singularities are first identified and extracted from the radar rainfall field. The resulting non-singular radar field is then used in the merging process and the singularities are subsequently and proportionally added back to the final reconstructed rainfall field. A full-scale testing of this methodology in an urban area in the UK has been conducted and the result suggests that the original Bayesian data merging technique (Todini, 2001) could be effectively improved by incorporating this singularity analysis.

References
Cheng, Q., et al., (1994) Journal of Geochemical Exploration, 51(2), 109-130.
Todini, E., (2001) Hydrology and Earth System Sciences, 5, 187-199.
Wang, L. et al., (2013) Water Science & Technology, 68(4), 737-747.
Wang, L. and Onof, C., (2013) Hydrofractals '13, Kos island, Greece.

# Multi-storm, multi-catchment investigation of rainfall spatial resolution requirements for urban hydrological applications

Susana Ochoa Rodriguez (1), Marie-Claire ten Veldhuis (2), Guendalina Bruni (2), Auguste Gires (3), Johan van Assel (4), Lipen Wang (5), Ricardo Reinoso-Rodinel (2), Abdellah Ichiba (3), Stefan Kroll (4), Daniel Schertzer (3), Christian Onof (1), and Patrick Willems (5)

(1) Imperial College London, London, United Kingdom, (2) Delft University of Technology, Delft, The Netherlands, (3) Ecole des Ponts ParisTech, LEESU, Paris, France, (4) Aquafin, Leuven, Belgium, (5) KU Leuven, Leuven, Belgium

Rainfall estimates of the highest possible resolution are required for urban hydrological applications, given the small size and fast response which characterise urban catchments. While significant progress has been made over the last few decades in high resolution measurement of rainfall at urban scales and in the modelling of urban runoff processes, a number of questions as to the actual resolution requirements for input data and models remain to be answered. With the aim of answering some of these questions, this work investigates the impact of rainfall estimates of different spatial resolutions and structures on the hydraulic outputs of models of several urban catchments with different characteristics. For this purpose multiple storm events, including convective and stratiform ones, measured by a polarimetric X-band radar located in Cabauw (NL) were selected for analysis. The original radar estimates, at 100 m and 1 min resolutions, were aggregated to coarser spatial resolutions of up to 1000 m. These estimates were then applied to the high-resolution semi distributed hydraulic models of four urban catchments of similar size (approx. 7 km2), but different morphological and land use characteristics; these are: the Herent catchment (Belgium), the Cranbrook catchment (UK), the Morée Sausset catchment (France) and the Kralingen District of Rotterdam (The Netherlands). When doing so, methodologies for standardising rainfall inputs and making results comparable were implemented. Moreover, the results were analysed considering different points at each catchment, while also taking into account the particular storm and catchment characteristics.

The results obtained for the storms used in this study show that flat and less compact catchments (e.g. polder areas) may be more sensitive to the spatial resolution of rainfall estimates, as compared to catchments with higher slopes and compactness, which in general show little sensitivity to changes in spatial resolution. While this study provides interesting insights, further investigation is still required in order to obtain a more complete answer regarding rainfall resolution requirements for urban hydrological applications. Future work should include testing on higher resolution fully distributed hydro models, as well as the analysis of many more storm events.

# Nested 1D-2D approach for urban surface flood modeling

Damian Murla and Patrick Willems

KU Leuven, Civil Engineering, Hydraulics Laboratory, Leuven, Belgium (Damian.MurlaTuyls@bwk.kuleuven.be; Patrick.Willems@bwk.kuleuven.be)

Floods in urban areas as a consequence of sewer capacity exceedance receive increased attention because of trends in urbanization (increased population density and impermeability of the surface) and climate change. Despite the strong recent developments in numerical modeling of water systems, urban surface flood modeling is still a major challenge. Whereas very advanced and accurate flood modeling systems are in place and operation by many river authorities in support of flood management along rivers, this is not yet the case in urban water management. Reasons include the small scale of the urban inundation processes, the need to have very high resolution topographical information available, and the huge computational demands.

Urban drainage related inundation modeling requires a 1D full hydrodynamic model of the sewer network to be coupled with a 2D surface flood model. To reduce the computational times, 0D (flood cones), 1D/quasi-2D surface flood modeling approaches have been developed and applied in some case studies. In this research, a nested 1D/2D hydraulic model has been developed for an urban catchment at the city of Gent (Belgium), linking the underground sewer (minor system) with the overland surface (major system). For the overland surface flood modelling, comparison was made of 0D, 1D/quasi-2D and full 2D approaches. The approaches are advanced by considering nested 1D-2D approaches, including infiltration in the green city areas, and allowing the effects of surface storm water storage to be simulated. An optimal nested combination of three different mesh resolutions was identified; based on a compromise between precision and simulation time for further real-time flood forecasting, warning and control applications. Main streets as mesh zones together with buildings as void regions constitute one of these mesh resolution (3.75m2 – 15m2); they have been included since they channel most of the flood water from the manholes and they improve the accuracy of interactions within the 1D sewer network. Other areas that recorded flooding outside the main streets have been also included with the second mesh resolution for an accurate determination of flood maps (12.5m2 – 50m2). Permeable areas have been identified and used as infiltration zones using the Horton infiltration model. A mesh sensitivity analysis has been performed for the low flood risk areas for a proper model optimization. As outcome of that analysis, the third mesh resolution has been chosen (75m2 – 300m2). Performance tests have been applied for several synthetic design storms as well as historical storm events displaying satisfactory results upon comparing the flood mapping outcomes produced by the different approaches. Accounting for the infiltration in the green city spaces reduces the flood extents in the range 39% - 68%, while the average reduction in flood volume equals 86%.

# Temporal interpolation of radar rainfall fields: meeting the stringent requirements of urban hydrological applications

Li-Pen Wang (1), Susana Ochoa-Rodriguez (2), Christian Onof (2), and Patrick Willems (1)

(1) Hydraulics Laboratory, KU Leuven, 3001, Heverlee (Leuven), Belgium , (2) Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK

Radar rainfall estimates are playing an increasingly important role in urban hydrological applications due to their better description of the spatial and temporal characteristics of rainfall. However, the operational radar rainfall products provided by national weather services (typically at 1 km / 5 min resolution) still fail to meet the stringent resolution requirements of urban hydrological applications. While the spatial and temporal resolution of rainfall inputs are strongly related, recent studies suggest that the latter generally constitutes a more critical factor and that temporal resolutions of ∼1-2 min (i.e. below those currently available) are required for urban hydrological applications, while spatial resolutions of ∼1 km (i.e. close to those currently available) appear to be sufficient. Traditional strategies for obtaining higher temporal resolution radar rainfall estimates include changes in radar scanning strategies and stochastic downscaling. However, the former is not always possible, due to hardware limitations, and the latter results in large ensembles members which hinder practical use. In this work a temporal interpolation method, based upon the multi-scale variational optical flow technique, is proposed to generate high temporal-resolution (i.e. 1-2 min) radar rainfall estimates. The proposed method has been successfully applied to obtain radar rainfall estimates at 1 and 2 min temporal resolutions from UK Met Office C-band radar products originally at 5 and 10 min temporal resolution and varying spatial resolutions of 1 km, 500 m and 100 m. The performance of the higher temporal-resolution radar rainfall estimates was assessed through comparison against local rain gauge records collected at a pilot urban catchment (size ∼ 865 ha) in North-East London. A further evaluation was conducted by applying the different rainfall products as input to the hydraulic model of the pilot catchment and comparing the hydraulic outputs against available flow and depth records. The results show that the temporally-interpolated rainfall estimates can better reproduce the small-scale dynamics of the storm events, leading to better reproduction of urban runoff.

# Spatial-temporal rainfall input resolution requirements for urban drainage modelling: a multi-storm, multi-catchment investigation

Susana Ochoa-Rodriguez (1), Li-Pen Wang (2), Auguste Gires (3), Rui Pina (1), Ricardo Reinoso-Rondinel (4), Guendalina Bruni (5), Abdellah Ichiba (3), Santiago Gaitan (5), Elena Cristiano (5), Johan van Assel (6), Stefan Kroll (6), Damian Murlà-Tuyls (2), Daniel Schertzer (3), Ioulia Tchiguirinskaia (3), Patrick Willems (2), Christian Onof (1), and Marie-Claire ten Veldhuis (5)

(1) Imperial College London, Urban Water Research Group, Civil and Environmental Engineering, London, United Kingdom, (2) Hydraulics Laboratory, KU Leuven, Leuven, Belgium, (3) Laboratoire Eau Environnement et Systèmes Urbains (LEESU), École des PontsParisTech, Paris, France, (4) Department of Geosciences and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands, (5) Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands, (6) Aquafin NV, Aartselaar, Belgium

Urban hydrological applications require high resolution precipitation and catchment information in order to well represent the spatial variability, fast runoff processes and short response times of urban catchments. Although fast progress has been made over the last few decades in high resolution measurement of rainfall at urban scales, including increasing use of weather radars, the resolution of the currently available rainfall estimates (typically 1 x 1 km2 in space and 5 min in time) may still be too coarse to meet the stringent spatial-temporal scales characteristic of urban catchments. In addition, current evidence is still insufficient to provide a concrete answer regarding rainfall input resolution requirements of urban hydrological applications. With the aim of providing further evidence in this regard, in the framework of the EU Interreg RainGain project a collaborative study was conducted which investigated the impact of rainfall estimates for a range of spatial and temporal resolution combinations on the outputs of operational semi distributed models of seven urban catchments in North-West Europe. Nine storm events measured by a dual polarimetric X-band weather radar, located in the Cabauw Experimental Site for Atmospheric Research (CESAR) of the Netherlands, were selected for analysis. Based on the original radar estimates, at 100 m and 1 min resolutions, 15 different combinations of coarser spatial and temporal resolutions, up to 3000 m and 10 min, were generated. These estimates were applied to the hydraulic models of the urban catchments, all of which have similar size (between 3 and 8 km2), but different morphological, hydrological and hydraulic characteristics. When doing so, methodologies for standardising model outputs and making results comparable were implemented. Results were analysed in the light of storm and catchment characteristics. Three main features were observed in the results: (1) the impact of rainfall input resolution decreases as catchment drainage area increases; (2) in general, the variation in temporal resolution of rainfall inputs affects hydrodynamic model results more strongly than variations in spatial resolution; (3) there is a strong interaction between the spatial and temporal resolution of rainfall input estimates and in order to avoid losing relevant information from the rainfall fields, the two resolutions must be in agreement with each other. Based on these results, initial models to quantify the impact of rainfall input resolution as a function of catchment size and spatial-temporal characteristics of storms are proposed and discussed.

**Towards a high resolution stochastic rainfall generator for urban applications**

C. Muñoz[1]*, L.-P. Wang[1], P. Willems[1]

1 Hydraulics section, Department of Civil Engineering, KU Leuven, Belgium

*Corresponding author: <u>carlos.munozlopez@bwk.kuleuven.be</u>

## Abstract

Traditionally, urban hydrological applications rely on spatially-uniform rainfall estimates derived from point measurements. However, several studies indicated that the performance of urban runoff and drainage simulations may largely depend upon the spatial and temporal variability of the rainfall input and that, therefore, rainfall input at high spatial and temporal resolutions is required. Historical data with such resolutions are, however, short in time. For this reason, long-term synthetic rainfall data generated by a stochastic model, accurately representing the real spatial-temporal rainfall properties, would be very beneficial. Such generated data is not constrained by the length of available historical data, hence would provide a better basis for urban-scale applications, such as urban pluvial flood risk analysis and urban drainage design. For this purpose, the early stages towards a development of a stochastic spatial-temporal rainfall generator for urban hydrological applications are presented in this work. The spatial stochastic generator for small spatial scales presented by Willems (2001) is employed as the starting point.

The core of Willems' generator is a conceptual rain storm model that aims to characterise rain storms with a number of physically-meaningful features (e.g. storm direction and velocity, rain cell extent, peak intensity and so on), and then to describe the statistical properties of each of them with a specific probability distribution. Based upon this, design rainfall with spatial variability can be simulated by firstly sampling a number of rain cell clusters over a 'simulation area', and then by moving the overall simulation area across the 'catchment' area with a given speed and direction. However, three main aspects where the model could be potentially improved were identified:

- The parameters of Willems' model were primarily calibrated based upon point rain gauge data, which could be insufficient to capture the real structure of rain storms and cells.
- The rain cells were conceptualised using bi-variate Gaussian model, which might oversimplify the real structures of small-scale rain cells and consequently smooth off the rain cell peaks.
- The temporal variability of the rain field was due to merely (stationary) field advection, so the temporal evolution of the field itself was not taken into account in Willems' model. This will lead to the 'unrealistic' isotropy in the spatial and temporal scaling behaviours of simulated storms (Seed et al. 1999).

To start tackling these deficiencies, the following strategies have been implemented:

- High-resolution radar images (provided by the Royal Meteorological Institute of Belgium) were used to better capture the spatial and temporal characteristics of rainfall fields. However, the use of radar images made the storm cell identification and tracking more challenging, in particular for small-scale rainfall details. To cope with this, two main algorithms were developed. First, a multi-threshold identification algorithm based upon the hierarchical threshold segmentation (HTS) method (Peak and Tag, 1994) was created. With this technique, adjacent storm cell clusters at small scales could be better identified and isolated. Secondly, an enhanced version of the TITAN algorithm (Dixon et al., 1993) by means of integrating optical flow techniques, was also developed. The performance of the enhanced TITAN tracking algorithm was evaluated by the ROC (Receiver Operating Curve) analysis (Fig. 1).
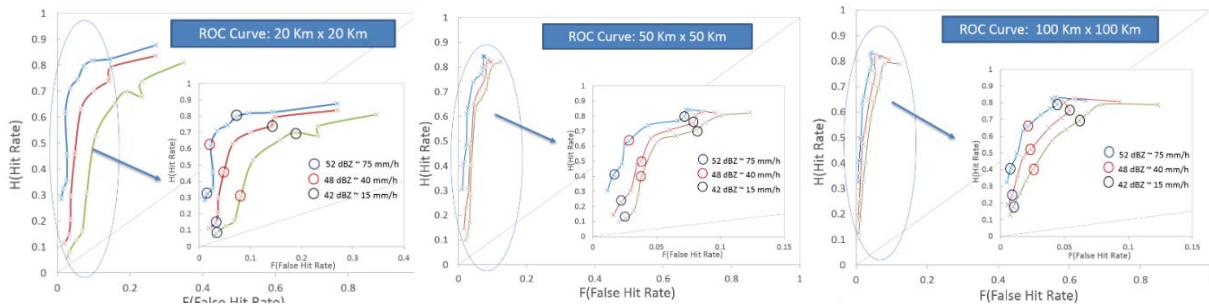
Fig 1: ROC plots for different threshold levels and catchment extents. The lower curve shows TITAN performance. Multi-threshold method inclusion improves the performance while optical flow integration gives the best results (Upper curve).

- A multi-layer conceptual model based upon the superposition of different rainfall entities (including high intensity peaks, rainfall cells, and small and large storm scales areas) was adopted. By making use of the improved TITAN algorithm, the rainfall fields were built by overlapping high intensity peaks within rain cells, which, in turn, were embedded in small mesoscales areas (Fig 2). Rain cells were still modelled using a bivariate Gaussian model. Small mesoscales areas were fitted as ellipses with constant intensity.
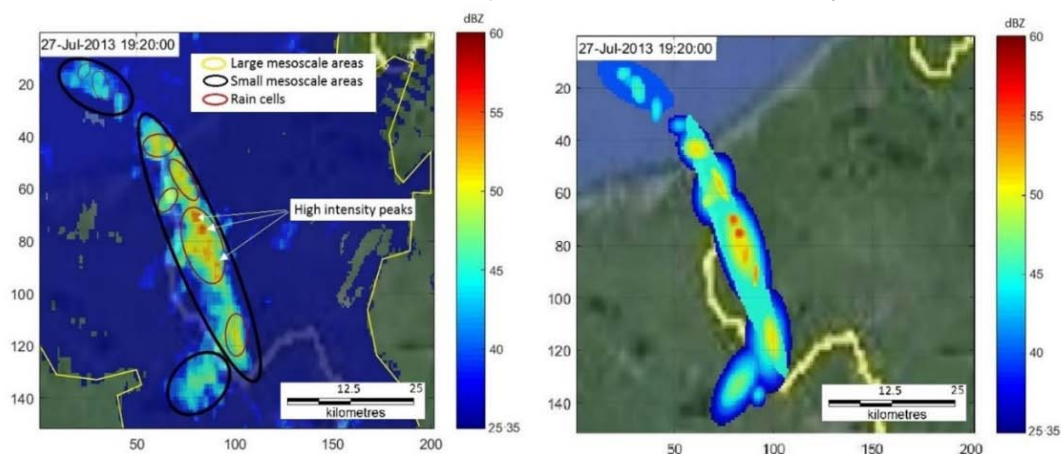


Fig 2: The original radar rainfall image (left) and the conceptualised rainfall image (right).

Results show that the integration of optical flow with a multi-threshold identification method considerably improved the performance of the original TITAN method. Furthermore, the conceptual model shows great potential to mimic the spatial distribution of rainfall intensities in convective rainfall fields. Therefore, the methods presented in this work enable better capturing the behaviour of small-scale and high-intensity storm cells, and suggest a great potential to provide added values to the implementation of Willems' rainfall generator.

### References

Dixon, M. and Wiener, G. (1993), TITAN: Thunderstorm identification, tracking, analysis, and now-casting—A radar-based methodology, J. Atmos. Oceanic Technol., 10, 785–797.

Peak, J. E. and Tag, P. M. (1994), Segmentation of satellite imagery using hierarchical thresholding and neural networks, J. Appl. Meteor., **33**, 605–616.

Seed, A. W., Srikanthan, R. and Menabde, M. (1999), A space and time model for design storm rainfall, J. Geophys. Res., 104(D24), 31623

Willems, P. (2001), A spatial rainfall generator for small spatial scales, J. Hydrol., 252, 126–44.

# Generation of high-temporal resolution QPEs through temporal interpolation of radar images: evaluation over multiple spatial-scales

L.-P. Wang*[1], S. Ochoa-Rodriguez[2], C. Onof[2], P. Willems[1]

1 Department of Civil Engineering, KU Leuven, Belgium
2 Department of Civil and Environmental Engineering, Imperial College London, UK

*Corresponding author: Lipen.Wang@bwk.kuleuven.be

## Abstract

Radar quantitative precipitation estimates (QPEs) are playing an increasingly important role in urban hydrology due to their better description of the spatial and temporal characteristics of rainfall. However, the operational radar QPE products provided by national weather services (typically at 1 km / 5-10 min resolution) still fail to meet the stringent resolution requirements of urban hydrological applications. While the spatial and temporal resolution of rainfall inputs are strongly related, recent studies suggest that the latter generally constitutes a more critical factor and that temporal resolutions of ~1-2 min are required for urban hydrological applications, while spatial resolutions of ~1 km appear to be sufficient (Ochoa-Rodríguez et al., 2015).

Traditional strategies for obtaining higher temporal-resolution radar QPEs include changes in radar scanning strategies and stochastic downscaling. However, the former is not always possible, due to hardware limitations, and the latter results in impractical large ensembles. In this work, an advection-based temporal interpolation method, based upon the multi-scale variational optical flow technique, is proposed to generate high temporal-resolution radar QPEs (Brox et al., 2004; Wang et al., 2015). The proposed method was used to generate radar QPEs at 1-min temporal resolutions from UK Met Office C-band radar QPEs originally at 5-min temporal resolution and varying spatial resolutions of 1 km, 500 m and 100 m (the former two are generated with C-band radar operating in 'long-pulse' mode, whereas the latter is generated with 'short-pulse' mode).
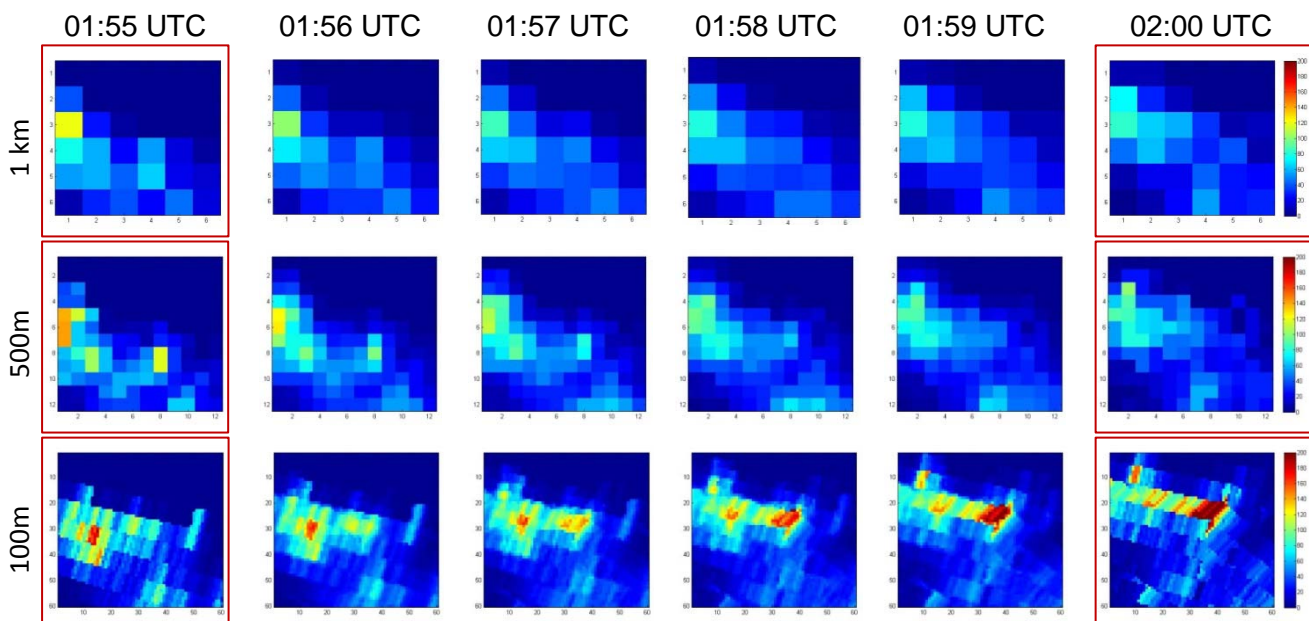
Fig 1: Snapshot images of the observed (images with red borders) and the temporally-interpolated radar rainfall fields across multiple spatial scales (from top to bottom: 1 km, 500 m and 100 m) during the peak intensity period of the event on 19th September 2014.

The performance of the temporally-interpolated radar QPEs, across a range of spatial resolutions, was assessed through comparison against local rain gauge records and through hydrological verification using as case study 3 storm events observed in a small urban catchment (~865 ha) in London for which dense rain gauge and sewer flow records, as well as a recently-calibrated high-resolution urban drainage model were available. Fig 1 shows the snapshot images of the observed and the interpolated radar images at different spatial resolutions (from top to bottom: 1 km, 500 m and 100 m) during the peak period of an event on 19th Sep 2014. As can be seen, the impact and added value of temporal interpolation is in particular evident for the radar images at higher spatial resolution. This is also confirmed by the comparison with local rain gauge records (Fig 2). Preliminary hydraulic results (which are not shown here) suggest that the temporally-interpolated rainfall estimates can better reproduce the small-scale dynamics of the storm events, leading to improved reproduction of urban runoff.
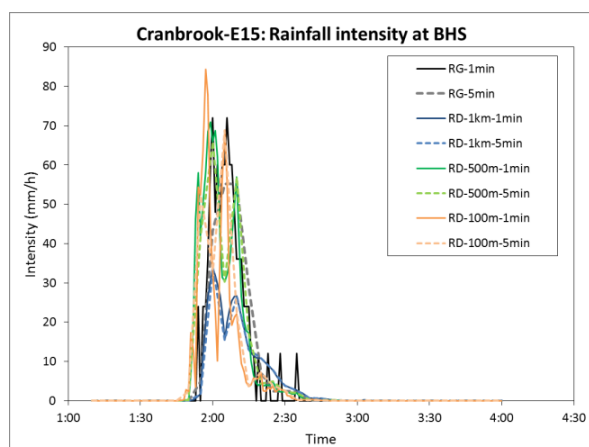


Fig 2: Comparison of rain gauge records (dark solid and dashed lines) against coincidental radar rainfall estimates at different spatial (1 km: blue lines, 500 m: green lines, 100m: orange lines) and temporal (5 min and 1 min) resolutions for the 19th September 2014 event.

**References**

Brox, T., Bruhn, A., Papenberg, N. and Weickert, J. (2004), High accuracy optical flow estimation based on a theory for warping, in: Pajdla, T., Matas, J. (Eds.), Computer Vision - ECCV 2004 SE - 3, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 25–36.
Ochoa-Rodríguez, S., Wang, L.-P., Gires, A., Pina, R., Reinoso-Rodinel, R., Bruini, G., van Assel, J., Kroll, S., Murlà-Tuyls, D., Ichiba, A., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P. and Ten Veldhuis, J.A.E. (2015), Impact of spatial and temporal resolution of rainfall inputs on semi-distributed urban drainage modelling outputs: a multi-catchment investigation, J. Hydrol. (revision)
Wang, L.-P., Ochoa-Rodríguez, S., Pina, R. D., Pessemier, M., Kroll, S., Van Assel, J., Willems, P. and Onof, C. (2015), Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment, J. Hydrol. (revision)

# Probabilistic urban inundation nowcasting

V. Ntegeka[1], D. Murla Tuyls[1], L-P. Wang[1], L. Foresti[2], M. Reyniers[2], L. Delobbe[2], K. Van Herck[3], L. Van Ootegem[3], P. Willems*[1]

1 KU Leuven, Hydraulics section, Leuven, Belgium
2 Royal Meteorological Institute of Belgium, Uccle, Belgium
3 KU Leuven, HIVA Research Institute for Work and Society, Leuven, Belgium

*Corresponding author: Patrick.Willems@bwk.kuleuven.be

## Abstract

A probabilistic model has been set up and evaluated for the nowcasting (short-term forecasting) of urban inundations. It consists of the following components:

- A rainfall nowcasting model based on the Short Term Ensemble Prediction System (STEPS), originally co-developed by the UK Met Office and Australian Bureau of Meteorology, but further customised for urban applications in Belgium (denoted STEPS-BE). It provides high-resolution (1 km / 5 min) rainfall nowcast ensembles with a 2-hour lead time.
- A hydraulic model that consists of the 1D sewer network and an innovative 'nested' 2D surface model to model 2D urban surface inundations at high resolution. The surface components are categorised into three groups and each group is modelled using triangular meshes at different resolutions; these include streets ($3.75 - 15$ m$^2$), high flood hazard areas ($12.5 - 50$ m$^2$) and low flood hazard areas ($75 - 300$ m$^2$).
- Functions describing urban flood damage and social consequences in relation to inundation depth. These functions were empirically derived based on questionnaires to people in the region that were recently affected by sewer floods.
- Statistical post-processing methods in order to produce probabilistic urban flood risk maps: spatial maps representing the probability of flooding.

The method has been implemented and tested for the villages Oostakker and Sint-Amandsberg, which are part of the larger city of Gent, Belgium. After each of the different above-mentioned components were evaluated, they were combined and tested for five recent historical flood events. The rainfall nowcasting, hydraulic sewer and 2D inundation modelling and socio-economical flood risk results each could be partly evaluated: the rainfall nowcasting results based on radar data and two rain gauges; the hydraulic sewer model results based on water level and discharge data at pumping stations; the 2D inundation modelling results based on limited data on some recent flood locations and inundation depths; the results for the socio-economical flood consequences of the most extreme events based on claims in the database of the national disaster agency. Different methods for visualisation of the probabilistic inundation results are proposed and tested.

# Sensitivity of urban drainage models to the spatial-temporal resolution of rainfall inputs: a multi-storm, multi-catchment investigation

S. Ochoa-Rodriguez*[1], L.-P. Wang[2], A. Gires[3], R. Reinoso Rondinel[4], R. D. Pina[1], J. van Assel[6], S. Kroll[6], D. Murlà-Tuyls[2], G. Bruni[5], A. Ichiba[3], S. Gaitan[5], E. Cristiano[5], D. Schertzer[3], I. Tchiguirinskaia[3], C. Onof[1], P. Willems[2], J.A.E. ten Veldhuis[5]

1 Urban Water Research Group, Imperial College London, UK
2 Hydraulics Laboratory, KU Leuven, Belgium
3 Ecole des Ponts ParisTech, Hydro-Meteorology and Complexity, Paris, France
4 Department of Geosciences & Remote Sensing, Delft University of Technology, the Netherlands
5 Department of Water Management, Delft University of Technology, the Netherlands
6 Aquafin NV, Belgium

*Corresponding author: s.ochoa-rodriguez@imperial.ac.uk

## Abstract

Urban hydrological applications require high resolution precipitation and catchment information in order to well represent the spatial variability, fast runoff processes and short response times of urban catchments (Berne et al., 2004). Although fast progress has been made over the last few decades in high resolution measurement of rainfall at urban scales, including increasing use of weather radars, recent studies suggest that the resolution of the currently available rainfall estimates (typically 1 x 1 km$^2$ in space and 5 min in time) may still be too coarse to meet the stringent requirements of urban hydrology (Gires et al., 2012). What is more, current evidence is still insufficient to provide a concrete answer regarding the added value of higher resolution rainfall estimates and actual rainfall input resolution requirements for urban hydrological applications. With the aim of providing further evidence in this regard, a collaborative study was conducted which investigated the impact of rainfall input resolutions on the outputs of the operational urban drainage models of four urban catchments in the UK and Belgium (Figure 1).
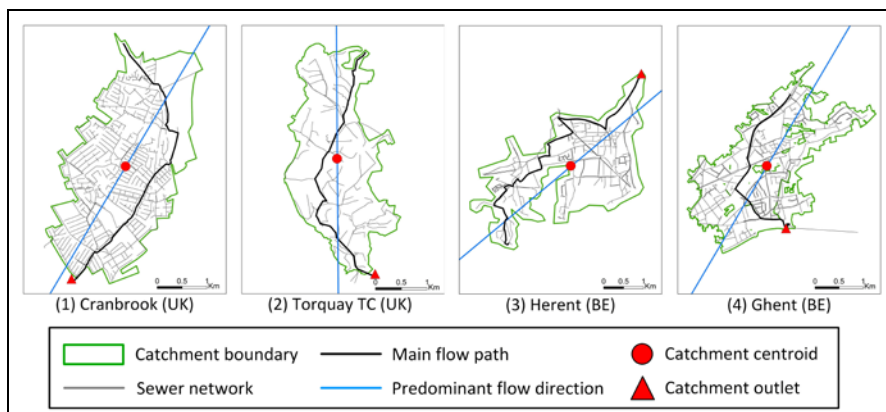
Fig 1: Boundary and sewer layout of the pilot urban catchments.

Nine storm events measured by a dual polarimetric X-band weather radar, located in the Cabauw Experimental Site for Atmospheric Research (CESAR) of the Netherlands, were selected for analysis. Based on the original radar estimates, at 100 m and 1 min resolutions, 15 different combinations of coarser spatial and temporal resolutions, up to 3000 m and 10 min, were generated. Coarser spatial resolutions were generated by averaging in space, whereas coarser temporal resolutions were generated through two different strategies: (1) by sampling radar images at the desired temporal resolution, thus replicating radar scanning strategies; (2) by averaging in time. The

resulting rainfall estimates were applied as input to the operational semi-distributed models of the urban catchments, all of which have similar size (between 5 and 8 km$^2$), but different morphological, hydrological and hydraulic characteristics (Figure 1). When doing so, methodologies for standardising model outputs and making results comparable were implemented. Hydrodynamic response behaviour was summarised using dimensionless performance statistics and was analysed in the light of drainage area and critical spatial temporal resolutions computed for each of the storm events. The main features observed in the results are the following (Figure 2):

- The impact of rainfall input resolution decreases rapidly as catchment drainage area increases.
- In general, the coarsening of temporal resolution of rainfall inputs affects hydrodynamic model results more strongly than the coarsening of spatial resolution. This is particularly the case when coarser temporal resolution rainfall estimates are generated through sampling of radar images; however, in the case of averaging in time, temporal resolution still shows a dominant effect over spatial resolution.
- There is a strong interaction between the spatial and temporal resolution of rainfall input estimates and in order to avoid losing relevant information from the rainfall fields, the two resolutions must be in agreement with each other.
- For the storms, models and drainage areas under consideration, temporal resolutions below 5 min appear to be required for urban hydrological applications, whereas spatial resolutions of the order of 1 km appear to be sufficient.

Based on these results, initial models to quantify the impact of rainfall input resolution as a function of catchment size and spatial-temporal characteristics of storms are proposed and discussed.
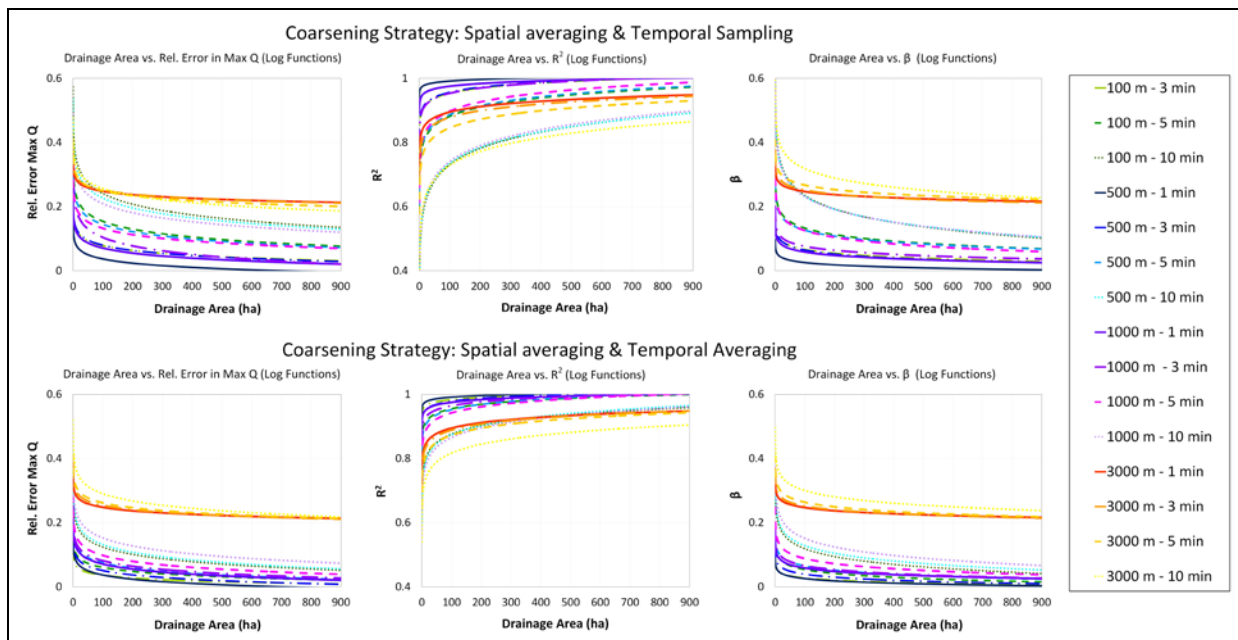


Fig 2: Logarithmic functions fitted to performance statistics of hydraulic outputs (relative error in maximum flow peak, coefficient of determination (R$^2$) and regression coefficient (β)) as a function of drainage area size, for different space-time resolution combinations. *Line type denotes different temporal resolutions (1 min = solid; 3 min = dash-dot; 5 min = dashed; 10 min = dotted) and colour range denotes different spatial resolutions (100 m = green; 500 m = blue; 1000 m = purple; 3000 m = orange).*

## References

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.

Gires, A., Onof, C., Maksimović, Č., Schertzer, D., Tchiguirinskaia, I. & Simoes, N. (2012). Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. Journal of Hydrology, 442, 117-128.

# Evaluation of radar-rain gauge merging methods for urban hydrological applications: relative performance and impact of gauge density

S. Ochoa-Rodriguez*[1], L.-P. Wang[2], A. Bailey[3,4], A. Schellart[4], P. Willems[2], C. Onof[1]

1 Urban Water Research Group, Imperial College London, UK
2 Hydraulics Laboratory, KU Leuven, Belgium
3 Clear Environmental Consultants, UK
4 Pennine Water Group, University of Sheffield, UK

*Corresponding author: s.ochoa-rodriguez@imperial.ac.uk

## Abstract

Rainfall estimates of very high accuracy and resolution are required for urban hydrological applications, given the high impermeability, small size and fast response which characterise urban catchments. Traditionally, urban drainage modelling applications have relied mainly upon rain gauge data as input, given that these sensors provide relatively accurate point rainfall estimates near the ground. However, they cannot capture the spatial variability of rainfall, which has a significant impact on the urban hydrological system and thus on the modelling of urban runoff. With the advent of weather radars, radar quantitative precipitation estimates (QPEs) with higher temporal and spatial resolution have become increasingly available and have started to be used operationally for urban storm-water modelling. Nonetheless, the insufficient accuracy of radar QPEs, arising from the indirect measurement of rainfall -often significantly high above ground-, has proven problematic and has hindered its widespread practical use (Schellart et al., 2012). In order to improve the accuracy of radar rainfall estimates while preserving their spatial description of rainfall fields, it is possible to dynamically adjust them based on rain gauge measurements. Gauge-based adjustment of radar QPEs, also referred to as radar-rain gauge combination or merging, has been an active topic of research over the last few decades and has proven effective to improve the accuracy of radar QPEs, thus improving their applicability for hydrological applications. However, most gauge-based adjustment methods have been tested and applied at large spatial and temporal scales -of the order of thousands of square kilometres and at temporal resolutions $\geq$ 1 h - (e.g. (Goudenhoofdt & Delobbe, 2009)), and their suitability for small-scale urban hydrology is seldom explored.

In this work we evaluate the performance of several radar-rain gauge merging techniques of various degrees of complexity at urban scales. The techniques under investigation were selected on the grounds of their widespread use and/or their relative performance against other existing techniques, as reported in previous studies (Goudenhoofdt & Delobbe, 2009; Wang *et al.*, 2013; Jewell & Gaussiat, 2015). The tested techniques include the simple mean field bias (MFB) correction, the Kriging with external drift (KED), and the more advanced Bayesian (BAY) merging (Todini, 2001) and singularity-sensitive Bayesian (SIN) merging (Wang *et al.*, 2015). The study area is a sub-catchment of Birmingham (drainage area ~ 67 km$^2$), UK, for which Met Office C-band radar QPEs (at 1 km / 5 min resolution), as well as records from 20 rain gauges (at 2 min resolution) and 41 flow gauges (at 2 min resolution) over a 6 month period are available. The relative performance of the different merging methods is first assessed on an event basis through comparison against rain gauge records and through hydrological verification (Figure 1). Moreover, the effect of rain gauge density on the performance of the merging methods is investigated. For this purpose, a simple approach of removing gauges from an initially dense network of rain gauges was applied; the selected approach ensures realistic configuration of rain gauge networks of different densities. The initial conclusions of this study are the following:

- All adjustment methods improve the applicability of the original radar and rain gauge QPEs estimates to urban hydrological applications (Figure 1); however, the degree of improvement varies for each method.

- In general, MFB is insufficient for satisfactorily correcting the errors in radar QPEs (Figure 1 (a, b)) and this is evident in the associated hydraulic outputs, which fail to properly reproduce peak depths and flows. This suggests that more dynamic and spatially-varying adjustment methods are required for urban hydrological applications.
- At high rain gauge densities (~1 rain gauge every 3 km$^2$), KED, BAY and SIN rainfall estimates show very good quantitative performance, both in terms of comparison against rain gauge records and in terms of their ability to reproduce observed urban runoff (Fig. 1). The SIN QPEs perform particularly well at reproducing peak rainfall intensities and associated depths & flows.
- At low rain gauge densities (~ 1 rain gauge per 16 km$^2$) KED, which is one of the most popular methods (Goudenhoofdt & Delobbe, 2009; Jewell & Gaussiat, 2015), performs poorly, and the advantage of the BAY and in particular the SIN method becomes more evident (Fig. 2).
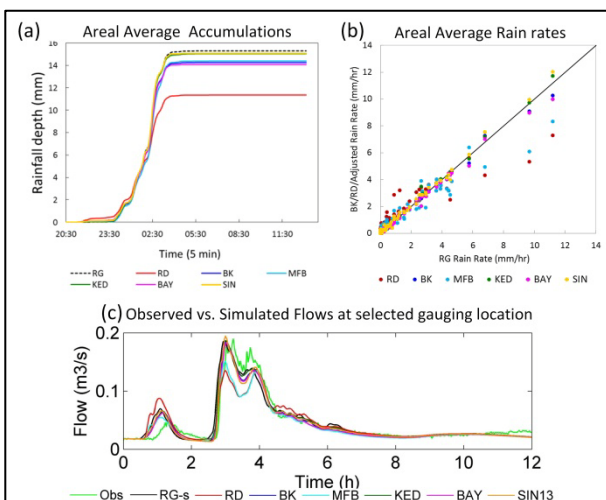


Fig 1: Quantitative performance of different merged QPEs – Storm 1: (a) Areal average total rainfall accumulations; (b) Areal average rain gauge vs. radar/ merged QPEs instantaneous rain rates; (c) Observed vs. simulated flows.
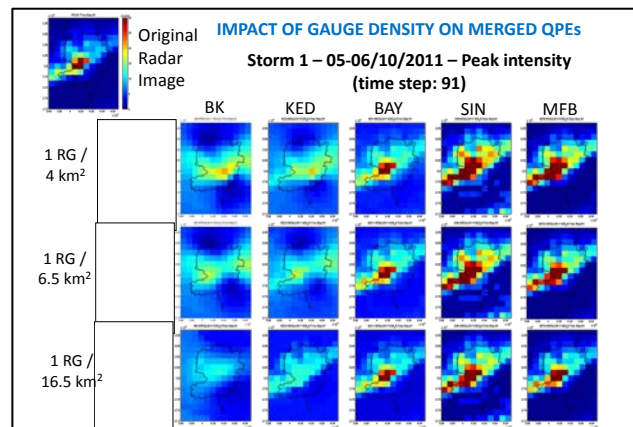


Fig 2: Impact of rain gauge density on the rain gauge interpolated (BK: block-kriging) and radar-rain gauge merged QPEs.

## References

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

Jewell, S. A. & Gaussiat, N. (2015). An assessment of kriging-based rain-gauge-radar merging techniques. Quarterly Journal of the Royal Meteorological Society.

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. Advances in Water Resources, 45 (0), 65-75.

Todini, E. (2001). A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. Hydrology and Earth System Sciences, 5 (2), 187-199.

Wang, L.-P., Ochoa-Rodríguez, S., Onof, C. & Willems, P. (2015). Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications. Hydrology and Earth System Sciences Discussions, 12 (2), 1855-1900.

Wang, L.-P., Ochoa-Rodríguez, S., Simoes, N., Onof, C. & Maksimović, Č. (2013). Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. Water Science & Technology, 68 (4), 737-747.

# The need for high resolution data to improve urban flood risk assessment

**J.A.E. ten Veldhuis[1*], S. Ochoa-Rodriguez[2], A. Gires[3], J. van Assel[4], A. Ichiba[3,5], S. Kroll[4], L. Wang[6], I. Tchiguirinskaia[3], D. Schertzer[3], P. Willems[6]**

[1]{Delft University of Technology, Delft, the Netherlands}

[2]{Imperial College London, London, United Kingdom}

[3]{Ecole des Ponts-ParisTech, Paris, France}

[4]{Aquafin NV, Aartselaar, Belgium}

[5]{Conseil Général du Val-de-Marne, France}

[6]{KU Leuven, Leuven, Belgium}

Correspondence to: J.A.E. ten Veldhuis (j.a.e.tenveldhuis@tudelft.nl)

## Abstract

Cities are particularly vulnerable to rainfall-generated floods that are typically characterised by their rapid onset and localised nature. This implies that precipitation and catchment information need to be available at high resolution to reliably predict hydrological response and potential flooding. On the contrary, urban areas constitute a major knowledge gap as most flood risk studies have concentrated on natural basins and records of rain gauges and water level gauges in cities are scarce. While increase in intense precipitation as a result of climate change is expected in many areas around the world, it is at present not possible to assess how this will affect urban pluvial flood risk. Collection of reliable, high resolution data in cities needs to start urgently to build up datasets in support of urban flood risk assessment and to enable detection of changes in flood risk whether these are induced by climate change, urbanisation or other future developments. This study shows how implementation of polarimetric X-band radar can contribute to filling the knowledge gap of flood risk quantification in cities.

## 1   Introduction

Cities are particularly sensitive to flooding induced by short-duration, high-intensity precipitation, due to their high degree of imperviousness, resulting in fast runoff processes

and lack of available water storage. Moreover, the high density of population and economic assets in urban areas results in high vulnerability to flooding. Ongoing urbanisation and urban densification further contribute to exacerbating flood vulnerability, thus increasing urban pluvial flood risk.

Based on climate models, increases in the frequency and intensity of heavy precipitation are projected for the 21st century in several regions of the world. It is "likely that the frequency of heavy precipitation will increase in the 21st century, particularly in the case of high latitudes and tropical regions and in winter in the northern mid-latitudes" (Kundzewicz, 2014). Increase in heavy precipitation could in turn be expected to contribute to increases in precipitation-generated flood risk; however, insufficient evidence is currently available on both flood frequency and magnitudes as well as on flood losses to assess climate-driven changes (Kundzewicz, 2014). Urban areas in particular represent a major knowledge gap, since most flood risk studies refer to "natural" basins and records of rain gauges and water level gauges in cities are scarce and essentially insufficient to represent the fine-scale urban hydro-meteorological variability (Schellart et al., 2012, Jensen and Pedersen, 2005).

The focus of this paper is on risks associated with urban, rainfall-generated flooding. It introduces new approaches and technologies to characterise high resolution temporal and spatial characteristics of rainfall, hydrological response and flood vulnerability in urban catchments. Information on these characteristics is crucial to be able to reliably quantify urban flood risk. This is in turn a first requirement to be able to build up the datasets required to assess potential changes in rainfall-generated flood risk in cities. The final challenge will be to identify drivers from the complex interactions of rainfall, urban development, concentration of asset value and development of man-made drainage infrastructure in cities.

**High resolution data at urban scales**

The spatial-temporal characteristics of urban catchments and stormwater drainage systems are generally small, often of the order of 1-10 km$^2$ and a few minutes, respectively (Arnbjerg-Nielsen et al., 2013; Ochoa-Rodriguez et al., submitted). Cities typically display high spatial variability in land-use, small catchment areas and a high degree of imperviousness. Stormwater drainage systems are predominantly man-made and consist of complex networks of channel and pipe networks. For the hazard component of flood risk assessment this implies that precipitation information needs to be available at high resolution to reliably predict hydrological response and potential flooding. With respect to flood vulnerability, cities

typically comprise a high concentration of assets (infrastructure, buildings), economic value and, most importantly, people. Unlike river and coastal flooding, potential locations of rainfall-generated flooding are not concentrated along line-elements like coasts of river banks, as localised storms can occur anywhere in a catchment. Moreover, artificial drainage networks can lead to redistribution of flows, away from natural drainage paths. Not only the position and capacity of drainage systems is important, also their condition needs to be known, as man-made drainage systems in cities (especially underground sewer pipes) are sensitive to blockage and flow disruption which can induce localised floods in areas that would not normally be susceptible to flooding (ten Veldhuis et al., 2009).

For the damage component of flood risk assessment, collection of damage data associated with urban pluvial flooding is complicated by the fact that in cities damage is borne by a wide range of individual property owners, industries as well as governmental authorities. Insurance databases constitute a valuable source of damage information, albeit difficult to access due to privacy issues and data quality (Spekkers et al., 2014). Kundzewicz at al. (2014) show that globally about 9% of flood damage in the period 2001-2011 was insured, which implies that dedicated methods for damage data collection need to be found to obtain comprehensive insight into urban pluvial flood damage. Reports of localised flooding and damage reports need to be collected in a structured way over long periods of time before changes in urban flood risk can be assessed and drivers of changes can be identified.

In summary, analysis of rainfall-generated flooding in cities requires high resolution data in space and time of precipitation intensities, catchment characteristics, hydrological response, vulnerability and historical flood damage. This will support quantification of current urban pluvial flood risk and will constitute a starting point for building up time series of high resolution rainfall intensities, occurrence of flood events and flood damage which will enable detection of changes to date and forecasting of future changes in urban pluvial flood risk.

The EU-funded RainGain project set out to collect reliable, high resolution precipitation data in four cities in North-West-Europe, based on innovative technology: state-of-the-art dual-polarimetric weather radars for retrieval or rainfall information. Ten pilot sites were selected in the four cities based on availability of recent topographic datasets and detailed hydrodynamic models representing the drainage networks and drainage area characteristics. Through collaboration with local authorities holding operational knowledge of the urban drainage networks and historical flood events, valuable information for flood risk assessment

was generated. In this paper, we will present findings of a recent study in which high resolution precipitation datasets derived from a dual-polarimetric X-band radar were used to examine hydrological response patterns in seven of the 10 pilot catchments. We will build upon these findings and ongoing work of the RainGain project to identify some critical requirements for data collection in support of urban flood risk analysis and detection of changes in flood risk characteristics. We will also outline some approaches and methodologies to meet such requirements.

## 2 Methods and dataset

### 2.1 Urban catchment characteristics

Seven urban catchments, located at each of the four RainGain partner countries, were adopted as pilot locations in this study. With the aim of facilitating inter-comparison of results, catchment areas of similar size (3-8 km$^2$) were selected for testing. The main characteristics of the selected pilot catchments are summarised in Table 1. Moreover, images of the boundaries and sewer layouts of all pilot catchments can be found in Figure 1. More detailed information on each of these catchments is available on the RainGain project website: http://www.raingain.eu/en/actualite/learn-more-about-ten-locations-where-raingain-solutions-will-be-implemented. As can be seen, the selected pilot catchments cover a wide range of morphological, topographic and land use conditions.

### 2.2 Precipitation dataset: dual-polarimetric X-band radar

One of the aims of the RainGain project is to obtain high resolution precipitation estimates in cities. To this end, four different radar-rain gauges configurations are set up for precipitation estimation in Leuven, London, Paris and Rotterdam (figure 2). Configurations vary from a single polarisation radar and a network of rain gauges for ground truthing in Leuven, providing rainfall estimates at 125x125m$^2$ and 1 minute resolution. In London, pilot sites are within coverage of 2 radars of the national C-band radar network, equipped and being upgraded to dual-polarisation, where super-resolution protocols are applied, i.e. by adjusting signal pulse length, to obtain high resolution precipitation estimates. In Paris and in Rotterdam, new dual polarisation X-band radars were installed, a pulse radar and a Frequency-Modulated Continuous Wave (FMCW) radar, respectively. All sites are equipped with a network of rain gauges; additionally, disdrometers are installed in Paris and Rotterdam.

While radar rainfall products are under development in the four pilot cities, high resolution data were obtained for this study from a polarimetric radar located in Cabauw, the Netherlands (Leijnse et al., 2010). Data were derived for nine storms in the period 2011-2014 and were used to conduct analyses of the space-time scales of storm cells and study hydrological modelling response at a range of space-time resolutions.

The estimated spatial and temporal characteristics of the nine storm events are summarised in table 2. For information on the method applied for estimating space-time scales of the storms we refer to Ochoa-Rodrigues et al. (submitted). Mean velocity of the nine storms varies from 6.4 m/s to 19.3 m/s (34 to 69 km/h) and storm ranges vary from 1700 to 4660 m. The combination of storm velocity and storm range, together with catchment dimensions, determines the time during which the storm core passes through a given catchment. For the storms and catchments considered in this study, this time varied between ~2-12 min. Moreover, based upon the estimated spatial and temporal characteristics of the storms, and making use of communication theory concepts (Shannon, 1948), the minimum required space-time resolutions for precipitation sampling were computed for the nine storms under consideration. This resulted in required precipitation sampling resolution varying between ~1 minute and ~6 minutes and spatial resolution varying from ~700m to ~2000m. This implies that the typical resolutions of rainfall estimates provided by national radar networks (i.e. 5 min, 1000 m) matches required space-time resolution for only four out of the nine storms under consideration. It should be noted that within the spatial scale of the storm core, high intensity storm cells can still be detected (see for instance figure 3). Given the small size of drainage areas in the urban catchments, down to below the 100 m scale, such cells can still be potential triggers for localised flooding, in cases where local drainage capacity is limited.

## 2.3   Hydrological response: sensitivity to space-time resolution

In order to investigate the sensitivity of urban drainage models to the spatial-temporal resolution of rainfall inputs, the high-resolution precipitation data for the nine (9) storm events, initially at 100 m and 1 min, were aggregated to a number of coarser temporal and spatial resolutions (up to 3000 m and 10 min) and were applied as input to the urban drainage models of the seven (7) pilot catchments. Results were analysed at different drainage areas of varying sizes (~ 1 ha to ~ 800 ha) within each pilot catchment. Some of the main findings of the hydrological response analysis are summarised in this paper; for an in-depth analysis please refer to Ochoa-Rodriquez et al. (submitted).

The results showed that hydrodynamic response behaviour in urban catchments is highly sensitive to combinations of temporal and spatial resolutions of rainfall input: resolution combinations that do not properly reflect storm dynamics lead to large deviations in hydrodynamic model outcomes. For the storms investigated in this study, hydrodynamic response behaviour was more sensitive to temporal than to spatial resolution. Temporal resolution coarsening beyond the estimated required resolution (between 2 and 6 min) led to under- and overestimations of flow peaks by up to 100% with respect to the original 100 m, 1 minute rainfall input. Similarly, it resulted in low explained variance (down to 20% explained variance, median value, at the level of the entire hydrograph) and flow underestimation at the level of the entire hydrograph (figure 4, illustration of results for 1 catchment, 1 storm). Spatial resolution coarsening led to underestimation of hydrographs for spatial scales between 500 m and 1 km for drainage areas of 1 to 100 ha. A special feature observed in the analysis is the strong interaction between the spatial and temporal resolution of rainfall estimates: the two resolutions must be consistent with each other to prevent loss of information from the higher resolution (more detail in Ochoa-Rodriguez et al., submitted).

## 3 Implications for urban pluvial flood risk analysis

Analysis of critical space-time scales based on high resolution (100 m in space, 1 minute in time) rainfall data obtained from dual-polarimetric X-band radar suggested that, for the nine investigated storms, sampling frequency for rainfall measurement should be at least of the order of ~700m to ~2000m in space and ~1 minute to ~6 minutes in time to capture most of the space-time variability. Still higher space-time resolutions are needed to fully reproduce rainfall variability. Urban hydrological response proved to be highly sensitive to space-time resolution of rainfall input: large deviations in flow peak and hydrograph volumes were found for coarser rainfall input resolutions, especially for small drainage basins of 1 to 100 ha, typical of urban neighbourhoods.

For flood risk analysis, this implies that rainfall as well as catchment data need to be available at the same high resolutions to be able to explain damage-generating processes, to predict and prevent flood risk. If flood risk analysis is conducted based on lower resolution data, large uncertainties occur in hydrological response prediction and in the assessment of expected flood frequencies and magnitude. Moreover, damage-generating mechanisms cannot be well

understood at the coarser level of data resolution (see for instance Spekkers et al, 2014), resulting in highly unreliable flood risk predictions.

Detecting changes in urban flood risk will not be possible at the current state of knowledge and data availability. Flood risk generating mechanisms, frequency, vulnerability and damage are too poorly understood to conduct reliable flood risk assessment even in the current situation. First, high resolution, reliable datasets need to be set up to properly understand damage generating mechanisms for urban rainfall-generated floods and to obtain time series of hazard (rainfall), vulnerability (catchment) and damage. Once such datasets have been built up, time series analyses can be conducted, aiming at detecting changes and their underlying mechanisms, whether climate change, urbanisation, ageing infrastructures or other.

Polarimetric X-band radar has proved to provide valuable, high resolution information in support of hydrological response analysis in urban areas, a first crucial step in urban flood risk assessment.

## Acknowledgements

## References

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Gregersen, I. B., ... & Nguyen, V. T. V. (2013). Impacts of climate change on rainfall extremes and urban drainage systems: a review. Water Science & Technology, 68(1), 16-28.

Jensen, N. E., & Pedersen, L. (2005). Spatial variability of rainfall: Variations within a single radar pixel. Atmospheric Research, 77(1), 269-277.

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., ... & Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. Hydrological Sciences Journal, 59(1), 1-28.

Leijnse, H., Uijlenhoet, R., van de Beek, C. Z., Overeem, A., Otto, T., Unal, C. M. H., Dufournet, Y., Russchenberg, H. W. J., Figueras i Ventura, J., Klein Baltink, H. & Holleman,

I. (2010). Precipitation Measurement at CESAR, the Netherlands. Journal of Hydrometeorology, 11 (6), 1322-1329

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., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., ten Veldhuis, M. (submitted). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation.

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. Advances in Water Resources, 45 (0), 65-75.

Spekkers, M. H., Kok, M., Clemens, F. H. L. R., & ten Veldhuis, J. A. E. (2014). Decision tree analysis of factors influencing rainfall-related building damage. Natural Hazards and Earth System Sciences Discussions, 2(4), 2263-2305.

ten Veldhuis, J. A., Clemens, F. H., & van Gelder, P. H. (2011). Quantitative fault tree analysis for urban water infrastructure flooding. Structure and Infrastructure Engineering, 7(11),809-821.

Table 1. Summary characteristics of pilot urban catchments

| | Cran-brook | Torquay | Morée-Sausset | Sucy-en-Brie | Herent | Ghent | Kralin-gen |
|---|---|---|---|---|---|---|---|
| Area [ha] | 865.2 | 570.03 | 560.4 | 269 | 511.5 | 649.33 | 670 |
| Catchment length and width [km]* | 6.1/1.4 | 5.4/1.1 | 5.3/1.1 | 4.0/0.7 | 8.2/0.6 | 4.7/1.4 | 2.1/3.2 |
| Catchment shape factor [-] | 0.23 | 0.2 | 0.2 | 0.17 | 0.08 | 0.29 | 1.49 |
| Slope [m/m]*** | 0.0093 | 0.0262 | 0.0029 | 0.0062 | 0.0083 | 0.0001 | 0.0003 |
| Main flow direction [deg] | 239 | 270 | 198 | 138 | 40 | 235 | 152 |
| Type of drainage system | Mostly separate, branched | Mostly combined, branched | Mostly separate, branched | Separate, branched | Mostly combined, branched | Mostly combined, branched | Mostly combined, looped |
| Is flow mainly driven by gravity? | Yes | Yes | Yes | Yes | Yes | Yes | No |
| Control elements | 3 storage lakes | 3 storage tanks, 1 pumping station | 2 storage tanks | 1 storage basin, 1 pumping station | 5 main CSO's with control | 15 pumping stations | 20 pumping stations |
| IMP (%)**** | 52% | 26% | 37% | 34% | 27% | 41% | 48% |
| Predominant land-use | R&C | R&C | R&C | R&C | R | R | R&C |
| Population density [pp/ha] | 47 | 60 | 70 | 95 | 20 | 24 | 154 |

*Length = Length of longest flow path (through sewers) to catchment outfall; Width = Catchment Area / Catchment Length;

**Shape factor = Width / Length (this parameter is lower for elongated catchments)

***Catchment slope = Difference in ground elevation between upstream most point and outlet / catchment length

****IMP: total proportion of impervious areas in relation to total catchment area

Table 2: Estimated spatial and temporal characteristics and required resolutions of the storm events under consideration

| Storm event ID | Spatial Range | Mean Velocity | Anisotropic Coefficient | Required Spatial Resolution | Required Temporal Resolution |
|---|---|---|---|---|---|
| | [m] | [m/s] | | [m] | [min] |
| E1 | 4056.69 | 9.76 | 0.38 | 1694.77 | 5.79 |
| E2 | 3524.76 | 9.91 | 0.38 | 1472.54 | 4.95 |
| E3 | 4655.10 | 14.04 | 0.55 | 1944.77 | 4.62 |
| E4 | 3218.91 | 11.71 | 0.34 | 1344.77 | 3.83 |
| E5 | 2061.98 | 14.11 | 0.59 | 861.43 | 2.03 |
| E6 | 3737.52 | 11.68 | 0.26 | 1561.43 | 4.46 |
| E7 | 1702.93 | 13.95 | 0.24 | 711.43 | 1.70 |
| E8 | 3644.43 | 18.40 | 0.36 | 1522.54 | 2.76 |
| E9 | 2354.53 | 16.97 | 0.08 | 983.66 | 1.93 |
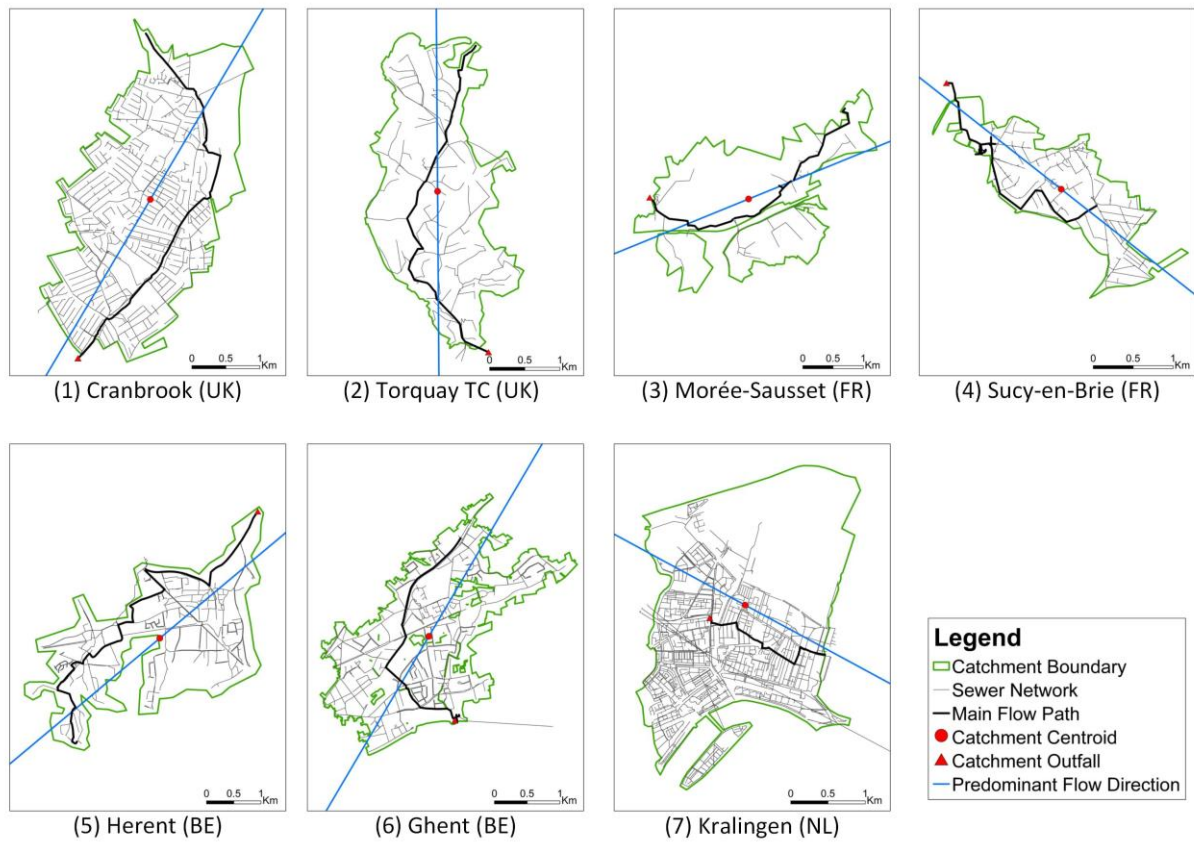
Figure 1. Catchment boundary and sewer layout for the pilot urban catchments



Figure 2. Radars implemented at the pilots sites of RainGain (from left to right): X-band single pol radar implemented in Leuven, Chenies C-band radar of the UK national network, dual-pol X-band radar installed in Paris, dual-pol X-band radar to be installed in Rotterdam.
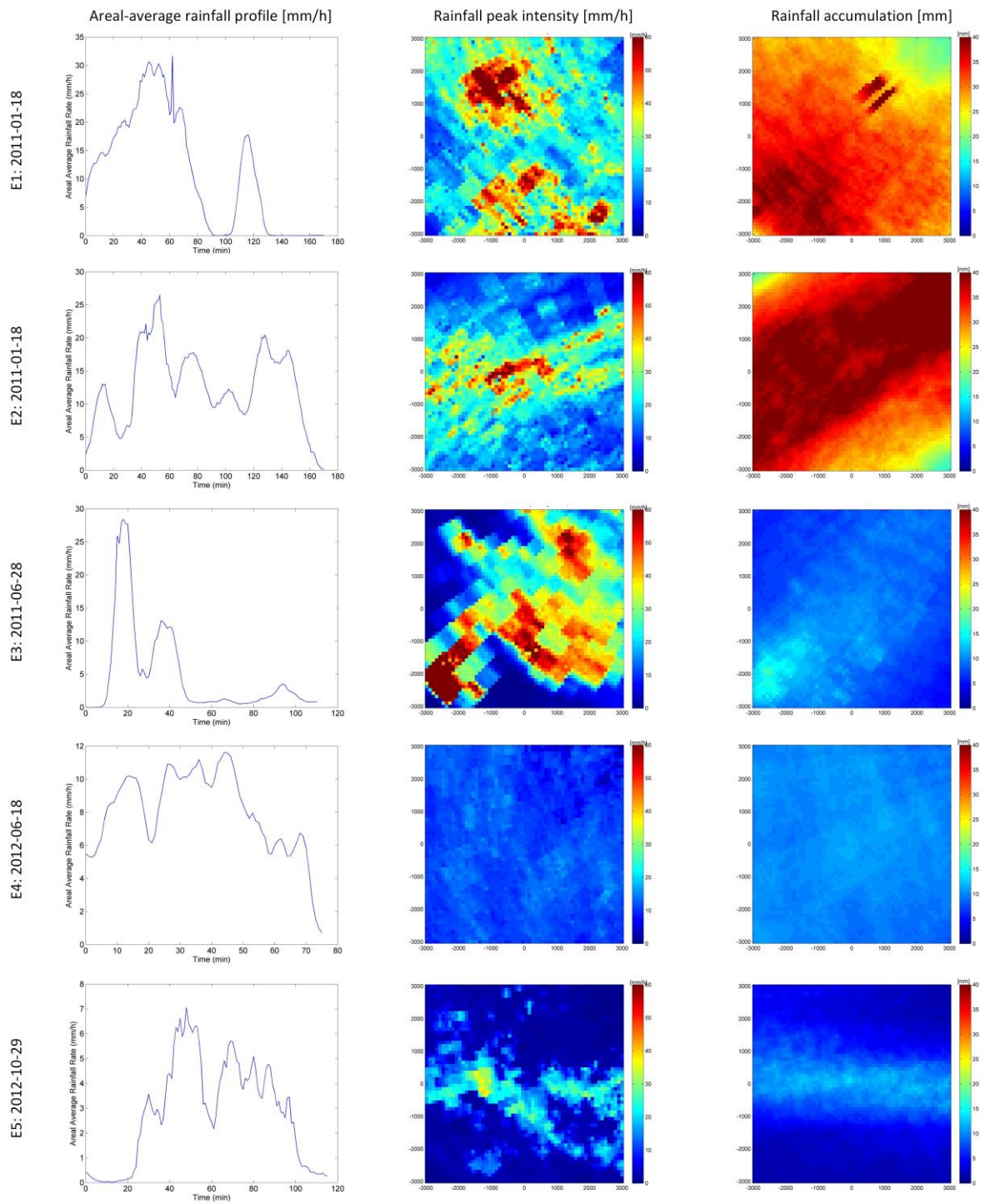
Figure 3. Areal average storm profile (left column), snapshot image during the peak intensity period of the storm (middle column) and total event accumulations for the storm events under consideration.
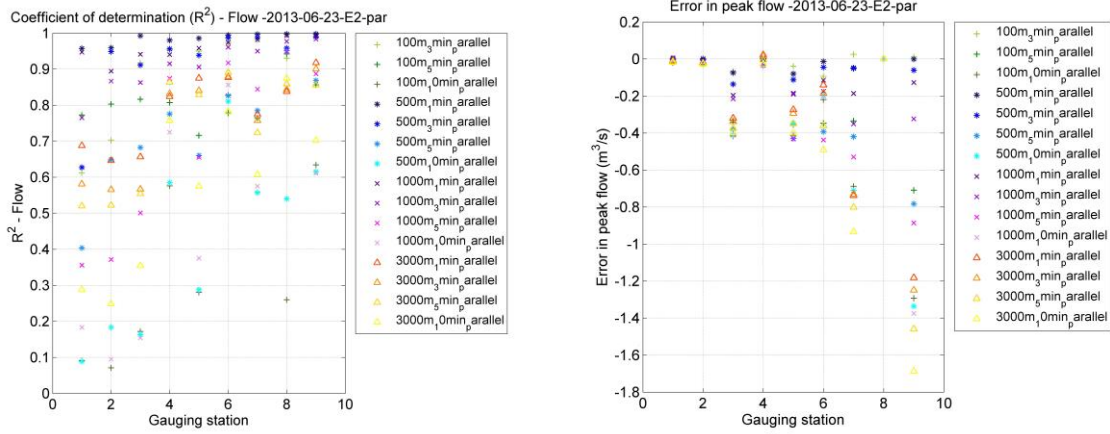
Figure 4. Example of hydrological response characteristics for 1 storm event, 1 catchment, 16 resolution combinations. Coefficient of determination and absolute error in peak flow are plotted for nine gauging stations corresponding with drainage area size increasing from ~1 ha to ~ 800 ha. The original input resolution of 100 m, 1 minute resolution is taken as a reference.

# High Resolution Radar Rainfall for Urban Pluvial Flood Control
### Lessons Learnt From Ten Pilots in North-West Europe Within the RainGain Project

Marie-claire TEN VELDHUIS[1*], Susana OCHOA-RODRIGUEZ[2], Guenda BRUNI[1],
Auguste GIRES[3], Johan VAN ASSEL[4], Abdellah ICHIBA[3], Stefan KROLL[4], LiPen
WANG[5], Ioulia TCHIGUIRINSKAIA[3], Agathe GIANGOLA-MURZYN[3], Julien
RICHARD[3], Daniel SCHERTZER[3], Patrick WILLEMS[5]

[1]*Delft University of Technology, Delft, The Netherlands*
[2]*Imperial College London, London, United Kingdom*
[3]*Ecole des Ponts-Paristech, Paris, France*
[4]*Aquafin NV, Antwerp, Belgium*
[5]*KU Leuven, Leuven, Belgium*
*Corresponding author*
*Email: j.a.e.Tenveldhuis@tudelft.nl*

## ABSTRACT

Precipitation and catchment information need to be available at high resolution to reliably predict hydrological response and potential flooding in urban catchments. Due to recent advances in weather radar technology and DTM availability for urban flood modelling, the question arises whether these are sufficient to provide reliable predictions for urban pluvial flood control. The RainGain project (EU-Interreg IVB NWE) brings together radar technologists and hydrologists to explore a variety of rainfall sensors, rainfall data processing techniques and hydrodynamic models for the purpose of fine-scale prediction of urban hydrodynamic response. High resolution rainfall and hydrodynamic modelling techniques were implemented at ten different pilot locations under real-life conditions. In this article, the pilot locations, configurations of rainfall sensors (including X-Band and C-Band radars, rain gauges and disdrometers) and modelling approaches used in the RainGain project were introduced. Initial results  presented the hydrodynamic modelling using high resolution precipitation inputs from dual-polarisation X-band radar, followed by a discussion of differences in hydrodynamic response behaviour between the pilots.

## KEYWORDS
Radar rainfall, urban hydrology, urban flood modelling

## INTRODUCTION
Urban catchments are characterised by high spatial variability, fast runoff processes and short response times. This implies that precipitation and catchment information needs to be available at high resolution to reliably predict urban hydrological processes (Aronica & Cannarozzo, 2000; Einfalt, 2005; Segond et al., 2007). Several studies have shown that despite recent advances in the use of weather radar, the resolution of the currently available rainfall estimates (typically 1 x 1 $km^2$ in space and 5 min in time)  may still be too coarse to match the spatial-temporal scales of urban catchments (Fabry et al., 1994; Gires et al., 2012a). In this regard and in the light of recent developments, new questions arise, such as: what rainfall resolution is needed for different urban applications? How do rainfall data resolution and data reliability interrelate? What reliability can be delivered by different configurations of radar and rain gauges in cities? What modelling approaches are best suited to obtain reliable results in terms

of water level and flood predictions? How sensitive are hydrodynamic models to rainfall spatial variability? What is the influence of catchment variability? With the aim of answering some of these questions, the RainGain project (EU-Interreg IVB NWE) has set to explore the use of a variety of rainfall sensors (including X-Band and C-Band radars, rain gauges and disdrometers), to develop and test a number of rainfall data processing techniques and to test the response of hydrodynamic models with different characteristics to varying rainfall inputs. In addition, the needs of the stakeholders involved in flood risk management are assessed and ways of using high resolution rainfall and hydrodynamic model outputs for improving flood risk management are explored.

In this paper, the main characteristics of the 10 pilot locations adopted within the RainGain project are presented. Initial experiences and results are presented with respect to implementation of high resolution radars in urban settings and to application of resolution precipitation estimation in hydrodynamic modelling at different catchments.

## EXPERIMENTAL SITES – 10 PILOT LOCATIONS

Ten experimental sites have been implemented within the RainGain project; pilot sites have been selected so as to represent a range of varying urban catchment characteristics and different types of pluvial flooding problems. Characteristics of the pilot sites are summarised in table 1. Most of the sites are highly urbanised and vary in size from about 1.4 to 34 km². Half of the sites are fairly flat, the other half are characterised by a combination of plateaus and steep slopes along river banks. Some of the sites are located in urban polders, without natural drainage outlets; in these areas stormwater needs to be locally stored and evacuated through pumps. Applied model software includes semi-distributed and fully distributed modelling approaches, one-dimensional and two-dimensional overland flow modules.

**Table 1.** General characteristics of pilot urban catchments

| Pilot site | Catchment size [km²] | General catchment characteristics | General characteristics of drainage system | Modelling approach and software |
|---|---|---|---|---|
| Cranbrook catchment (London Borough of Redbridge) | 8.65 | Highly urbanised, mildly sloping, coincidental fluvial and pluvial flooding | Mostly separate, main brook has been culverted | Semi distributed, dual drainage (both 1D-1D and 1D-2D models; rainfall applied through subcatchments), InfoWorks CS-2D |
| Purley Area (London Borough of Croydon) | 6.5 | Highly urbanised, great density of receptors, slopes drain to natural depression | Mostly separate, combination of natural drainage channels, culverted river and sewers | Semi distributed, sewer system only, simplified modelling of exceedance flow .InfoWorks CS-2D |
| Torquay Town Centre (Devon Borough of Torbay) | 14.5 | Coastal city, steep slopes drain to natural depression, flooding worsened by high tides. | Combined sewer system; two CSO's, discharging into Torquay Harbour under storm conditions. | Semi distributed, 1D 2D dual drainage (with rainfall applied through subcatchments). InfoWorks CS-2D |
| Morée Sausset, incl. Kodak subcatchment (Seine-Saint-Denis, | 34 Kodak: 1.44 | Highly urbanised, rather flat. Several retention basins for flood control. | Mostly separate, main brook has been culverted, several storm water retention | Semi-distributed, sewer system only, simplified exceedance flow (Canoe) Kodak: Fully distributed, |

| | | | | |
|---|---|---|---|---|
| Paris region) | | | basins | 1D 2D dual drainage (rainfall applied directly on 2D model of surface) Multi-Hydro |
| Jouy en Josas (Seine-Saint-Denis, Paris region) | 2.5 | Combination of residential and green areas. River bank, steep slopes (100m elevation difference) and plateau. | Mostly separate, several storm water retention basins | Fully distributed, 1D-2D dual drainage: Multi-Hydro |
| Sucy en Brie (Val de Marne, Paris region) | 2.69 | Residential and industrial use. River bank, steep slopes (32 m elevation difference) and plateau. | Mostly separate, new retention basin (interest on RT control of it) | Current semi-distributed (Canoe). New: fully distributed, 1D 2D dual drainage: Multi-Hydro |
| Herent (Leuven, northern part) | 4.75 | Densely built village centres and rural areas; fairly flat. | Mostly combined sewer system, CSOs discharging to two local rivers running through the city | Current semi- distributed. New: semi distributed, 1D 2D dual-drainage ( rainfall applied through subcatchments). InfoWorks ICM |
| Kralingen- (Rotterdam) | 6.70 | Residential and industrial use, flat polder area | Combined, looped system; CSOs discharging to local channels, sewer pumps evacuate water from urban polder | Semi-distributed, simplified modelling of exceedance flow (Sobek Urban) |
| Spaanse Polder (Rotterdam) | 1.9 | Industrial area, densely urbanised, flat polder area | Combined, looped; CSOs discharging to local channels, sewer pumps evacuate water from urban polder | Semi-distributed, simplified modelling of exceedance flow (Sobek Urban) |
| Centrum district (Rotterdam) | 3.7 | Residential and commercial area, 2 urban parks, flat polder area | Combined, looped; CSOs discharging to local channels, sewer pumps evacuate water from urban polder | Semi-distributed, simplified modelling of exceedance flow (Sobek Urban) |

## HIGH RESOLUTION PRECIPITATION DATASETS

Four different radar-rain gauges configurations are used for precipitation estimation in Leuven, London, Paris and Rotterdam (figure 1). In Leuven, a single polarisation radar has been operational since 2008 providing rainfall estimates at $125 \times 125 m^2$ and 1 minute resolution. Original data processing algorithms are adjusted under the project, in order to improve the quality of radar rainfall estimates. Pilot sites in London are within coverage of 2 radars of the national C-band radar network, equipped and being upgraded to dual-polarisation. Experiments are being conducted for improving resolution of the radar rainfall estimates by adjusting signal pulse length and shortening the repetition cycle. In addition, a short testing of a single polarisation X-band radar was carried out in London between May and October 2014. In Paris and in Rotterdam, new, dual polarisation X-band radars are installed, a pulse radar and a Frequency-Modulated Continuous Wave (FMCW) radar

respectively. All sites are equipped with a network of rain gauges; additionally, disdrometers are installed in Paris and Rotterdam.
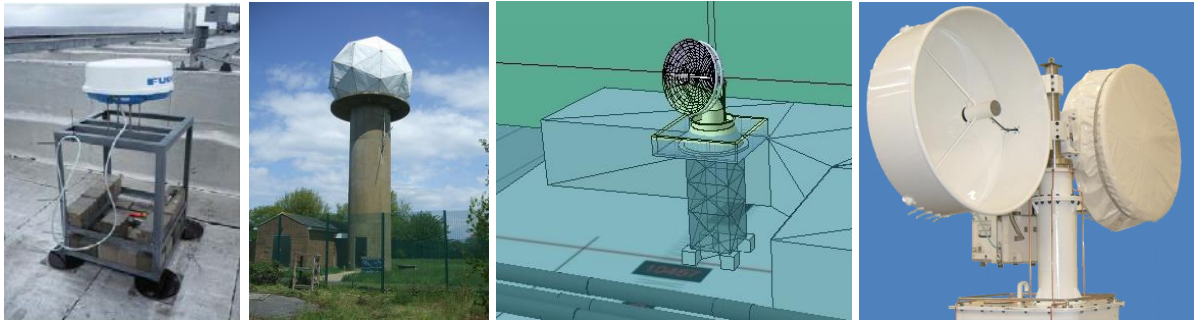


**Figure 1**. Radar implemented at the pilots sites of RainGain (from left to right): X-band single pol radar implemented in Leuven, Chenies C-band radar of the UK national network, impression of dual-pol X-band radar under construction in Paris, dual-pol X-band radar to be installed in Rotterdam.

**Implementation of radar in densely urbanised environments, experiences**
Through the installation of X-band radars at heart of the highly urbanised RainGain pilot locations, many lessons have been learned. Weather radars used for high resolution precipitation estimation are preferably installed within a city area, above the urban canopy. This generally means installation on existing high-rise, in agreement with constraints set by building owner, architect, signal emission standards and other radar applications, especially near airports. Clutter correction is especially important in urban areas due to the relatively frequent presence of objects and other signals compared to a rural setting.
Radar signal correction for single polarisation radar to obtain quantitative precipitation estimates has proven complicated and the added value compared to rain gauge networks has found to be small in several cases (e.g. Goormans and Willems, 2013; Shrestha et al., 2013; Ochoa-Rodriguez et al. 2014). Additional Doppler and dual-polarisation measurements provide valuable information to improve reliability of precipitation estimates (Van de Beek et al., 2010; Otto and Russchenberg, 2013). Another important aspect that the project is investigating is the effect of wind drift on rainfall patterns. High resolution precipitation estimates are more sensitive to this effect, which plays an important role in urban areas due to their highly variable microclimate induced by urban structures.

**Rainfall data downscaling**
The availability of rainfall data at different spatio-temporal resolutions in the RainGain project provide the opportunity to compare characteristics of downscaled rainfall data from C-band weather radar networks to high resolution rainfall data from X-band radar. One of the downscaling processes implemented within the RainGain project relies on Universal Multifractals which have been extensively used to characterize and simulate geophysical fields extremely variable over wide range of scales such as rainfall (see Schertzer and Lovejoy 2011 for a recent review). In this framework rainfall is expected to be generated through a scale invariant cascade process. This framework is very convenient for downscaling (Biaou et al., 2003), which can be done by first assessing the relevant features of the underlying cascade process on the available range of scales and second continuing the cascade process beyond the observation scale. See Gires et al. (2014) for a validation with networks if point measurement devices deployed over 1 km$^2$ areas and Gires et al. (2012) for applications in urban hydrology.

## HIGH RESOLUTION MODELLING APPROACHES

Initial results of modelling studies conducted at the pilot sites in the RainGain projects, are summarised in this paper. Modelling results of rainfall input from X-band radar are presented for different pilot sites as well as results of a comparison between fully and semi-distributed approaches. For more details on modelling results, the authors refer to relevant papers.

The modelling approaches adopted at each pilot site are as summarised in table 1. Semi-distributed models have been current practice at most locations. Semi-distributed one-dimensional sewer and two-dimensional overland flow models are tested at 4 pilot sites. Two types of overland flow models are tested; a fast, one-dimensional model for real-time prediction and a detailed, two-dimensional model aiming at accurate water level predictions. A fully distributed model, Multi-Hydro, is being tested at 3, potentially 4 sites. This model is under development at Ecole des Ponts ParisTech, see also Giangola et al., 2012). The model includes a 2-dimensional model representing surface runoff, infiltration and overland flow, as well as a one-dimensional sewer model which interacts with the surface model through connecting elements such as manholes or gullies. Fully distributed hydrologic models are based on a gridded input structure that can be directly adjusted to the spatial resolution of rainfall input. In semi-distributed models, rainfall input values are routed through subcatchments of varying size and shape, with a lumped representation of hydrological run-off processes.

### High resolution rainfall from X-band radar: hydrodynamic modelling results at four pilot catchments

Two storm events, one convective and one stratiform, measured by a polarimetric X-band radar located in Cabauw (The Netherlands) were used as input into semi-distributed models at four pilot locations of similar size (between 5 and 8 km2; more catchments characteristics in table 2), the Cranbrook catchment (UK), the Herent catchment (Belgium), the Morée Sausset catchment (France) and the Kralingen District (The Netherlands). Storm events were applied in such a way that: (1) the centroid of the selected rainfall area coincides with the centroid of each catchment, and (2) storm direction is approximately perpendicular to the main flow direction at each catchment (in order to avoid variations in response due to differences in relative storm/flow direction (Singh, 1997)). For each of the model runs the simulated flow and water depth time series at the downstream end of three pipes located in the upstream, mid-stream and downstream sections of the catchments were selected for analysis (see table 3). The looped nature of the Dutch catchment and the fact that flows may change direction throughout a storm event make it difficult to determine an exact area drained by a given pipe.

**Table 2.** Summary catchment characteristics of 4 pilot catchments used for high resolution hydrodynamic modelling

| Pilot site | Catchment size [km²] | Catchment length* and width** [km] | Catchment shape factor*** [-] | Catchment slope**** [m/m] | Imperviousness (%) |
|---|---|---|---|---|---|
| Cranbrook, UK | 8.65 | 6.10/1.42 | 0.23 | 0.0093 | 66 |
| Morée-Sausset, FR | 5.60 | 5.28/1.06 | 0.20 | 0.0029 | 37 |
| Herent, BE | 4.75 | 8.16/0.58 | 0.07 | 0.0220 | 18 |

| Kralingen, NL | 6.70 | 2.12/3.16 | 1.49 | 0.0003 | 48 |

*Length of longest flow path (through sewers) to catchment outfall;
**Width = Catchment Area / Catchment Length;
****Shape factor = Width / Length (this parameter is lower for elongated catchments)
****Catchment slope=Difference in ground elevation between upstream most point and outlet / catchment length

Figure 2 shows response hydrographs and depth time series for the two storm events, at the upstream pipes selected for analysis at each pilot catchment. The results show that the catchments respond quite differently to the convective storm event precipitation. The Cranbrook and Moree-Sausset catchments' hydrographs have a well-defined single response peak, while the Kralingen hydrograph has multiple peaks and the Herent hydrograph has a quick response peak followed by very slow increase and decrease of the flow. The atypical response behaviour of the Herent and Kralingen catchments can be explained by their specific features: the Herent catchment is equipped with a throttle device in the main sewer transport line to maximise in-sewer storage. This strongly delays the flow upstream and smooths the flow peak. The Kralingen catchment is located in a polder area where in the absence of natural flow directions, sewer networks tend to be strongly looped. As a result, the overall behaviour of the catchments is determined by a filling process of in-sewer storage, as evidenced by a fast rise in water depth leading to surcharged pipes. During the filling process, flow directions can change, as flow first moves towards a pumping station, then, once pumping capacity is exceeded, moves towards combined sewer overflows. Hydrological response of the four catchments shows similar behaviour for the stratiform storm event (not shown here). Response characteristics were also investigated for different rainfall spatial resolutions (100m and 1000m), for a discussion of these results we refer to ten Veldhuis et al. (2014).
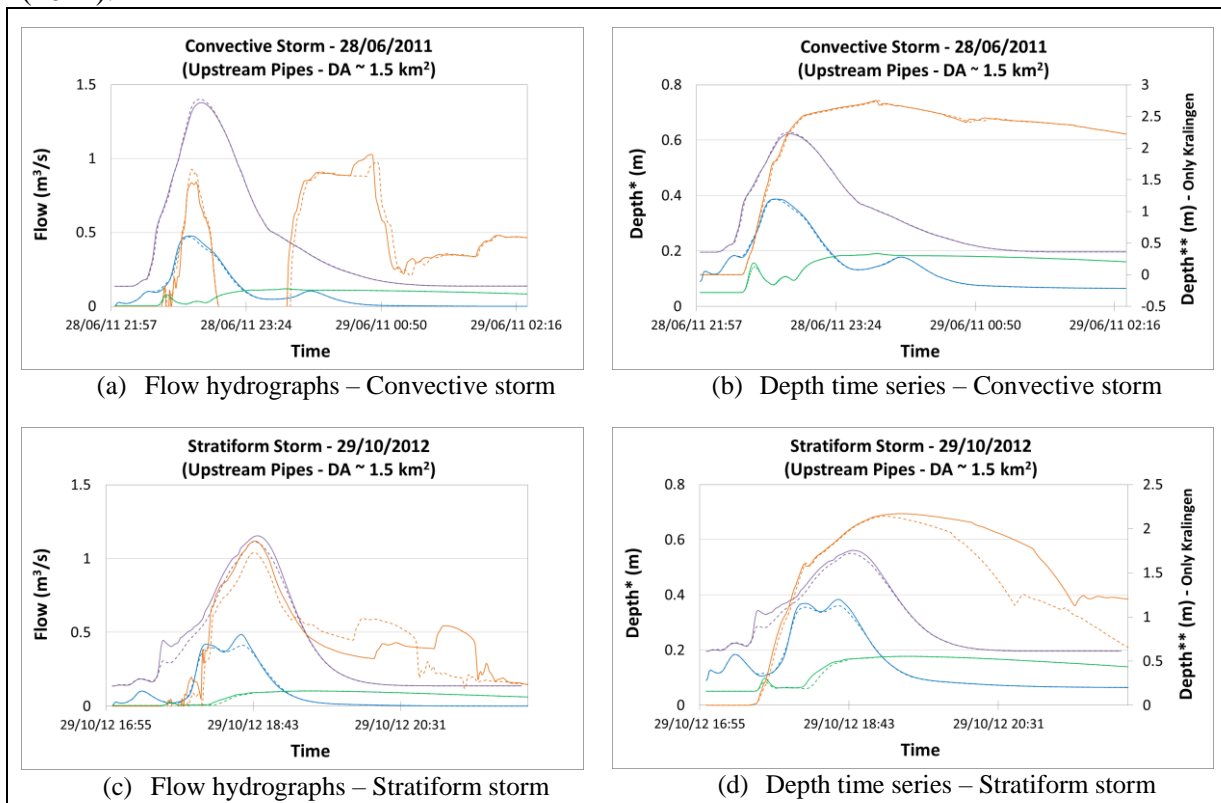


(a)  Flow hydrographs – Convective storm

(b)  Depth time series – Convective storm

(c)  Flow hydrographs – Stratiform storm

(d)  Depth time series – Stratiform storm

| ----Cranbrook 1000 m | ——Cranbrook 100 m | ----Moree-Sausset 1000 m | ——Moree-Sausset 100 m |
| ----Herent 1000 m | ——Herent 100 m | ——Kralingent 100 m | ----Kralingen 1000 m |

**Figure 2:** Response hydrographs and water depths at the downstream end of the upstream pipes selected for analysis at each pilot location (with drainage area (DA) ∼ 1.5 km$^2$). The solid lines correspond to the 100 m resolution outputs and the dashed lines to the 1000 m ones. * Water depth scale used for the depths observed in the Cranbrook (UK), Morée-Sausset (FR) and Herent (BE) pilot locations; **Water depth scale used for the depths observed in the Kralingen (NL) pilot location. In order to avoid distortion, a different y-axis was used for the water depths observed in Kralingen, as these were significantly higher than the ones observed at other locations.

Table 3 provides a summary of the measures which characterise the overall hydrological/hydraulic response of the catchments to rainfall. The results show that characteristic total flow volumes and peak values vary strongly between pilot sites. These variations are mainly explained by different settings in the rainfall-runoff model, especially runoff coefficients applied for impervious areas have an important influence.

**Table 3:** Response variables of each pilot catchment for each storm event. Characteristic runoff volume (total volume / drainage area) and characteristic peak flow (peak flow / drainage area) values are provided for the three pipe locations selected at each pilot catchment (Upstream/Mid-stream/Downstream)

| Pilot site | Model location* | Drainage area [km²] | Convective Storm – 28/06/11 | | | Stratiform Storm – 29/10/12 | | |
|---|---|---|---|---|---|---|---|---|
| | | | Vchar [m3/m2] | Qchar [m3/m2/s] | Tc [min] | Vchar [m3/m2] | Qchar [m3/m2/s] | Tc [min] |
| Cranbrook, UK | US | 1.65 | 0.86 | 0.29 | 45 | 0.017 | 0.29 | 49 |
| | MS | 3.24 | 0.89 | 0.27 | 45 | 0.015 | 0.21 | 49 |
| | DS | 5.67 | 0.91 | 0.25 | 45 | 0.013 | 0.17 | 49 |
| Morée-Sausset, FR | US | 1.99 | 3.55 | 1.4 | 48 | 3.5 | 0.6 | 52 |
| | MS | 3.83 | 3.88 | 3.0 | 48 | 3.5 | 0.6 | 52 |
| | DS | 5.60 | 3.59 | 3.7 | 48 | 2.8 | 0.5 | 52 |
| Herent, BE | US | 1.51 | 1.19 | 0.08 | 307 | 1.0 | 0.07 | 292 |
| | MS | 3.80 | 1.36 | 0.04 | 307 | 1.4 | 0.04 | 292 |
| | DS | 4.75 | 1.31 | 0.1 | 307 | 1.1 | 0.06 | 292 |
| Kralingen, NL | US | 1.30 | 7.05 | 0.79 | 213 | 0.11 | 0.86 | 169 |
| | MD | 3.10 | 6.71 | 0.76 | 213 | 0.08 | 0.52 | 169 |

**Semi-distributed versus fully distributed modelling: sensitivity to small-scale rainfall variability**

The uncertainty associated with small scale rainfall variability on urban catchments was assessed through the analysis of the sensitivity to rainfall resolution of hydrologic/hydraulic models. Two models were tested on the same 1.44 km2 Kodak catchment (see Table 1); the fully distributed Multi-Hydro model (grid with 10 m pixels) (Giangola et al. 2012) and the semi-distributed Canoe model (sub-catchments with size ranging from 4 to 16 ha) (Allison et al. 2005) . Only a brief summary of this study is reported here, and more details can be found in Gires et al. (2013). The methodology implemented consists in first generating an ensemble of downscaled rainfall fields with the help of discrete Universal. The raw data is the available Météo-France radar mosaic whose resolution is 1 km in space and 5 min in time, and the final resolution is 12.3 m and 18.75 s for the Multi-Hydro model and 111 m and 1.25 min for the Canoe model (given the size of the sub-catchments it was not relevant to further downscale

the data). Then each realisation of the downscaled rainfall field is inputted into the models. Finally the variability among the obtained hydrographs is analysed. To achieve this for each time step the 95, 75, 25 and 5% quantile are estimated. This enables to compute the envelop curves (Q0.1, Q0.25 Q0.75 and Q0.9) corresponding to their temporal evolution. Figure 3 displays these curves along with Qradar (flow simulated with raw radar data) at the outlet of the catchment for the February 2009 event (total depth 8.3 mm). The observed uncertainty reflects a significant impact of small scale rainfall variability on simulated discharge. The uncertainty increases with upstream conduits. Furthermore it appears that the uncertainty revealed by the fully distributed model is much greater. It means the semi-distributed model would not be able to fully benefit from improved rainfall data.
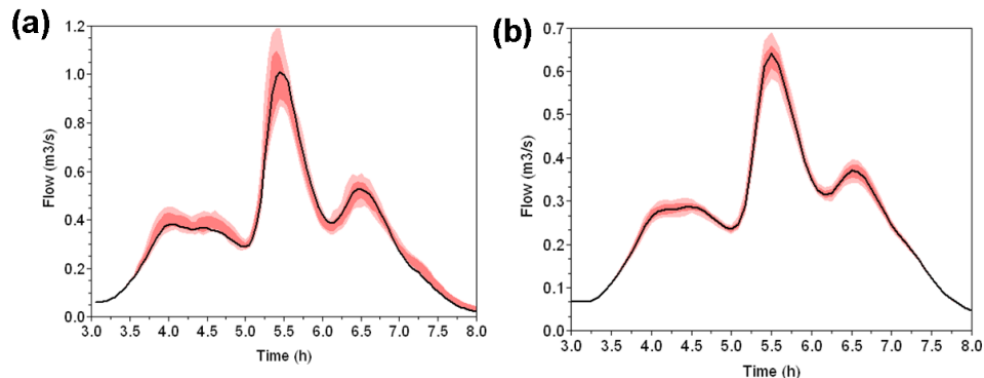


**Figure 3**. Simulated flow with the raw radar data (black), Q0.25 and Q0.75 (dark colour), Q0.1 and Q0.9 (light colour) for the outlet of the Kodak catchment. (a) Multi-Hydro 10 m, 2009 event; (b) 1D model, 2009 event; (adapted from Gires et al., 2013)

## CONCLUSIONS
These first results suggest that model settings, catchment and drainage infrastructure characteristics have a strong influence on hydrological response. Differences in catchment slope and drainage infrastructures have shown to result in entirely different response behaviors. Also, semi-distributed models seem not to be able to fully benefit from high resolution rainfall input data. Further studies into the impact of rainfall input resolution in relation to catchment characteristics, hydrological input data and model features will be conducted to gain more insights into these interactions.

## ACKNOWLEDGEMENT

## REFERENCES
Allison, 2005. Manuel d'utilisation de canoe, Conception et évaluation de réseaux d'assainissement, Simulation de pluies, des écoulements et de qualité des eaux. INSA Lyon, SOGREAH Consultants, 476 p. Available from: www.canoe-hydro.com [Accessed 7 September 2012].

Aronica, G. & Cannarozzo, M. (2000). Studying the hydrological response of urban catchments using a semi-distributed linear non-linear model. Journal of Hydrology, 238 (1-2), 35-43.

Biaou, A., Hubert, P., Schertzer, D., Tchiguirinskaia, I., Bendjoudi, H., 2003. Fractals, multifractals et prévision des précipitations. Sud Sciences et Technologies 10, 10-15.

Einfalt, T. (2005). A hydrologists' guide to radar use in various applications. In 10th International Conference on Urban Drainage, Copenhagen, Denmark.

Fabry, F., Bellon, A., Duncan, M. R. & Austin, G. L. (1994). High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. Journal of Hydrology, 161, 415-428.

Giangola-Murzyn, A., Gires, A., Hoang, C.T., Tchiguirinskaia, I. and Schertzer, D., 2012. Multi-Hydro: physically based modelling to assess flood resilience across  scales, case studies in Paris region. Proceedings of 10th Urban Dranaige conference, Belgrade 3-7 Sept.  2012, Serbia.

Gires, A., Onof, C., Maksimović, Č., Schertzer, D., Tchiguirinskaia, I. & Simoes, N. (2012a). Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. Journal of Hydrology, 442, 117-128.

Gires, A., Tchiguirinskaia, I., Schertzer, D. & Lovejoy, S. (2012b). Multifractal analysis of an urban hydrological model on a Seine-Saint-Denis study case. Urban Water Journal.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., Berne, A., Lovejoy, S., 2014. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. Atmospheric Research, 138(0), 125-138

Gires A., Giangola-Murzyn A., Abbes J.B., Tchinguirinskaia I., Schertzer D., Lovejoy.S. (2014)  Impacts of small scale rainfall variability in urban areas: a case study with 1D and 1D/2D hydrological models in a multifractal framework. Urban water journal, Under review.

Goormans, T., & Willems, P. (2012). Using Local Weather Radar Data for Sewer System Modeling: Case Study in Flanders, Belgium. Journal of Hydrologic Engineering, 18(2), 269-278.

Ochoa-Rodriguez et al. (2014). X-band radar monitoring campaign in Central London: an experiment to assess and improve the applicability of a low-cost portable X-band radar. Int. Symposium on Weather Radar and Hydrology, Washington, 2014.

Shrestha, N. K., Goormans, T., and Willems, P. (2013). Evaluating the accuracy of C-and X-band weather radars and their application for stream flow simulation. Journal of Hydroinformatics, 15(4).

Schertzer, D. and Lovejoy, S., 2011. Multifractals, Generalized Scale Invariance and Complexity in Geophysics. International Journal of Bifurcation and Chaos, 21(12): 3417–3456.

Segond, M. L., Neokleous, N., Makropoulos, C., Onof, C. & Maksimović, Č. (2007). Simulation and spatio-temporal disaggregation of multi-site rainfall data for urban drainage applications. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 52 (5), 917-935.

Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrological Processes, 11 (12), 1649-1669.

Ten Veldhuis, J.A.E., Ochoa-Rodriguez, S., Bruni, G., Gires, A., Van Assel, J., Wang, L., Reinoso Rodinel, R., Kroll, S., Schertzer, D., Onof, C., Willems, P. (2014). Weather radar for urban hydrological applications: lessons learnt and research needs identified from 4 pilot catchments in North-West Europe. Int. Symposium on Weather Radar and Hydrology, Washington, 2014.

# The need for high resolution precipitation data to improve urban drainage modelling

J.A.E. ten Veldhuis[1]*, S. Ochoa-Rodriguez[2], A. Gires[3], J. van Assel[4], A. Ichiba[3,5], S. Kroll[4], L. Wang[6], I. Tchiguirinskaia[3], D. Schertzer[3], P. Willems[6]

[1]Department of Watermanagement, Delft University of Technology, Delft, the Netherlands
 (Email: *j.a.e.tenveldhuis@tudelft.nl*)
[2] Imperial College London, London, United Kingdom
[3] Ecole des Ponts-ParisTech, Paris, France
[4] Aquafin NV, Aartselaar, Belgium
[5] Conseil Général du Val-de-Marne, France
[6] KU Leuven, Leuven,

**Abstract**
In this study high resolution precipitation data are used, derived from polarimetric X-band radar at 100 m, 1 min resolution. The data are used to study the impact of different space-time resolutions of rainfall input on urban hydrodynamic modelling response for 9 storms, in 7 urban catchments. The results show that hydrodynamic response behaviour was highly sensitive to variations in rainfall space-time resolution, more strongly so for changes in temporal than in spatial resolution. Under- and overestimations of flow peaks amounted to up to 100% with respect to the original 100 m, 1 minute rainfall input.

**Keywords**
Rainfall space-time resolution; Radar rainfall; Urban hydrology; Urban hydrological response modelling;

## INTRODUCTION
Cities are particularly sensitive to flooding induced by short-duration, high-intensity precipitation, due to their high degree of imperviousness, resulting in fast runoff processes and lack of available water storage. Moreover, the high density of population and economic assets in urban areas results in high vulnerability to flooding. Based on climate models, increases in the frequency and intensity of heavy precipitation are projected for the 21st century in several regions of the world (Kundzewicz, 2014).

### High resolution data at urban scales
The spatial-temporal characteristics of urban catchments and stormwater drainage systems are generally small, often of the order of 1-10 km2 and a few minutes, respectively (Arnbjerg-Nielsen et al., 2013; Ochoa-Rodriguez et al., submitted). Cities typically display high spatial variability in land-use, small catchment areas and a high degree of imperviousness. Stormwater drainage systems are predominantly man-made and consist of complex networks of channel and pipe networks. This implies that precipitation information needs to be available at high resolution to reliably predict hydrological response and potential flooding.

## METHODS AND DATASET

### Urban catchment characteristics
Seven urban catchments, located at each of the four partner countries of the EU-funded RainGain project, were adopted as pilot locations in this study. With the aim of facilitating inter-comparison of results, catchment areas of similar size (3-8 km$^2$) were selected for testing. The main characteristics of the selected pilot catchments are summarised in Table 1. Moreover, images of the boundaries and sewer layouts of all pilot catchments can be found in Figure 1.

**Table 1**. Summary characteristics of pilot urban catchments

|  | Cran brook | Tor- quay | Morée- Sausset | Sucy-en- Brie | Herent | Ghent | Kralingen |
|---|---|---|---|---|---|---|---|
| Area (ha) | 865.2 | 570.03 | 560.4 | 269 | 511.5 | 649.33 | 670 |
| Catchment length and width (km)* | 6.1/1.4 | 5.4/1.1 | 5.3/1.1 | 4.0/0.7 | 8.2/0.6 | 4.7/1.4 | 2.1/3.2 |
| Catchment shape factor (-) | 0.23 | 0.2 | 0.2 | 0.17 | 0.08 | 0.29 | 1.49 |
| Slope (m/m)*** | 0.0093 | 0.0262 | 0.0029 | 0.0062 | 0.0083 | 0.0001 | 0.0003 |
| Main flow direction (deg) | 239 | 270 | 198 | 138 | 40 | 235 | 152 |
| Type of drainage system | Mostly separate branched | Mostly combined branched | Mostly separate branched | Separate branched | Mostly combined branched | Mostly combined branched | Mostly combined looped |
| Is flow mainly driven by gravity? | Yes | Yes | Yes | Yes | Yes | Yes | No |
| Control elements | 3 storage lakes | 3 storage tanks, 1 pumping station | 2 storage tanks | 1 storage basin, 1 pumping station | 5 main CSO's with control | 15 pumping stations | 20 pumping stations |
| IMP (%)**** | 52% | 26% | 37% | 34% | 27% | 41% | 48% |
| Predominant land-use | R&C | R&C | R&C | R&C | R | R | R&C |
| Population density [pp/ha] | 47 | 60 | 70 | 95 | 20 | 24 | 154 |

*Length = Length of longest flow path (through sewers) to catchment outfall; Width = Catchment Area / Catchment Length;
**Shape factor = Width / Length (this parameter is lower for elongated catchments)
***Catchment slope = Difference in ground elevation between upstream most point and outlet / catchment length
****IMP: total proportion of impervious areas in relation to total catchment area

### Precipitation dataset: dual-polarimetric X-band radar
While radar rainfall products are under development in the four pilot cities or RainGain, high resolution data were obtained for this study from a polarimetric research radar located in Cabauw, the Netherlands (Leijnse et al., 2010). Data were derived for nine storms in the period 2011-2014

and were used to conduct analyses of the space-time scales of storm cells and study hydrological modelling response at a range of space-time resolutions.

The estimated spatial and temporal characteristics of the nine storm events are summarised in table 2. For information on the method applied for estimating space-time scales of the storms we refer to Ochoa-Rodrigues et al. (2015).

**Table 2**. Summary characteristics of pilot urban catchments

| Storm Event ID | Spatial range | Mean Velocity | Anisotropic coefficient | Required Spatial Resolution | Required Temporal Resolution |
|---|---|---|---|---|---|
| | (m) | (m/s) | (-) | (m) | (min) |
| E1 | 4057 | 9.8 | 0.38 | 1695 | 5.8 |
| E2 | 3525 | 9.9 | 0.38 | 1473 | 4.9 |
| E3 | 4655 | 14.0 | 0.55 | 1945 | 4.6 |
| E4 | 3219 | 11.7 | 0.34 | 1345 | 3.8 |
| E5 | 2062 | 14.1 | 0.59 | 861 | 2.0 |
| E6 | 3738 | 11.7 | 0.26 | 1561 | 4.5 |
| E7 | 1703 | 14.0 | 0.24 | 711 | 1.7 |
| E8 | 3644 | 18.4 | 0.36 | 1523 | 2.8 |
| E9 | 2355 | 17.0 | 0.08 | 984 | 1.9 |
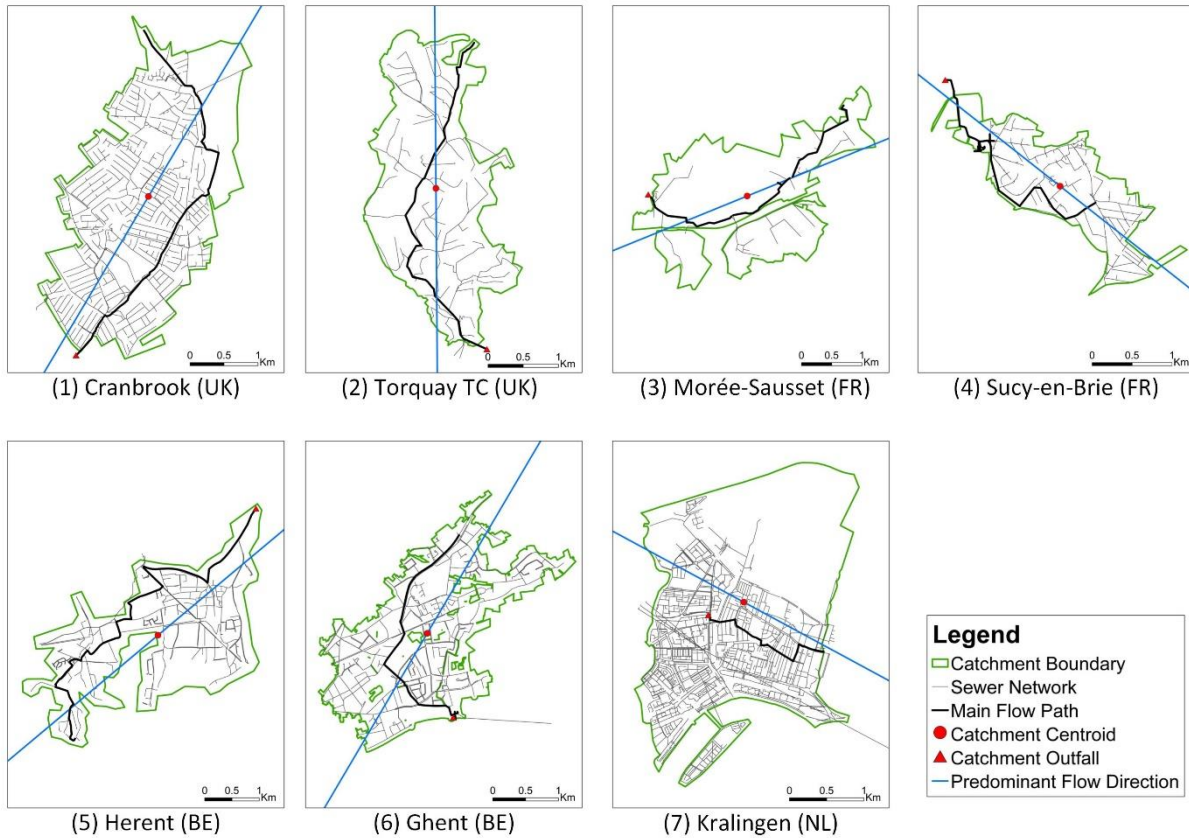
**Figure 1.** Catchment boundary and sewer layout for the pilot urban catchments

## RESULTS AND DISCUSSION

### Storm characteristics

Mean velocity of the nine storms varies from 6.4 m/s to 19.3 m/s (34 to 69 km/h) and storm ranges vary from 1700 to 4660 m. The combination of storm velocity and storm range, together with catchment dimensions, determines the time during which the storm core passes through a given catchment. For the storms and catchments considered in this study, this time varied between ~2-12 min.

### Hydrological response: sensitivity to space-time resolution

In order to investigate the sensitivity of urban drainage models to the spatial-temporal resolution of rainfall inputs, the high-resolution precipitation data for the nine (9) storm events, initially at 100 m and 1 min, were aggregated to a number of coarser temporal and spatial resolutions (up to 3000 m and 10 min) and were applied as input to the urban drainage models of the seven (7) pilot catchments. Results were analysed at different drainage areas of varying sizes (~ 1 ha to ~ 800 ha) within each pilot catchment. Some of the main findings of the hydrological response analysis are summarised in this paper; for an in-depth analysis please refer to Ochoa-Rodriquez et al. (2015).

The results showed that hydrodynamic response behaviour in urban catchments is highly sensitive to combinations of temporal and spatial resolutions of rainfall input. For the storms investigated in this study, hydrodynamic response behaviour was more sensitive to temporal than to spatial resolution. Temporal resolution coarsening beyond the estimated required resolution (between 2 and 6 min) led to under- and overestimations of flow peaks by up to 100% with respect to the original 100 m, 1

minute rainfall input. Similarly, it resulted in low explained variance (down to 20% explained variance, median value, at the level of the entire hydrograph) and flow underestimation at the level of the entire hydrograph (figure 2, illustration of results for 1 catchment, 1 storm). Spatial resolution coarsening led to underestimation of hydrographs for spatial scales between 500 m and 1 km for drainage areas of 1 to 100 ha. A special feature observed in the analysis is the strong interaction between the spatial and temporal resolution of rainfall estimates: the two resolutions must be consistent with each other to prevent loss of information from the higher resolution (more detail in Ochoa-Rodriguez et al., 2015).
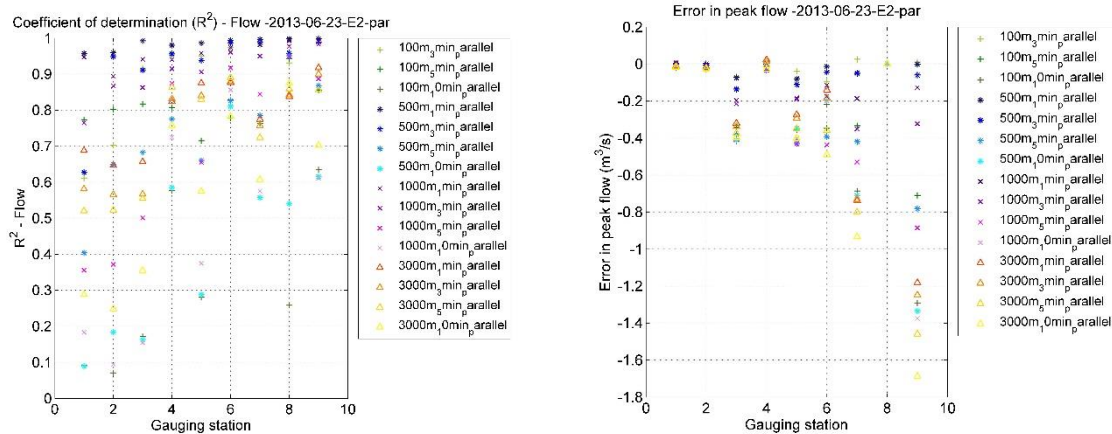


**Figure 2.** Example of hydrological response characteristics for 1 storm event, 1 catchment, 16 resolution combinations. Coefficient of determination and absolute error in peak flow are plotted for nine gauging stations corresponding with drainage area size increasing from ~1 ha to ~ 800 ha. The original input resolution of 100 m, 1 minute resolution is taken as a reference.

**ACKNOWLEDGEMENTS**

**REFERENCES**

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Gregersen, I. B., ... & Nguyen, V. T. V. (2013). Impacts of climate change on rainfall extremes and urban drainage systems: a review. Water Science & Technology, 68(1), 16-28.

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., ... & Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. Hydrological Sciences Journal, 59(1), 1-28.

Leijnse, H., Uijlenhoet, R., van de Beek, C. Z., Overeem, A., Otto, T., Unal, C. M. H., Dufournet, Y., Russchenberg, H. W. J., Figueras i Ventura, J., Klein Baltink, H. & Holleman, I. (2010). Precipitation Measurement at CESAR, the Netherlands. Journal of Hydrometeorology, 11 (6), 1322-1329

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., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., ten Veldhuis, M. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. Journal of Hydrology, in press.

# Computationally efficient modelling approach and software for sewer quantity simulations

Vincent Wolfs[1] and Patrick Willems[1]

[1]Department of Civil Engineering, KU Leuven, Kasteelpark Arenberg 40 box 2448, 3001 Leuven, Belgium (Email corresponding author: vincent.wolfs@bwk.kuleuven.be)

**Key-findings**
- This paper presents a novel modular and lumped computationally efficient modelling approach for sewer system quantity simulations
- An accompanying software tool with GUI was developed for model set-up
- Tests with the developed modelling approach and software on a real-life case study demonstrate that the methodology can account for backwater effects and deliver accurate CSO predictions, while the simulation time is reduced by $4 \cdot 10^5$ times compared to a full hydrodynamic model

**Keywords**
Computational efficiency, CSO, Quantity modelling, Sewer, Software, Surrogate modelling

## BACKGROUND AND RELEVANCE

Mathematical models play a crucial role in management of urban drainage systems. To assess the efficiency of alternative management strategies for different return periods correctly, long term simulations should be performed followed by a statistical post-processing of the results. This approach allows for the explicit incorporation of antecedent conditions. To perform such simulations, models with a very short calculation time are required. This need is also shared by numerous applications that demand for a large number of simulations, such as optimization problems, impact assessments and real-time control. In addition, the importance and potential of integrated urban water system models becomes more and more recognized, in which not only the urban drainage system itself, but many other aspects are being considered simultaneously (e.g. Fratini et al., 2012; Bach et al. 2014). This evolution towards integrated models has an impact on the way systems are modelled. In summary, future urban drainage management imposes the following four key requirements on models: (1) a short calculation time, (2) yield accurate predictions, (3) easily linkable with other modules, and (4) allow for different levels of model detail tailored to the application in mind, since not all dynamics are relevant for certain applications.

Conventional full hydrodynamic models, such as InfoWorks and SWMM, emulate reality accurately, but do not fulfil the other three specified requirements. Their long calculation time and difficulty to link models made in different software programs in particular makes their use infeasible for many applications. Different simplified modelling approaches (e.g. Saagi et al., 2014) and software packages exist (e.g. CITY DRAIN (Achleitner et al., 2007), SMUSI (Muschalla et al., 2006)), but these suffer from fundamental shortcomings when emulating complex flow regimes, such as backwater effects and pressurized or reverse flows (Vanrolleghem et al., 2009). This paper presents a novel conceptual mechanistic modelling approach and accompanying software tool that can account for such behaviour, and strives for model parsimony and pragmatism by focusing on the dominating processes.

## METHODOLOGY

The developed modelling approach emulates the results of detailed models. The methodology is based on the storage cell concept, in which the sewer network is lumped and divided into multiple cells. The water balance is explicitly closed in each cell, while several structures with different complexity are employed to predict the inter-cell flows: predefined discharges in combination with logic controls, a static-dynamic virtual storage approach, multi-input transfer functions representing arrangements of linear reservoirs, and structures incorporating artificial neural networks. This modular approach is very flexible and enables the modeller to pick the most suitable model structure depending on the system's dynamics. Since some of these structures are data-driven and hence adapt themselves to the data provided during configuration, many complex flow dynamics can be mimicked accurately.

A software tool with GUIs was developed that guides the user through the configuration of the model. Finally, the model itself is automatically generated in a C-script together with files containing all boundary data. Such script can easily be expanded with other modules and is computationally very efficient. The workflow is schematized in Figure 1. Please contact the authors concerning software availability.
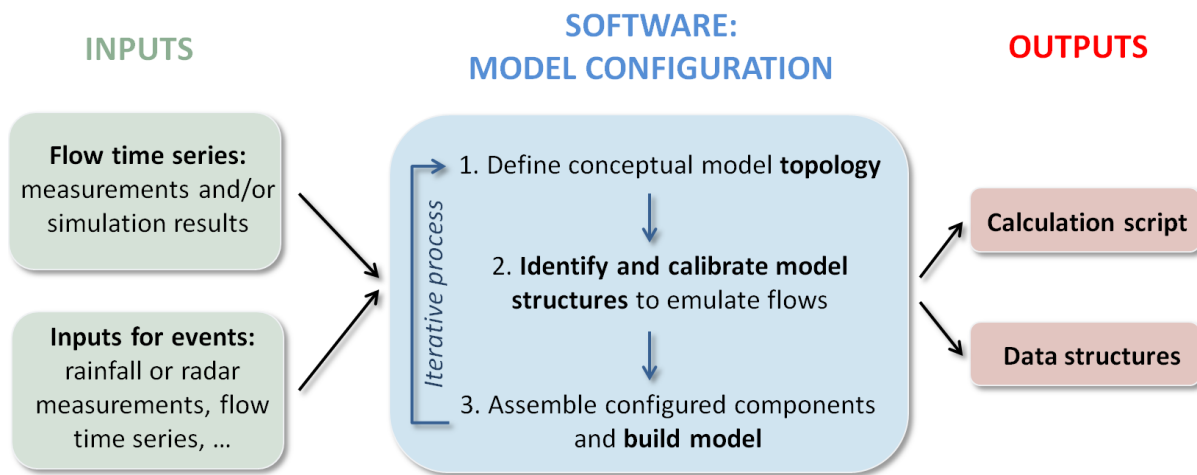


**Figure 1.** Schematic overview of the workflow and functionality of the software tool

## RESULTS AND DISCUSSION

To demonstrate the approach, a surrogate model was created that emulates an InfoWorks CS model of the sewer system of the cities Sint-Amandsberg and Oostakker in Belgium (Figure 2). The detailed model is relatively flat and contains over 6000 conduits, 43 pumps and about 100 other hydraulic structures. The surrogate model aims to predict the CSO flows and other outflows of the system accurately, both in magnitude and timing. Simulation results of three synthetic events with a frequency of 7 times a year (f07), and return periods of 2 (T02) and 10 (T10) years were used for calibration, while three others (f10, T05 and T20) were used for validation.
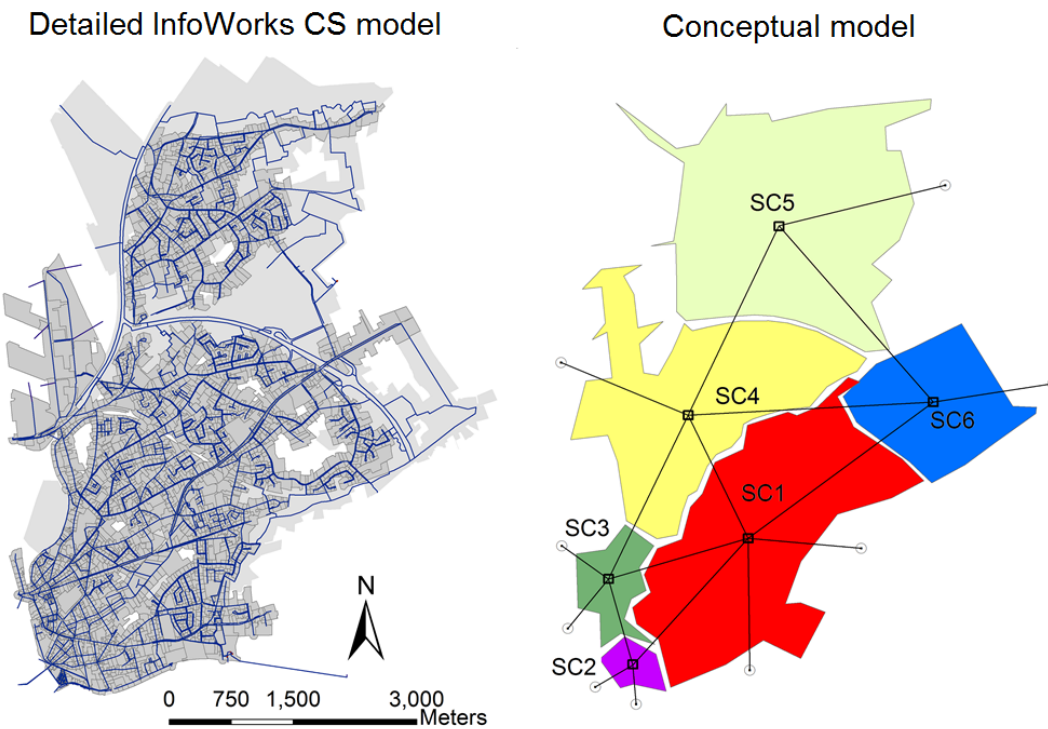
Detailed InfoWorks CS model

Conceptual model



**Figure 2.** Network topology of the detailed InfoWorks CS model (left) and the configured conceptual model (right). The inter-cell fluxes and fluxes to outfalls are also indicated.

The topology of the conceptual model consists of 6 cells or subcatchments (SC; see also Figure 2). The boundaries of these SCs were selected such that the number of inter-cell fluxes is minimal. Some flows over CSO's and other outflows that are situated close to each other were aggregated to minimize the amount of variables that has to be calculated. This yielded 9 flows to outfalls and as many inter-cell fluxes. For each flux, the most suitable model structure was identified and calibrated. The Nash-Sutcliffe efficiencies (NSE) of the simulated flows to outfalls of the surrogate model are shown in Table 1. It is clear that most NSE-values are close to unity, indicating a good fit, even for the most extreme events (f10 and T20) which fall outside the calibration range. The prediction of flow 'E' is characterized by lower NSE-values. This flow is determined by a pump with a predefined discharge (see also Figure 3). Indeed, small divergences in pump operation can result in large NSE-value deviations.

The calculation time of the obtained surrogate model amounts approximately 0.003 seconds for simulating a two-day event with a time step of just 10 seconds. This very short computation time, together with its potential to emulate complex flow dynamics and its flexibility, make the proposed modelling approach a suitable tool for many applications in urban drainage management.

**Table 1.** NSE-values indicating the goodness-of-fit of the simulated fluxes to outfalls in the conceptual model and the InfoWorks CS model. The letters "C" and "V" denote calibration and validation events respectively.

| Flow | Symbol | f10 (V) | f07 (C) | T02 (C) | T05 (V) | T10 (C) | T20 (V) |
|------|--------|---------|---------|---------|---------|---------|---------|
| SC1 – Outfalls – a | A | 0.97 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |
| SC1 – Outfalls – b | B | 0.86 | 0.87 | 0.87 | 0.90 | 0.90 | 0.90 |
| SC2 – Outfalls – a | C | 0.94 | 0.96 | 0.96 | 0.96 | 0.97 | 0.97 |
| SC2 – Outfalls – b | D | 0.80 | 0.78 | 0.76 | 0.76 | 0.76 | 0.74 |
| SC3 – Outfalls – a | E | 0.57 | 0.52 | 0.29 | 0.36 | 0.33 | 0.33 |
| SC3 – Outfalls – b | F | - | - | 0.96 | 0.98 | 0.98 | 0.98 |
| SC4 – Outfalls | G | 0.96 | 0.97 | 0.99 | 0.99 | 0.99 | 0.99 |
| SC5 – Outfalls | H | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 |
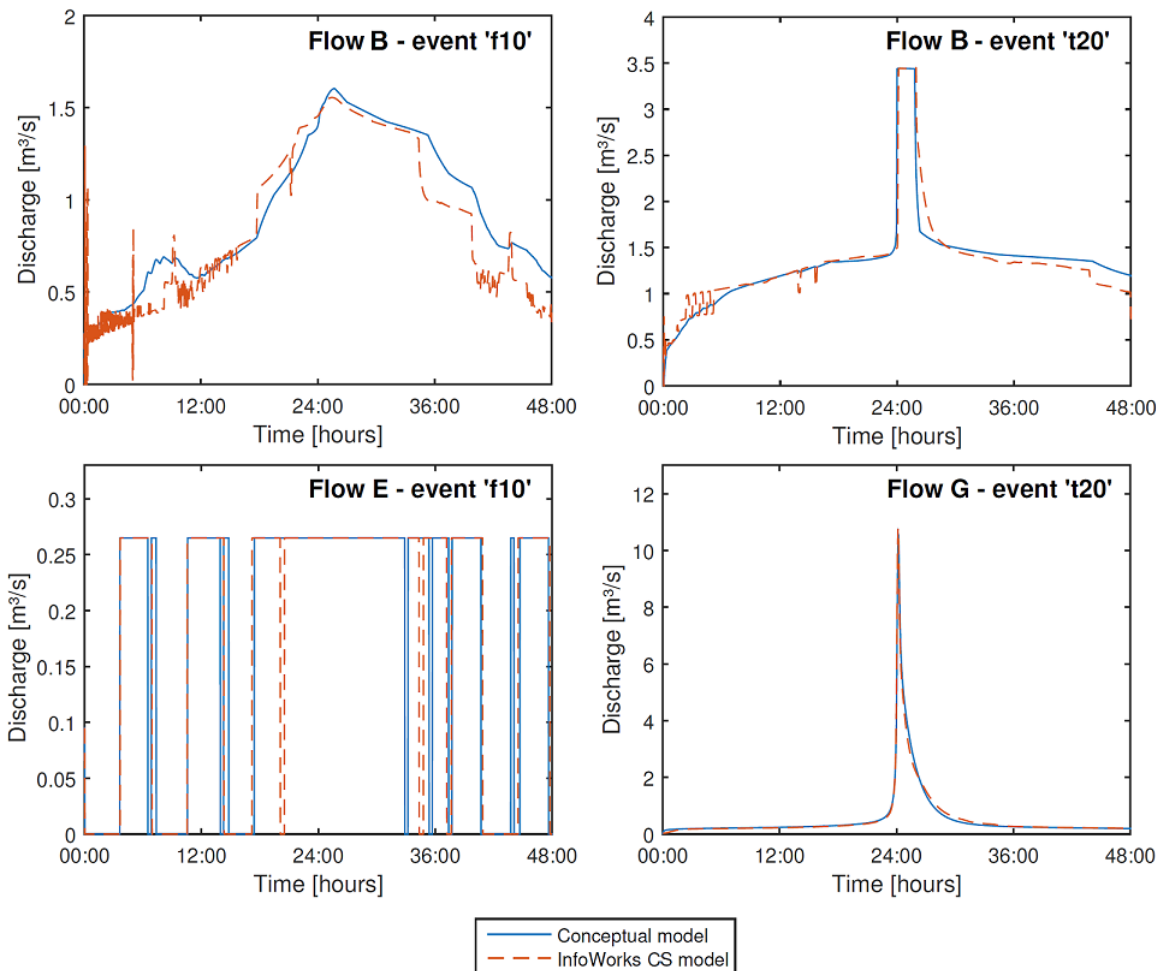


**Figure 3.** Simulated fluxes to outfalls in the conceptual and InfoWorks CS model for different locations and events

# REFERENCES

Achleitner, S., Möderl, M. and Rauch, W. (2007). CITY DRAIN © - An open source approach for simulation of integrated urban drainage systems. Environ. Model. Softw. 22, 1184-1198.

Bach, P.M., Rauch, W., Mikkselsen, P., McCarthy, D. and Deletic, A. (2014). A critical review of integrated urban water modelling – Urban drainage and beyond. Environ. Model. Softw. 54, 88-107.

Fratini, C.F., Geldof, G.D., Kluck, J. and Mikkelsen, P.S. (2012). Three Points Approach (3PA) for urban flood risk management: A tool to support climate change adaptation through transdisciplinary and multifunctionality. Urb. Wat. Journ. 9(5), 317-331.

Muschalla, D., Reussner, F., Schneider, S. and Ostrowski, M.W. (2006). Dokumentation des Schmutzfrachtmodells SMUSI Version 5.0. Technische Universität Darmstadt, Germany. (in German)

Saagi, R., Flores-Alsina, X., Fu, G., Benedetti, L., Gernaey, K.V., Jeppsson, U., Butler, D. (2014). Benchmarking integrated control strategies using an extended BSM2 platform. Proc. 13th Int. Conf. on Urban Drainage. Sarawak, Malaysia, Sept. 7-12 2014.

Vanrolleghem, P.A., Kamradt, B., Solvi, A.M. and Muschalla, D. (2009). Making the best of two Hydrological Flow Routing Models: Nonlinear Outflow-Volume Relationships and Backwater Effects Model. Proc. 8th Int. Conf. on Urban Drainage Modelling (8UDM). Tokyo, Japan, Sept. 7-11 2009.

# A framework for probabilistic pluvial flood nowcasting for urban areas

Victor Ntegeka (1), Damian Murla (1), Lipen Wang (1), Loris Foresti (2), Maarten Reyniers (2), Laurent Delobbe (2), Kristine Van Herk (3), Luc Van Ootegem (3), and Patrick Willems (1)

(1) KU Leuven, Hydraulics section, Leuven, Belgium, (2) Royal Meteorological Institute of Belgium, Uccle, Belgium, (3) KU Leuven, HIVA Research Institute for Work and Society, Leuven, Belgium

Pluvial flood nowcasting is gaining ground not least because of the advancements in rainfall forecasting schemes. Short-term forecasts and applications have benefited from the availability of such forecasts with high resolution in space (∼1km) and time (∼5min). In this regard, it is vital to evaluate the potential of nowcasting products for urban inundation applications. One of the most advanced Quantitative Precipitation Forecasting (QPF) techniques is the Short-Term Ensemble Prediction System, which was originally co-developed by the UK Met Office and Australian Bureau of Meteorology. The scheme was further tuned to better estimate extreme and moderate events for the Belgian area (STEPS-BE). Against this backdrop, a probabilistic framework has been developed that consists of: (1) rainfall nowcasts; (2) sewer hydraulic model; (3) flood damage estimation; and (4) urban inundation risk mapping.

STEPS-BE forecasts are provided at high resolution (1km/5min) with 20 ensemble members with a lead time of up to 2 hours using a 4 C-band radar composite as input. Forecasts' verification was performed over the cities of Leuven and Ghent and biases were found to be small. The hydraulic model consists of the 1D sewer network and an innovative 'nested' 2D surface model to model 2D urban surface inundations at high resolution. The surface components are categorized into three groups and each group is modelled using triangular meshes at different resolutions; these include streets (3.75 – 15 m2), high flood hazard areas (12.5 – 50 m2) and low flood hazard areas (75 – 300 m2). Functions describing urban flood damage and social consequences were empirically derived based on questionnaires to people in the region that were recently affected by sewer floods. Probabilistic urban flood risk maps were prepared based on spatial interpolation techniques of flood inundation.

The method has been implemented and tested for the villages Oostakker and Sint-Amandsberg, which are part of the larger city of Gent, Belgium. After each of the different above-mentioned components were evaluated, they were combined and tested for recent historical flood events. The rainfall nowcasting, hydraulic sewer and 2D inundation modelling and socio-economical flood risk results each could be partly evaluated: the rainfall nowcasting results based on radar data and rain gauges; the hydraulic sewer model results based on water level and discharge data at pumping stations; the 2D inundation modelling results based on limited data on some recent flood locations and inundation depths; the results for the socio-economical flood consequences of the most extreme events based on claims in the database of the national disaster agency. Different methods for visualization of the probabilistic inundation results are proposed and tested.

# Prototype system for nowcasting of urban flood risks

## Prototype d'un système de prévision immédiate des risques d'inondation en milieu urbain

Patrick Willems*, Laurent Delobbe**, Maarten Reyniers**, Luc Van Ootegem***

\*    Hydraulics Division, Department of Civil Engineering, KU Leuven.
     Kasteelpark Arenberg 40, 3001 Leuven, Belgium
     (Patrick.Willems@bwk.kuleuven.be)
\*\*   Royal Meteorological Institute of Belgium, Uccle, Belgium.
\*\*\*  KU Leuven, HIVA Research Institute for Work and Society, Leuven, Belgium.

## RÉSUMÉ

Un système prototype a été mis en place pour la prévision immédiate des probabilités de risques d'inondation dans les zones urbaines. Il se compose d'un modèle de prévision des chutes de pluie à court terme sur la base des données radar et pluviométriques, d'un modèle d'inondation hydraulique et surface 1D-2D, un modèle pour évaluer les dommages économiques et les conséquences sociales des inondations urbaines, et les méthodes de post-traitement statistique pour produire des cartes de probabilités des risques d'inondation urbaines. Le système a été mis en œuvre et testé pour les villages d'Oostakker et Sint-Amandsberg, qui sont rattachés à la ville de Gand, en Belgique. Le système a été évalué pour des événements historiques récents d'inondations. Les résultats sur la prévision des précipitations, les inondations de surface, et les risques socio-économiques pourraient être en partie évalués sur la base des observations historiques des précipitations, les débits des égouts et les niveaux d'eau, les inondations de surface et des dommages dus aux inondations ayant fait l'objet de réclamations auprès de l'agence nationale pour la gestion des catastrophes. Différentes méthodes de visualisation des résultats sur les probabilités d'inondation ont été proposées et testées.

## ABSTRACT

A prototype system has been set up for the probabilistic nowcasting of the inundation risks in urban areas. It consists of a rainfall nowcasting model based on radar and rain gauge data, a nested 1D-2D sewer hydraulic and surface inundation model, a model to assess the damages and social consequences of the urban inundations, and statistical post-processing methods to produce probabilistic urban inundation risk maps. The system has been implemented and tested for the villages Oostakker and Sint-Amandsberg, which are part of the larger city of Gent, Belgium. The system has been evaluated for recent historical flood events. The results on the rainfall nowcasting, the surface inundations, and the socio-economical risks could be partly evaluated based on historical observations of rainfall, sewer flows and water levels, surface inundations and flood damages claimed for the national disaster agency. Different methods for visualisation of the probabilistic inundation results were proposed and tested.

## KEYWORDS

Floods, Forecasting, Rainfall, Radar data, Sewer system, Socio-economic consequences

# 1    INTRODUCTION

Within the scope of the interdisciplinary research project PLURISK on "Forecasting and management of extreme rainfall induced risks in the urban environment" for the Belgian Science Policy Office (BELSPO), a probabilistic model has been set up and evaluated for the nowcasting (short-term forecasting) of urban inundations. It consists of the following components: (1) a rainfall nowcasting model; (2) a hydraulic model that consists of the 1D sewer network and an innovative 'nested' 2D surface model to model 2D urban surface inundations at high resolution; (3) functions describing urban flood damage and social consequences in relation to inundation depth or rainfall intensity; (4) statistical post-processing methods in order to produce probabilistic urban flood risk maps: spatial maps representing the probability of flooding. The prototype system has been implemented and tested for the villages Oostakker and Sint-Amandsberg, which are part of the larger city of Gent, Belgium (Figure 1).
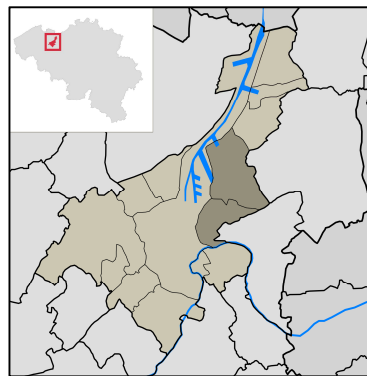


Figure 1: Study area of the villages Oostakker and Sint-Amandsberg in the larger city of Gent, Belgium: (left) location in Belgium; (right) sewer network and location FURUNO's WR-2100 X-band radars.
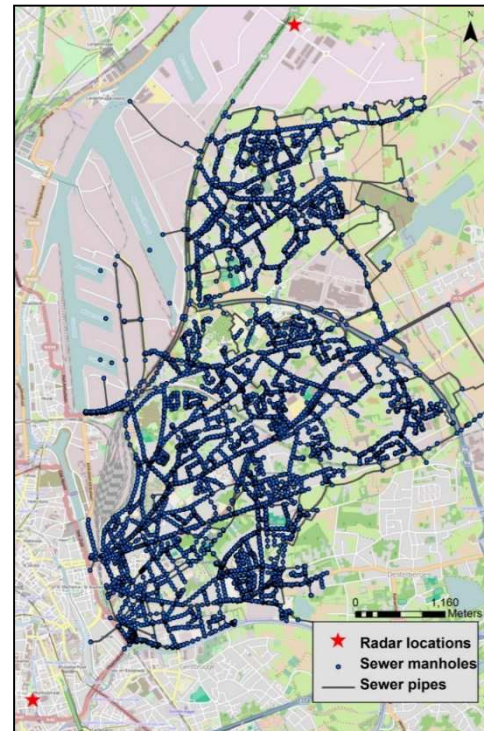
# 2    RAINFALL NOWCASTING

The rainfall nowcasting makes use of the Short-Term Ensemble Prediction System (STEPS), originally co-developed by the UK Met Office and Australian Bureau of Meteorology, but further customised for urban applications in Belgium (denoted STEPS-BE). The STEPS-BE nowcasting is based on temporal extrapolation of rain storms observed by radar data. The extrapolation is probabilistic in the sense that the deterministic radar extrapolation is perturbed with stochastic noise (Bowler et al., 2006). This noise aims to account for the unpredictable rainfall growth and decay processes. It hence represents the nowcast errors together with its spatial and temporal correlations.

To obtain uncertainty estimates for the rainfall nowcasts, 20 runs are conducted (20 member ensemble nowcast). So far, results are obtained at 1 km and 5 min resolutions up to 2 hours lead time using as input the composite image of the 4 C-Band radars located at Wideumont (RMI), Zaventem (Belgocontrol), Jabbeke (RMI) and Avesnois (MétéoFrance) (example Figure 2). Future developments involve the use of higher resolution X-band radar data at Leuven and Gent (FURUNO WR-2100 radars). In comparison with the original STEPS, some improvements were made to obtain smoother velocity fields and to generate stochastic rainfall within the boundaries of the advected radar composite.

Comparison of the rainfall nowcasts with radar observations for historical events has shown that nowcasts of rainfall intensities exceeding 0.5 mm/h are reliable up to 120 minutes lead time, or shorter for higher intensities (e.g. 30 minutes for intensities exceeding 5 mm/h) (Foresti et al., 2015). As expected, convective rainfall features have less predictability than large areas of stratiform rain.
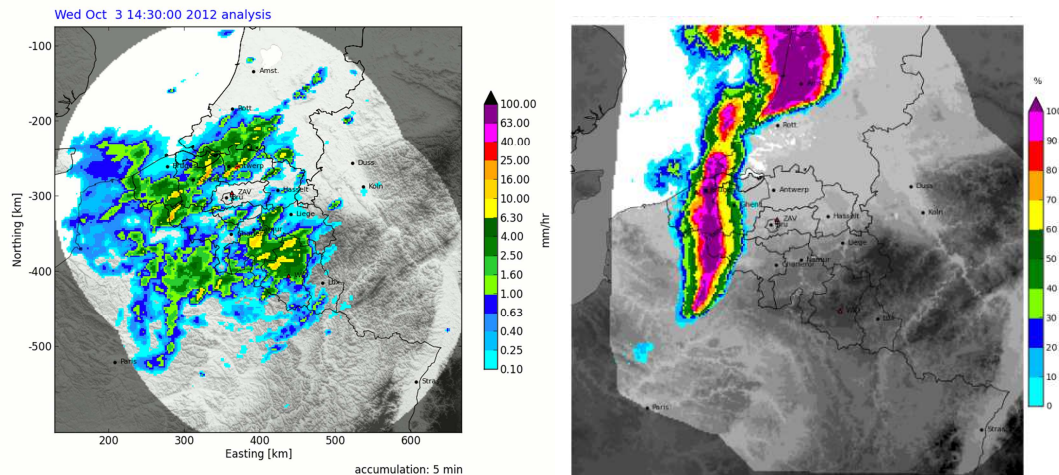
Figure 2: Example of the STEPS-based rainfall nowcasts: (left) rainfall nowcast for 3/10/2012 14:30 for one ensemble member; (right) probability of rainfall intensities higher than 1 mm/h.

## 3    2D SURFACE INUNDATION NOWCASTING

The rainfall nowcasts are converted in 2D surface inundation maps by means of hydraulic modelling. A 1D full hydrodynamic model for the underground sewer system, implemented in InfoWorks ICM, was coupled to a 2D hydrodynamic model for the surface flow. A digital elevation model (DEM) of very high spatial resolution of 1 meter was used as topographical input. In order to limit computational times, a nested 2D surface model was applied by developing the 2D surface mesh in diverse resolutions (Murla & Willems, 2015). Main streets as mesh zones together with buildings as void regions were modelled at the highest mesh resolution ($3.75m^2 - 15m^2$). They channel most of the flood water from the manholes and are crucial for the interactions within the 1D sewer network. The areas that recorded flooding outside these main streets were implemented with a second mesh resolution ($12.5 m^2 - 50 m^2$). The low flood risk areas were finally simulated at a coarser mesh resolution ($75 m^2 - 300 m^2$)



(example Figure 4). Permeable areas were identified and used as infiltration zones using the Horton infiltration model. The nested approach aims to take the advantages of both modelling systems to optimize the accuracy and computational time in describing urban surface inundations and allow an optimal representation of surface flooding in 2D and of the corresponding coupled 1D underground sewer.
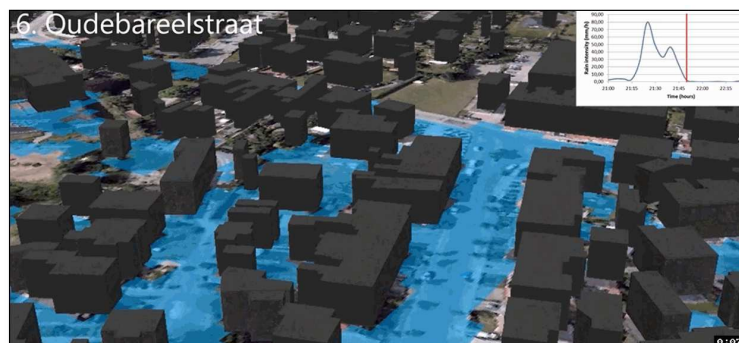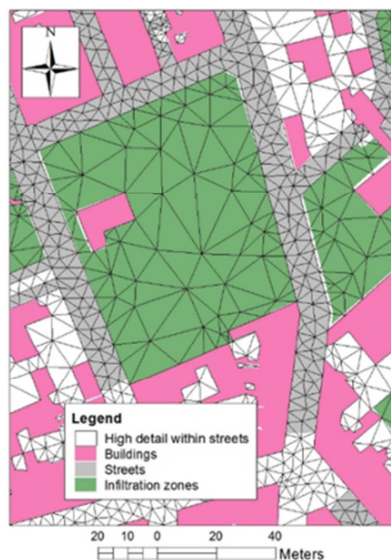


Figure 4: (Left) Surface inundation model with three triangular mesh resolutions; (right) example of urban inundation result visualized based on the high resolution 1-m DEM.

The 1D sewer simulation results were validated based on water level and discharge observations obtained from a 1-month in-sewer monitoring campaign, and water level data continuously measured at pumping stations for longer periods. Surface inundation results were validated for recent historical floods based on scattered information on inundation locations and depth assessments from photographs, interviews and fire brigade interventions. These validations were based on best estimates of the real rainfall over the sewer catchments for these events.

## 4    FLOOD RISK NOWCASTING

In order to convert the hydraulic flood variables into risk information, both the tangible and non-tangible consequences of urban floods were studied. For the tangible consequences (monetary damage), both "depth-damage" and "rainfall-damage" models (Figure 5) were derived after statistical analysis of the survey data of 346 Flemish households that were victims of pluvial floods (Van Ootegem et al., 2015). These survery data were collected after sending questionnaires to 3963 victims of urban floods (info obtained from newspapers and from the historical database of claims to the Belgian Federal Disaster Fund). 24.5% completed questionnaires were received, from a total of 973 households.

They statistical analysis shows an average increase in the damage between 3% and 5.3% per cm increase in inundation depth at the ground floor, depending on the type of damage (building, content or total). Other factors accounted for are building related properties such as the location (e.g. village), type of house (single, in row), etc., and person related properties such as the salary or income, whether the person suffered already from historical flood events, hence whether the person is aware of the flood problem and whether the person has self-coping capacity, e.g. by taking precautionary measures, etc. The uncertainty in the models was also derived from the survey data. For the consequences on cultural heritage in the urban environment, a separate study was conducted. Validation could be done by comparing the damage estimates for a number of historical floods with the amounts claimed and paid by the Federal Disaster Fund Agency.
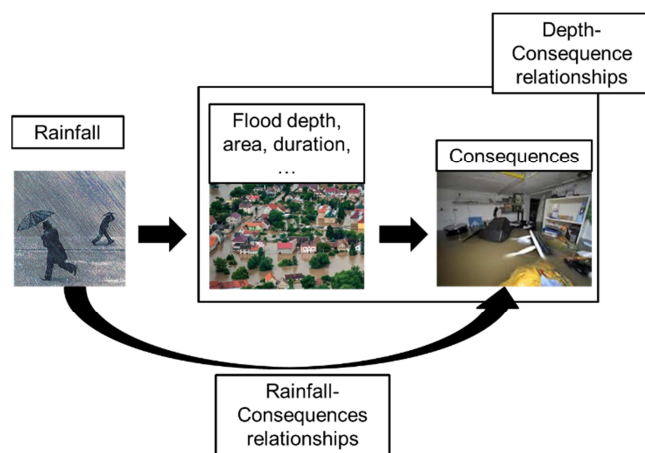
The non-tangible consequences were assessed by analyzing the relation between the self-reported "happiness in life" of the people that completed the questionnaire, when they are asked to look backwards and into the future, and the explanatory variables mentioned above.

Figure 5: Depth-consequence versus rainfall-consequence relationships.

## 5    VISUALISATION AND COMMUNICATION OF NOWCASTS

The urban inundation risk nowcasts are subject to high uncertainties as a result of the uncertainties in the rainfall nowcast, the hydraulic sewer modelling and urban surface inundation modelling, and the assessment of the consequences. For each of these three steps, uncertainties were quantified and propagated to obtain an overall uncertainty estimate of the risk nowcast. Given the importance to clearly but carefully communicate this nowcast uncertainty, different visualisation methods were tested and evaluated after consultation of the end users. Rainfall and flood probability or return periods maps and flood risk maps were finally selected, as well as plots showing the temporal evolution of the flood probability and risk.

## LIST OF REFERENCES

Bowler, N.E.H., Pierce, C.E., Seed, A.W. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, *Q. J. Roy. Meteorol. Soc.*, 132, 2127-2155.

Foresti, L., Reyniers, M., Delobbe, L. (2015). Probabilistic and ensemble verification of the Short-Term Ensemble Prediction System in Belgium. In: *UrbanRain15 - 10th International Workshop on Precipitation in Urban Areas*, Pontresina, Switzerland, 1-5 December 2015.

Murla, D. Willems, P. (2015). Nested 1D-2D approach for urban surface flood modeling. *Geophys. Res. Abstr.*, 17, EGU2015-7162: EGU General Assembly 2015, Vienna, Austria, 12-17 April 2015.

Van Ootegem, L., Van Herck, K., Creten, T., Verhofstadt, E., Foresti, L., Goudenhoofdt, E., Reyniers, M., Delobbe, L., Murla Tuyls, D., Willems, P. (2015). Pluvial flood damage models: Exploring the potential of multivariate rainfall models. *J. Flood Risk Manag.*, Submitted.

# CMD: fast and tailored conceptual river and sewer models for integrated water management

## CMD : modèles conceptuels rapides et adaptés de la rivière et du réseau pour une gestion intégrée de l'eau

Vincent Wolfs*, Patrick Willems*

\*  Hydraulics Division, Department of Civil Engineering, KU Leuven.
Kasteelpark Arenberg 40, 3001 Leuven, Belgium
(vincent.wolfs@bwk.kuleuven.be)

## RÉSUMÉ

En appui de la gestion de l'eau, il faut des modèles efficaces afin d'évaluer et d'optimiser des stratégies différentes au niveau du bassin, de quantifier l'impact des tendances telles que le changement climatique ou l'urbanisation en hausse, et enfin, de prendre en compte plusieurs incertitudes et risques. Pour les gestionnaires de l'eau, les modèles hydrodynamiques sont devenus l'outil standard. Néanmoins, ces modèles présentent des lacunes considérables, dont les principales sont : les temps de simulation très longs, les possibilités d'interfaçage limitées et la caractérisation souvent trop complexe pour les applications visées. Pour cette raison, nous avons développé une nouvelle méthodologie flexible en vue de convertir ces modèles détaillés de la rivière et du réseau en modèles conceptuels. Plusieurs études de cas ont montré que les modèles conceptuels sont entre $10^3$ et $10^6$ fois plus rapides que les modèles hydrodynamiques classiques, à la suite de la jonction efficace des processus et en s'appuyant sur les structures de modèles avancées. En parallèle, les modèles conceptuels imitent amplement la dynamique complexe des modèles plus détaillés. Un outil logiciel semi-automatique, nommé CMD, a été développé pour configurer rapidement des modèles conceptuels. En conceptualisant deux modèles de réseau détaillés de InfoWorks CS pour les villes de Geel et de Mol, aussi qu'un modèle détaillé en MIKE11 de la rivière Molse Nete, cette communication illustre l'approche de la modélisation et du logiciel.

## ABSTRACT

Efficient models are needed in support of water management to evaluate and optimize different strategies on catchment level, quantify the impact of trends such as climate change or the increasing urbanization, and account for various uncertainties and risks. Detailed hydrodynamic models have become the standard tools of water managers. However, these models suffer from several fundamental shortcomings, of which very long simulation times, limited interfacing possibilities and an often overly complex characterization for the intended applications are arguably the main ones. Therefore, we developed a novel flexible framework to translate detailed river and sewer quantity models to lumped conceptual models that can be tailored to a specific purpose. Various case studies showed that these conceptual models are between $10^3$ and $10^6$ times faster than conventional hydrodynamic models due to the efficient lumping of processes and by relying on advanced model structures, while they emulate the complex dynamics of the more detailed models accurately. A semi-automatic tool, named CMD, was developed to quickly configure conceptual models. This paper illustrates the modelling approach and tool by conceptualizing two detailed InfoWorks CS sewer models for the cities of Geel and Mol, and a detailed MIKE11 model of the Molse Nete River.

## KEYWORDS

Calculation speed, conceptual modelling, decision support, software, water management.

# 1   INTRODUCTION

Water systems are characterized by many interacting processes on different scales. An integrated approach at catchment scale is required that can deal with these interactions to develop effective and sustainable strategies. Due to the inherent complexity of the water system and the broad scope of water management, models are needed to support decision making. These models need to be employable for very diverse analyses, such as optimization questions, uncertainty and risk quantification and impact estimation of trends. Each type of analyses poses specific model requirements. However, water managers rely almost exclusively on detailed hydrodynamic models for designing new strategies or performing analyses of river and sewer systems. Due to their prolonged calculation times, these models cannot be employed for running long term simulations, or for applications that require a large amount of simulations, such as optimization questions or uncertainty analyses. In addition, integration of such detailed models is limited to local scales, and often problematic for technical reasons. Therefore, we propose the complimentary use of simplified or so-called conceptual models besides the conventionally used detailed hydrodynamic models. These conceptual models lump processes and rely on an alternative characterization of the river and sewer systems instead of using the de Saint-Venant equations. Their parsimonious structure results in very short simulation times and facilitates model interfacing, enabling multidisciplinary analyses at catchment scale. These features make conceptual models ideally suited for applications requiring fast and integrated models, while the results could be linked back to hydrodynamic models if spatially detailed investigations are required.

A new, integrated modelling approach is being developed to set up such conceptual models. This framework incorporates a flexible hydrological modelling component that can be efficiently employed at different scales and that delivers consistent results with varying resolution and model detail (Tran and Willems, 2015). In addition, two data-based mechanistic and modular approaches are included, one tailored to river modelling (Wolfs et al., 2015) and another for modelling sewer systems (Wolfs and Willems, 2015). Research is also being done on conceptual water quality modelling for rivers. To ensure that the conceptual models can be configured quickly, a user-friendly tool named Conceptual Model Developer (CMD) has been set up. **Figure 1** schematizes the modelling framework. This paper focuses on the newly developed conceptual modelling methodology for rivers and urban drainage systems.
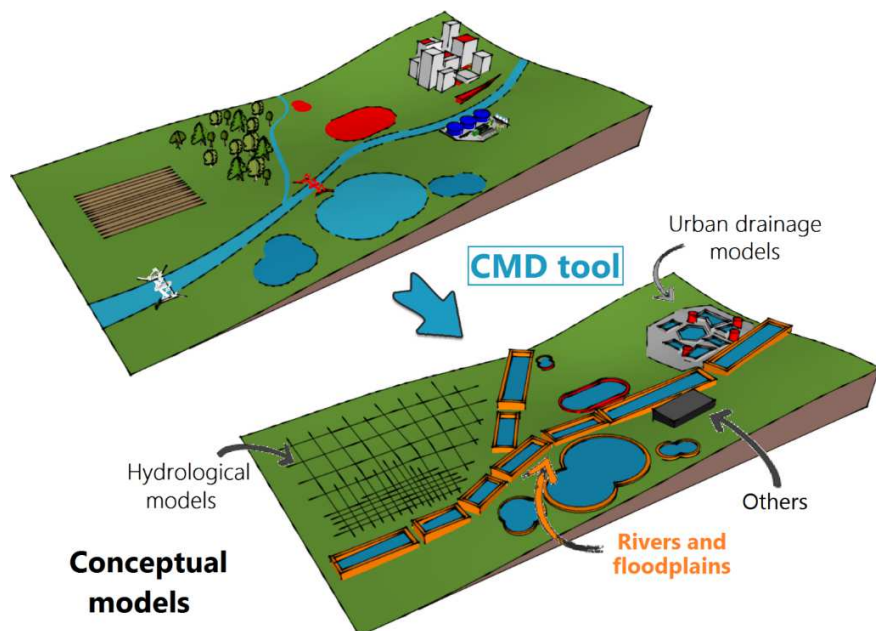


Figure 1: Schematization of the newly developed modelling framework to configure fast and tailored integrated conceptual models using the developed CMD software.

# 2   CONCEPTUAL MODELLING METHODOLOGY AND SOFTWARE

A new data-driven mechanistic and modular conceptual modelling approach was developed for rivers and sewer systems. The reader is referred to Wolfs et al. (2015) and Wolfs and Willems (2015) for details on the modelling approaches. The main characteristics are briefly discussed in this paragraph.

The conceptual models are derived from data, but most of the parameters are still interpretable in physically meaningful terms. For instance, dikes and hydraulic structures can be modelled explicitly in river systems. Such mechanistic representation is crucial for many scenario analyses and optimization questions. Due to the modularity of the methodology, the modeller can select the most appropriate model structures to emulate the dynamics of the system, including backwater effects and pressurized or reverse flows. Structures that are frequently applied in hydrology, such as the linear reservoir theory and transfer functions, are combined with advanced machine learning techniques such as artificial neural networks, adaptive neuro fuzzy inference systems, M5' model trees and state dependent parameter models. The approach is based on the storage cell concept, which implies that the entire system is divided in multiple interconnected cells. In each cell, the water balance is closed explicitly. Figure 2 shows the topologies of a conceptual river and sewer model. In addition, a new discrete and explicit solver was developed that employs a variable time step that differs in space and time. This solver maximizes computational efficiency and avoids numerical instabilities during simulations. The developed modelling methodology results in flexible, accurate and very fast models that are suitable for many scenario investigations and the optimization of strategies.
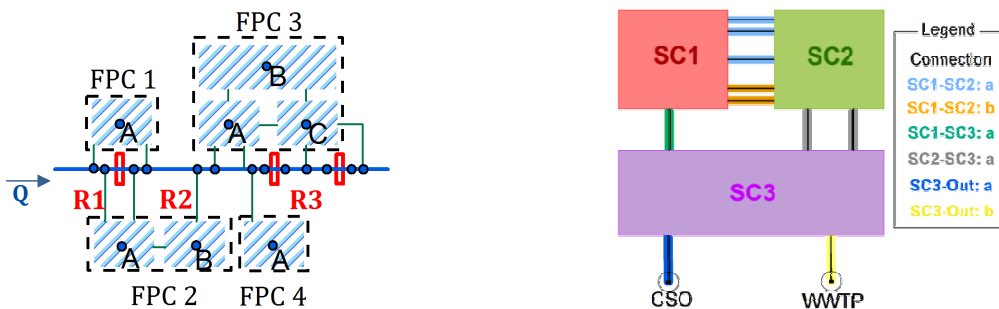


Figure 2: Example of a conceptual model topology for a river with floodplains and hydraulic structures indicated in red (left) and an urban drainage system (right) based on the storage cell concept. The broad gamut of included model structures is used to calculate the variables in the system.

The developed modelling approaches were incorporated in the CMD software tool to facilitate model set-up. Based on the delineated topology of the conceptual model, the tool "learns" how the different elements are interconnected. This enables fully automatic handling of the data sets and calibrated model structures. The tool guides the user through the step-wise calibration. The user can use built-in and tailored algorithms to identify the most suitable model structures. The tool is coded in MATLAB and equipped with GUIs. The models itself are generated in the C programming language to yield the fastest possible models. A close interfacing was foreseen with the InfoWorks and MIKE programs.

## 3   CASE STUDY AND APPLICATIONS

The functionality of the developed modelling approaches and software is hereafter illustrated based on a case study of the sewer systems of the cities Mol and Geel, and the receiving Molse Nete River. Conceptual models were set up and calibrated to existing full hydrodynamic models of the two sewer systems and the river system and consequently interfaced. The coupled river-sewer model was used to quantify the impact of the combined sewer overflows (CSOs) on the water quality of the receiving river (see Keupers et al., 2015, for more details on the water quality part). For the water quantity part, the main objective of the conceptual models was to predict CSO flows accurately, together with the flows and water levels in the river.

To overcome the lack of sufficient and accurate measurement data, simulation results of the existing full hydrodynamic models were employed. These detailed models were implemented in InfoWorks CS for the two sewer systems and MIKE11 for the river system. For the sewer models, the conceptual models were configured based on simulation results for design storms with return periods ranging between 1/20 and 20 years. For the river model, this was done based on a 1-year long term simulation in the MIKE11 model. These events and time series were chosen to ensure that the most important dynamics are included in the calibration and validation sets. The final conceptual model topology of the two sewer systems is shown in **Figure 3**. The average Nash-Sutcliffe efficiencies (NSE) for the simulated volumes of cells in the conceptual model are 0.81 and 0.94 for the sewer systems of Mol and Geel respectively for all validation events, indicating a good fit. The performance of the conceptual river model is exceptionally good, with NSE values for all simulated flows and stages exceeding 0.985 for all investigated events. Simulating a one-year period with the coupled conceptual river-sewer model takes merely 1.41 seconds, while this takes almost 60 hours with the hydrodynamic models.
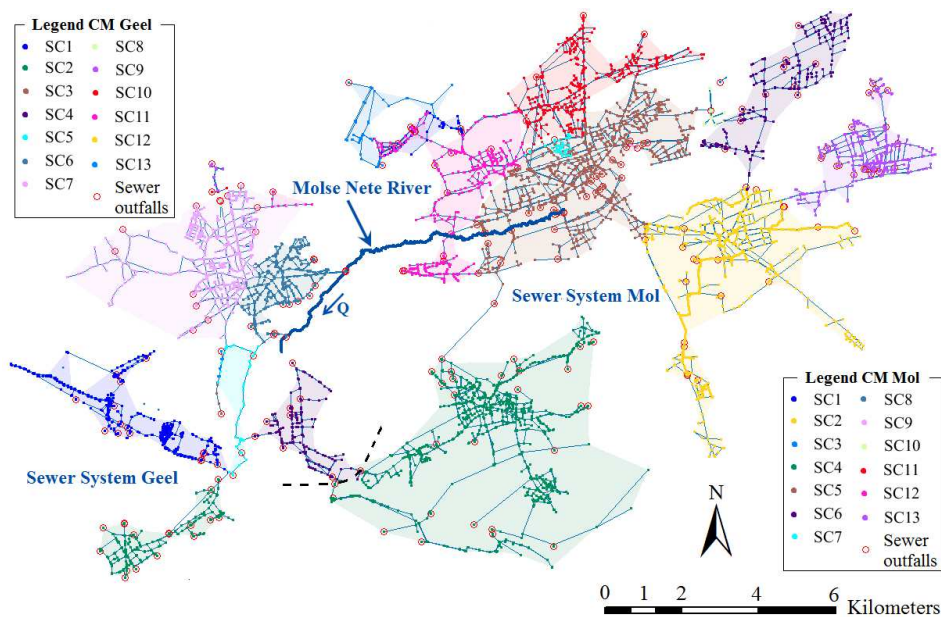
Figure 3: Topology of the conceptual models of the sewer systems of Mol and Geel, together with the modelled segment of the Molse Nete River.

Next to this integrated sewer-river modelling application, the methodology and CMD software were already employed for various other applications. A coupled sewer-river system was created to analyze different urban and river flood mitigation measures (De Vleeschauwer et al., 2014). A conceptual river model at catchment scale is used to control hydraulic structures in real time via model predictive control to reduce flood damage in the Demer basin (Vermuyten et al., 2015). Another conceptual river model was used for flood probability mapping in real time along the Dender River (Wolfs et al., 2012).

## 4   CONCLUSIONS

A modular and flexible conceptual modelling approach was developed to overcome the limitations of conventionally used detailed hydrodynamic models. Processes can be lumped on different scales. Due to the variety of incorporated model structures, the models can emulate complex dynamics such as backwater effects accurately. The parsimonious model structure and very short calculation times allow model interfacing on larger scales, enabling integrated system analyses. A software tool, named Conceptual Model Developer (CMD), was developed to quickly configure the conceptual models. This tool and modelling approach were demonstrated on a case study by modelling two sewer systems and the receiving river. The integrated conceptual model emulates the simulation results of hydrodynamic models precisely, but simulates events 150 000 times faster than the detailed models. Such short calculation time enables optimization of strategies, accounting for uncertainties and risks and performing long term simulations.

## LIST OF REFERENCES

De Vleeschauwer, K., Weustenraad, J., Nolf, C., Wolfs, V., De Meulder, B., Shannon, K. and Willems, P. (2014). *Green-blue water in the city: quantification of impact of source control versus end-of-pipe solutions on sewer and river floods.* Wat. Sci. & Tech., 70 (11), 1825-1837.

Keupers, I., Wolfs, V., Kroll, S. and Willems, P. (2015). *Impact analysis of sewer overflows on the receiving river water quality using an integrated conceptual model.* In: Proc. of the 10th Int. Conf. on Urban Drainage Modelling (UDM2015), Canada, 20-23 September.

Tran, Q.Q. and Willems, P. (2015). *Flexible hydrological modelling - Disaggregation from lumped catchment scale to higher spatial resolutions.* Submitted for review.

Vermuyten, E., Meert, P., Wolfs, V. and Willems, P. (2015). *Using a fast conceptual river model for floodplain inundation forecasting and real-time flood control – a case study in Flanders, Belgium.* In: Proc. of the 21st Int. Conf. On Modelling and Simulation (MODSIM2015), Queensland, Australia, 29 Nov. – 4 Dec.

Wolfs, V., Van Steenbergen, N. and Willems, P., (2012). *Flood probability mapping by means of conceptual modeling.* In: Proc. of the River Flow 2012 Conference, Costa Rica, 5-7 September, pp. 1081-1085.

Wolfs, V. and Willems, P. (2015). *Computationally efficient modelling approach and software for sewer quantity simulations.* In: Proc. of the 10th Int. Conf. on Urban Drainage Mod. (UDM2015), Canada, 20-23 September.

Wolfs, V., Meert, P. and Willems, P. (2015). *Modular conceptual modelling approach and software for river hydraulic simulations.* Environ. Model. Softw.,71, 60-77.

# Development of a computationally efficient urban flood modelling approach

V. Wolfs, V. Ntegeka, D. Murla Tuyls & P. Willems

*Hydraulics Division, Department of Civil Engineering, KU Leuven. Kasteelpark Arenberg 40, BE-3001, Leuven, Belgium.*

ABSTRACT: This paper presents a parsimonious and data-driven modelling approach to simulate urban floods. Flood levels simulated by detailed 1D-2D hydrodynamic models can be emulated using the presented conceptual modelling approach with a very short calculation time. In addition, the model detail can be adjusted, allowing the modeller to focus on flood-prone locations. This results in efficiently parameterized models that can be tailored to applications. The simulated flood levels are transformed into flood extent maps using a high resolution (0.5-meter) digital terrain model in GIS. To illustrate the developed methodology, a case study for the city of Ghent in Belgium is elaborated. The configured conceptual model mimics the flood levels of a detailed 1D-2D hydrodynamic InfoWorks ICM model accurately, while the calculation time is an order of magnitude of $10^6$ times shorter than the original highly detailed model. The proposed models can be used for numerous applications of urban water management requiring fast models, such as early warning systems, uncertainty analyses and optimization, e.g. to determine real-time storage operations.

## 1 INTRODUCTION

Urban flooding causes worldwide significant disruption to society, huge economic losses and imposes serious health risks. The rapidly increasing urbanization, aging storm networks and climate change will increase urban flood hazard and risk. Recent studies on future climate change impacts show that extreme rainfall intensities may strongly increase (Willems et al., 2012; IPCC, 2014) and as a result also the urban drainage flows and flood hazard (Willems, 2013). Increasing the resilience of urban areas to local rainfall-induced floods is therefore a major objective of present and future water management.

Next to the design of adaptation measures to reduce the flood hazard, measures that increase the preparedness of disaster agencies and the self-coping capacity of people becomes more and more important. One of the latter type of actions is the setup of an urban flood forecasting and warning system. For forecasting of floods in urban areas as a result of extreme rain storms, called pluvial floods, the forecasting has to focus on extreme convective rain storms. Given their short duration but also because of the quick responses of urban drainage systems to rainfall, forecasting of convective rain storms and related pluvial floods can only be done for short lead times. This type of forecasting is called "nowcasting".

One such operational nowcasting system is the operational Short-Term Ensemble Prediction System (STEPS), which was originally co-developed by the UK Met Office and Australian Bureau of Meteorology (Bowler et al.,2006). The system has recently been adapted for Belgium (STEPS-BE). STEPS-BE is based on a 4 C-band radar composite as input and provides rainfall forecasts at high resolution (1km/5min) with 20 ensemble members and a lead time of up to 2 hours (Foresti et al.,2016). As the forecasts are updated every few minutes, it is essential that pluvial inundation models are developed to suit operational forecasting requirements. More specifically, fast models are required such that the real-time flood impact simulations can be conducted every few minutes. Because uncertainties may be significant, especially for longer lead times, such uncertainties also have to be accounted for in the forecasting system. Since such quantification requires multiple runs, e.g. simulating the impacts of ensemble rainfall predictions, incorporating models with very short calculation times becomes even more important. In addition, the models need to be accurate, and "integratable" in existing systems. Given the specific characteristics of urban floods, high resolution models moreover are required to represent topography, the presence of houses, inlets, etc.

Various model types are combined and used to simulate urban floods (see e.g. Henonin et al. (2013) and René et al. (2014) for a comprehensive discussion). Although physically-based white box models are the most popular model type in urban drainage modelling (Parkinson and Mark, 2005), coupled 1D-2D models are commonly applied for urban flood studies (e.g. Schmitt et al., 2004; Carr & Smith, 2006). In such a setup, the underground drainage network is modelled in 1D, while the surface flow model is 2D. Such detailed 2D approach is necessary to accurately represent the full city surface. However, a 1D-2D approach is computationally very expensive and therefore does not meet the requisite functional model characteristics of flood forecasting systems and many other urban flood applications requiring fast models.

This paper tries to overcome this issue by introducing a new and parsimonious modelling approach for simulating urban flood water levels that is compatible with the "CMD" framework for efficient hydraulic modelling of sewers (Wolfs & Willems, submitted). A new module is introduced to emulate urban flood levels generated by highly detailed 1D-2D hydrodynamic models. The predicted water levels are translated to flood maps by means of GIS procedures using data from a high resolution digital terrain model (DTM). The approach is highly flexible and allows the user to focus on flood-prone areas. The proposed method and framework can deal with different temporal and spatial scales, and can for instance use (weather) radar data as input. A similar approach was already developed for river systems (Wolfs et al., 2015) and applied for various applications (e.g. De Vleeschauwer et al., 2014; Vermuyten et al., 2014; Wolfs & Willems, 2015).

To illustrate the developed approach, it was implemented for the case study of the sewer system of the city of Ghent, Belgium. It starts from a highly detailed 1D-2D InfoWorks ICM model to simulate urban floods at high spatial resolution. A fast conceptual hydraulic model was identified and calibrated to that detailed model. The conceptual model will then be applied in a surrogate, complementary way. It aims to accurately emulate water level simulation results of the detailed ICM model. Water level simulation results of the conceptual model are used as input for the GIS-DTM based flood mapping. The derived conceptual model is several orders of magnitude faster than the original 1D-2D hydrodynamic model, and is ideally suited to be used in real-time flood forecasting systems.

First, the methodology is presented, followed by a discussion of the case study area and available data. Next, the results of the case study are presented and discussed. Finally, conclusions and a view on future developments are given.

## 2 METHODOLOGY

The main objective of this research is to develop and test fast models that can accurately calculate instantaneous urban flood levels and visualize the flood extent. Simulating urban flood levels is very complicated due to the large number of interacting processes. Flood levels mainly depend on the hydraulic state of the subsurface urban drainage network, the topography, and interactions between the underground system and surface. Hence, in order to accurately calculate urban flood levels, the surface flow should also be simulated, which in turns necessitates spatially highly detailed maps of the topography. However, simulating such detailed models is computationally very expensive. Therefore, we suggest using mechanistic, fast and parsimonious conceptual models that aim to emulate urban flood level data sets. By emulating such data sets, the surface runoff does not have to be simulated explicitly, since the effect of the surface flow is inherently present in the data set. This strongly limits the calculation time, making these models suitable for flood forecasting systems and other applications requiring fast models. A three step approach is proposed to set up such models. Figure 1 illustrates the proposed methodology.
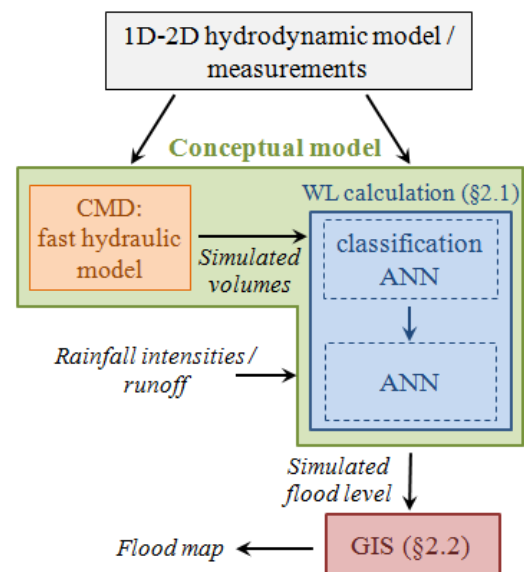


Figure 1: Schematization of the modelling approach.

In the first step, accurate flood level measurements are gathered to configure the model. Due to the lack of accurate level measurements, simulation results of hydrodynamic models are used in this study to develop and test the approach. The use of these "virtual" measurement sets is often advocated in literature to compensate for such data shortage in urban drainage modelling (e.g. Vaes, 1999; Meirlaen et al., 2001). Since urban flood levels are determined by fine-scale processes and due to the close interaction with the underground drainage network, highly detailed 1D-2D white box model must be employed to ensure realistic urban flood levels are obtained.

In the second phase, a parsimonious and fast model is configured to simulate the urban flood levels at selected locations. This model is set up using the simulation results obtained in the previous step. Note that the use of simulation results instead of in situ measurements to calibrate and validate the model does not put stringent limitations on the use and accuracy of the developed approach. Due to the data-driven character of conceptual emulation models, accurate and a sufficient amount of data that cover a wide range of system dynamics are essential to ensure proper configuration of these models. The presented approach is not tailored to emulate results generated by a specific model type or software package, but aims to be generically applicable.

Finally, the urban flood level predictions are translated to flood extent maps using GIS procedures.

The following sections elaborate on the second and third steps of the methodology.

## 2.1 *Fast model to simulate urban flood levels*

The proposed approach tries to emulate urban flood levels using a fast parsimonious model that does not model the surface flow during floods explicitly. By obviating such costly calculations, the calculation time can be significantly reduced compared to full hydrodynamic 1D-2D models. In order to configure such parsimonious models successfully, three critical criteria should be met. Firstly, the flood level set used to calibrate and validate the model should be accurate and realistic. The set should thus inherently account for the complex processes that influence the water level. Secondly, there must be a functional relationship between the sought flood levels and several states in the system that can be calculated accurately. Thirdly and finally, this relationship should be identified and parameterized using one or more model structures.

By using simulation results of a hydrodynamic model, it is assumed that the first condition is satisfied. The detailed 1D-2D hydrodynamic model should be able to generate sufficiently accurate flood levels. A careful selection of several system variables is necessary to comply with the second condition. The hydraulic state of the sewer system, such as the degree of filling prior to a storm event, has a major impact on floods. It is crucial to account for such (antecedent) conditions. It seems plausible to assume a dependency between the volume of part of the sewer system and the flood level at a selected location. To simulate such volumes quickly and accurately, the conceptual hydraulic modelling framework developed by Wolfs & Willems (submitted) is used. This framework is denoted as "CMD" and can account for various elements that influence the flow and thus volumes in the sewer system, including backwater effects, reverse and pressurized flows and

controllable structures such as pumps. Thus, the volume of part of the sewer system is used as first input. It is important to note that CMD allows varying levels of model detail. Thus, the modeller could use the volume of a single conduit up to the volume of an entire sewer system as input. The most appropriate level of lumping is determined using a trial-and-error procedure. Rainfall remains the main driver of urban floods. Therefore, the rainfall intensity (averaged over the response time of the system) is used as second input variable for simulating the flood level. Alternatively, the rainfall runoff flow could also be used as input, since such flow already accounts for antecedent surface conditions. Rainfall runoff flows are calculated in CMD using the Wallingford Model, which is by default used in the hydrodynamic InfoWorks software (Innovyze, 2014). Hence, it is assumed that there exists a functional relation between the two selected inputs, namely the averaged rainfall intensities and volume in selected areas of the sewer system, and the urban flood level at a specified location.

Finally, a model structure must be identified and calibrated to ensure that flood levels can be simulated using these two selected inputs. Given the large number of interacting processes, the use of machine learning techniques is proposed. These techniques are very flexible and can configure themselves using supervised learning to emulate complex data sets. The modelling approach incorporates a serial connection of two artificial neural networks (ANNs). The first is a binary classification ANN to determine if there is urban flooding at the selected location. This ANN has a single hidden layer with an adjustable number of nodes and uses the two selected inputs. It relies on sigmoid and softmax transfer functions in the hidden and output layers respectively, which are trained via a scaled conjugate gradient optimization to minimize the cross-entropy. This set-up improves classification performance. The outcome of the ANN is the probability at every time step during a simulation of having urban floods (i.e. the water level surpasses the ground level). If this probability exceeds a specified threshold, the second ANN is triggered to calculate the flood level. This second ANN uses a sigmoid and purely linear transfer functions in the hidden and output layers and are trained via Levenberg-Marquardt optimization using Bayesian Regularization and early stopping. To improve generalization and to limit the influence of the initially randomized parameters of the ANN, an ensemble ANN is used by averaging the outcome of several trained ANNs. After configuring both ANNs, the threshold probability that activates the second ANN is optimized by minimizing the root-mean-square error (RMSE). To facilitate and speed up ANN configuration, the set-up procedure was programmed in MATLAB using the Neural Network Toolbox.

This data-driven yet mechanistic approach inherently accounts for the complex dynamics and interactions between the sewer system and surface flow that are present in the calibration set. Given the data-driven character of the approach, having a sufficient amount of data that includes several urban flood events is a requisite for configuring accurate models with good generalization capabilities.

## 2.2  *Flood mapping*

Next to the sewer hydraulic computations, also the 2D surface inundation modelling and mapping are constrained with computational resources. Real-time urban flood nowcasting demands flood mapping techniques that are fast enough. In this regard, researchers and practitioners have developed GIS flood inundation techniques with minimal data requirements (Zhang & Pan, 2014; Sampson et al., 2012). For this study, flood mapping is based on a simplified methodology that generates flood extents in a relatively fast way based on flood levels and topographic information. The methodology is akin to the volume spreading algorithm applied in ISIS-FAST (Néelz & Pender, 2010), albeit with a focus on depth spreading. The steps involved in the methodology are described henceforth.

The first step involves the generation of a detailed topographic map of high resolution (1m or less). Steps should be taken to ensure that errors within the map are corrected to minimize the effects of data errors in the final flood extent maps.

In the second step, catchment zones with depressions (pits) are defined. Pits are those locations where all surrounding cells have flow directions pointing towards them. The identification of pits is done through an iterative approach with the use of parameters for storage volume, area and depth. However, not all pits are relevant as some pits are considered artificial. Therefore validation against observed flood extents or other sources of flood inundation is required. Without data on observed flood extents for urban inundation, results of 1D-2D inundation models help to locate regions that are flood prone and pinpoint locations of depressions. From these locations, better parameters for more realistic pits are estimated.

The third step involves the selection of control points for flood level extension. These locations do not necessarily have to be at pits but should fall within the catchment boundaries. For 1D-2D simulations, water level profiles along sewer pathways in the vicinity of flooded locations provide a suitable basis for selection. For instance, locations where the hydraulic grade line is close to the 2D water depth would be more appropriate for flood extension. For flood extents with more than one catchment or depression, a selection of different locations is done.

The fourth step involves the extension of the water levels based on the DTM. The assumption is that if the ground elevation within a depression is lower than the flood level, then the grid cell is considered flooded if it is topographically connected to the reference flood cell, which is the control point. Repeating this over all grid cells within a depression gives a flood extent map.

# 3  CASE STUDY AND AVAILABLE DATA

The developed approach was applied and tested for the urban drainage system of the village of Oostakker, a district of the city of Ghent in Belgium (Figure 2). A detailed 1D-2D full hydrodynamic model of the sewer network (1D) and the surface (2D) was setup, implemented in InfoWorks ICM. This ICM model covers the entire sewer system of the districts of Oostakker and Sint-Amandsberg. It counts in total 6025 conduits, 182 hydraulic structures (such as pumps, weirs, sluices and orifices) and 5855 manholes. The system releases water to receiving water bodies via 39 outfall nodes. This detailed 1D representation of the sewer network is coupled with a surface inundation model at high resolution based on a unique very high resolution DTM available for the study area. The surface inundation model makes use of nested triangular meshes at different resolutions; these include streets ($3.75 - 15$ m$^2$), high flood hazard areas ($12.5 - 50$ m$^2$) and low flood hazard areas ($75 - 300$ m$^2$). The mesh zones are generated from a 0.5x0.5x0.05m DTM (AGIV, 2015) with buildings defined as no flow zones. Additionally, land use areas are classified according to the dominant surface cover type into water, pervious and impervious areas. The interaction between the 1D underground sewer conduits and the 2D surface was through 2D manholes. A 2D manhole is conceptualized as a weir with a crest level taken as the ground level and a crest length equal to the shaft circumference of the node (Innovyze, 2014). In this way, water is exchanged to the 2D surface when the pressure head at the manhole exceeds the ground level.



Figure 2: Detailed 1D-2D InfoWorks ICM model of the case study area (left) and hydraulic conceptual CMD model (right).

Figure 3 illustrates this approach for a subzone of the study area. The street with the highlighted nodes 'SR07994104' and 'SR07995401' is prone to flooding. The nodes shown in the map are the water level calculation nodes in the InfoWorks ICM model. The node 'SR07994104' represents the highest situated manhole of the sewer system and floods first. From this location, water flows over the surface to other nodes in the street. The node 'SR07995401' represents a manhole which is situated in a depression in the topography. Hence, the flow at this location is both induced by sewer floods through the manhole at that specific location, but also by surface flows from other nodes. By selecting two nodes, it is possible to assess the performance for both types of flooding.
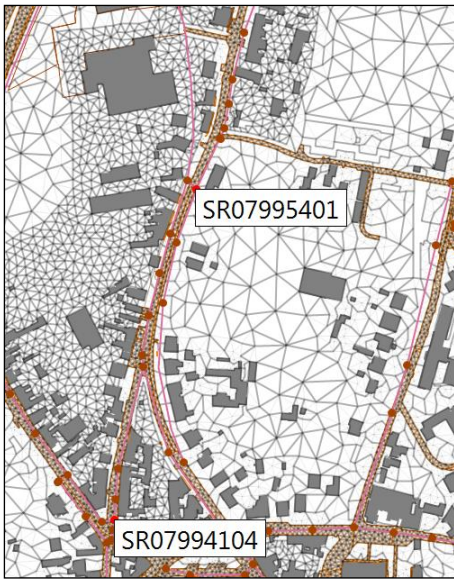


Figure 3: The most flood prone subzone of the study area, indicating the pipes of the 1D sewer network and the calculation nodes in the InfoWorks-ICM model, and the 2D triangular mesh zones. Buildings (grey) are also shown.

Six spatially uniform synthetic storm events with different frequencies of occurrence of 10 (denoted as 'f10') and 7 ('f07') times per year, and return periods of 2, 5, 10 and 20 (respectively 't02', 't05', 't10' and 't20') years were simulated. These six events lead to different flood levels. All models are calibrated for the 'f10', 't02', 't10' and 't20' events, and validated for the 'f07' and 't05' events.

## 4 RESULTS

### 4.1 *Hydraulic model*

A conceptual hydraulic model of the sewer system was configured based on simulation results of the detailed hydrodynamic models of the six synthetic storm events. This conceptual model is configured to simulate the volumes in the sewer system, which are in turn used as input to simulate the flood levels (see

§2.1 and §4.2). The conceptual hydraulic model divides the entire sewer system in six interconnected reservoirs (see Figure 2). The average Nash-Sutcliffe efficiencies (NSE; Nash and Sutcliffe, 1970) for the simulated volumes in these six cells for the calibration and validation events are 0.94 and 0.93 respectively. An NSE of unity indicates a perfect match between the conceptual and hydrodynamic models. Hence, the obtained NSE values indicate that the conceptual model manages to emulate the hydraulics and volumes of the detailed hydrodynamic model accurately. A comprehensive discussion of the calibration and validation results can be found in Wolfs and Willems (submitted).

### 4.2 *Flood levels*

Next, the two ANNs in series are configured for both investigated locations to calculate when flooding occurs and, if relevant, the flood levels. First, the binary neural network is trained according to the procedure outlined in §2.1. Water levels exceeding the flood level are given target values of unity, while others are zero. After training, the obtained ANN is visualized in a grid (see Figure 4a for the binary classification ANN and the target values for location 'SR07994104'). Target values where flooding occurred in the hydrodynamic ICM model are marked in red, while couples without flooding are shown in green. It is obvious that there is a clear segregation of the 2D input space possible into a subspace with no flooding and a subspace where the water level overtops the ground level, leading to floods. Next, the second ANN is configured, aimed to predict the magnitude of the flood levels. Only water levels above ground level are used as training data to ensure that the ANNs can purely focus on emulating the magnitude of the flood levels. An ensemble of four networks is trained. The obtained ensemble of ANNs is then translated to a single entity using simple averaging. By using such an averaged ensemble, the generalization capability is improved, since the variability of each ANNs' response is reduced. Finally, possible negative values in the response domain (thus representing water levels below the surface) are converted to zero. Note that negative values in the ANN's outcome will rarely occur, since the provided training values are strictly possible. Indeed, this second ANN is only being calculated when the first (binary) ANN indicates that flooding occurs. The obtained network for location 'SR07994104' is shown in Figure 4b. The ANN surface response misses only very few targets. In the third configuration step, the threshold probability that is used to determine precisely when flooding occurs is optimized. Note that the training of these ANNs only takes a few seconds due to the use of solely two inputs (volume and averaged rainfall in-

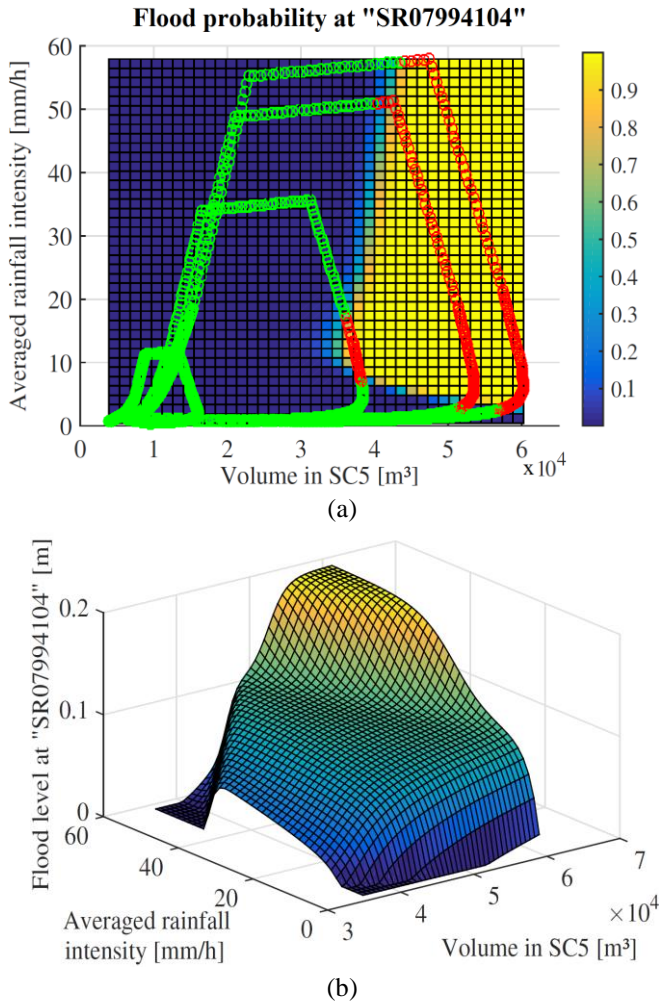tensity) and the low number of nodes in the hidden layer (≤10).



(a)



(b)

Figure 4: Trained ANNs for location "SR07994104"; (a) binary classification ANN which calculates the probability of flooding; (b) ANN which simulates the flood level in case the binary classification ANN assesses that flooding occurs.

Figure 5 shows the simulated flood levels for the four events that lead to flooding. It is clear that the conceptual model manages to emulate the results of the hydrodynamic ICM model very accurately.
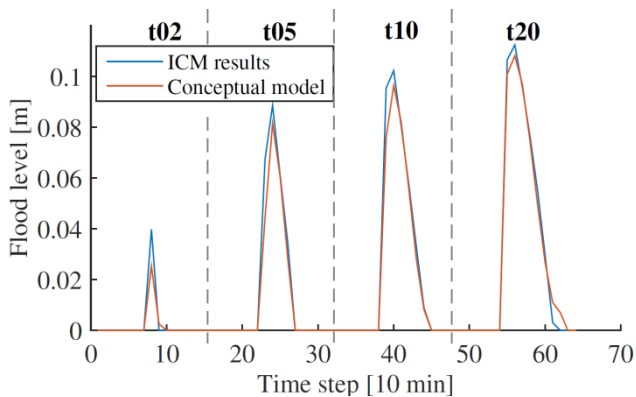


Figure 5: Flood level simulation results of the conceptual and 1D-2D ICM model for location 'SR07994104' for the events leading to flooding.

Table 1 summarizes the RMSE and NSE which are calculated by comparing the simulation results of the conceptual and 1D-2D ICM models. Note that only the 15-minute interval in which flooding occurs is used to calculate both goodness of fit statistics to ensure that only on the period of interest is focused. The results show that the conceptual model can accurately predict the flood level at both locations, although the accuracy of the 't05' event which is used for validation is lower. However, the deviations remain limited as indicated by the low RMSE value.

Table 1: Goodness of fit statistics of the water level simulations at the two selected locations for the different events (C = calibration; V = validation).

|  | 'SR07994104' | | 'SR07995401' | |
|---|---|---|---|---|
|  | RMSE [mm] | NSE | RMSE [mm] | NSE |
| f10 (C) | 0.0 | 1.00 | 0.0 | 1.00 |
| f07 (V) | 0.0 | 1.00 | 0.0 | 1.00 |
| t02 (C) | 2.1 | 0.86 | 6.6 | 0.92 |
| t05 (V) | 3.5 | 0.96 | 33.1 | 0.34 |
| t10 (C) | 2.9 | 0.99 | 6.8 | 0.99 |
| t20 (C) | 2.0 | 1.00 | 13.0 | 0.97 |

### 4.3 Flood extent

Flood extent maps were generated for the flood depth computed by the conceptual model for location 'SR07994104'. However, as the depth for the 2-year event was too shallow (less than 0.05m) it was not extended because the DTM vertical accuracy is around 0.05m. This implies that spreading flood depths close to 0.05m is not feasible. From Figure 6, it is evident that flood extents for the different return periods are almost indistinguishable which is explained by the close flood peak depths shown in figure 4. It is also apparent that GIS-based flood extents tend to be higher than the flood extents computed by InfoWorks-ICM, but the difference is small. Considering the 20-year flood extent around node 'SR07994104', the GIS-based flood extent is only about 8% higher than the flood extent by the InfoWorks-ICM. Highest differences are located in the low–lying areas downstream of the control node.
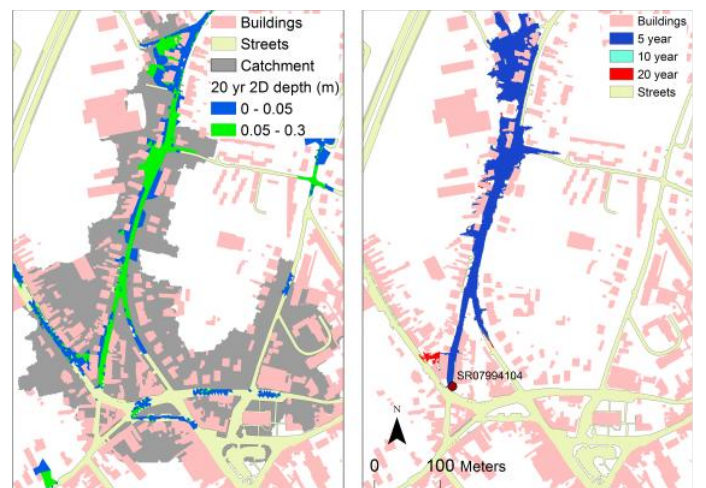


Figure 6: Maximum flood extent simulated by the InfoWorks-ICM for the 20-year storm (Left) and by the GIS-based approach for 5-,10-,20-year storms (Right).

# 5 DISCUSSION AND CONCLUSIONS

A new parsimonious and fast emulation approach was presented to simulate flood levels in urban areas. Application on a case study showed that the approach can successfully mimic the simulation results of a detailed 1D-2D hydrodynamic InfoWorks ICM model. The approach combines three modules: (1) a conceptual model (set up using the "CMD" framework; see Wolfs and Willems (submitted)) to simulate flows and volumes given rainfall intensities, (2) a serial connection of two ANNs to simulate flood levels at specified locations, and (3) a GIS module to visualize the flood extents. This paper focuses on the latter two components.

A connection of two ANNs is configured for each location where flood levels are simulated. The ANNs depend solely on rainfall intensities and the simulated volume of part of the sewer system to calculate the flood level. This volume is simulated by the hydraulic conceptual model (CMD). Given the flexibility of the CMD conceptual model, lumping of areas can range from vary small scales (i.e. the volume is being simulated for a combination of several pipes) up to entire districts (i.e. one volume is simulated for an entire district). Naturally, the applicability of the presented approach to simulate floods depends on the presence of a functional relationship between the simulated volume and rainfall intensities, and the sought flood levels. Given the flexibility of the approach, it is impossible to define a fixed set of crisp criteria to evaluate if such functional relationship exists. Instead, the existence of such relationship should be evaluated in an ad hoc fashion. Note that it is easily possible to use other inputs sets as well (e.g. add other variables to the input set, such as flows).

The results of the case study show that the derived set-up (conceptual CMD model extended with ANNs for flood level calculation) can simulate urban floods up to $10^6$ times faster than the original 1D-2D hydrodynamic model. This vast speed gain is achieved due to the data-driven character of the approach, which obviates detailed calculations of both the flow in the sewer system and the surface flow. To ensure that the data-driven approach can simulate the dynamics of the sewer and surface flow, the data used to set up the models needs to cover both dynamics. In addition, it is crucial to employ a large data set to configure a model with adequate generalization capabilities.

In a second phase, the urban flood levels were transformed to flood extent maps using a simple spreading algorithm in GIS based on the very fine resolution DTM. Flood extent mapping based on such algorithm was found to be a practical alternative to the more detailed 1D-2D full hydrodynamic modelling. Even though detailed flood models are more precise, calculation times remain prohibitive for pluvial forecasting over short time intervals. Nonetheless, comparing the flood extents simulated by the 1D-2D full hydrodynamic model with the GIS-based extents provides insights on selecting important parameters for defining depressions. The GIS-based technique allows for temporal flood maps, which is beneficial for understanding flood propagation. However, flood level spreading does not take into account flood volumes, which often lead to an overestimation of inundated areas especially in flat areas, unless this can be accounted for in the conceptual sewer model. Moreover, water movements and time delays at the surface runoff are not account for in the approach. The quality of the DTM obviously is of critical importance for accurately describing flood extents based on flood depths. In some cases, some infrastructure may appear flooded but in reality they could be above the flood level. It is crucial to stress that GIS-based flood extent maps generated from depths alone are aimed at quick assessment and should be treated as such.

The models configured with the presented methodology can be tailored to the intended application. Due to their flexibility, the model detail can easily be adjusted. For instance, flood-prone or high valuable areas can be modelled with enhanced accuracy, while processes in other regions can be lumped to a greater extent. The obtained models can be used for a wide range of applications that require fast simulation models, such as real-time forecasting systems, and to perform uncertainty or optimization analyses. The modelling approach can easily be semi-automated and merged with the CMD framework, which facilitates and speeds up the conversion of detailed hydrodynamic models into parsimonious and fast conceptual models.

Although the presented approach gives already promising results, additional research remains necessary. Flood extent maps can be generated based on several points for a better assessment of the sensitivity of mapping to different reference locations and to explore potential improvements in flood depth representation. The use of Digital Surface Models (DSMs) would help resolve the effects of flooding close to infrastructure. For instance, elevations of buildings are inherently incorporated in such maps. An approach that makes use of flood volume spreading between depressions would help to address the limitations of flood extents with regards to mapping temporal flood extents. The need for verification of inundation maps against historical flood related data would be important as well.

Clearly, further testing for different sewer systems, land uses and other storm events is necessary to validate and implement requisite improvements to the approach. In a next phase, models configured using the presented methodology will be incorporated in a large-scale urban flood nowcasting test case.

# 6 ACKNOWLEDGEMENTS

# 7 REFERENCES

AGIV, 2015. Digitaal Hoogtemodel Vlaanderen II (2013-2015). https://www.agiv.be/producten/digitaal-hoogtemodel-vlaanderen/meer-over-dhm-v/producten-onderdelen/digitaal-hoogte-model-vlaanderen-ii

Bowler, N.E., Pierce, C.E. & Seed, A.W., 2006. STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society* 132: 2127–2155.

Carr, R.S. & Smith, G.P., 2006. Linking of 2D and pipe hydraulic models at fine spatial scales. *7th International Conference on Urban Drainage Modelling and 4th International Conference on Water Sensitive Urban Design, 2-7 April 2006, Melbourne, Australia*, Vol. 2, pp. 361-368.

De Vleeschauwer, K., Weustenraad, J., Nolf, C., Wolfs, V., De Meulder, B., Shannon, K. & Willems, P., 2014. Green-blue water in the city: quantification of impact of source control versus end-of-pipe solutions on sewer and river floods. *Water Science and Technology* 70(11): 1825-1837.

Foresti, L., Reyniers, M., Seed, A., & Delobbe, L., 2016. Development and verification of a real-time sto-chastic precipitation nowcasting system for urban hydrology in Belgium. *Hydrology and Earth System Sciences* 20: 505-527.

Henonin, J., Russo, B., Mark, O. & Gourbesville, P., 2013. Real-time urban flood forecasting and modelling – a state of the art. *Journal of Hydroinformatics* 15(3): 717-736.

Innovyze, 2014. InfoWorks ICM Help v5.5.

IPCC 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. *Summary for Policymakers, Working Group II contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 44 p.

Meirlaen, J., Huyghebaert, B., Sforzi, F., Benedetti, L. & Vanrolleghem, P., 2001. Fast, simultaneous simulation of the integrated urban wasewater system using mechanistic surrogate models. *Water Science and Technology* 47(7): 301-308.

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

Néelz, S. & Pender, G., 2010. Benchmarking of 2D hydraulic modelling packages. UK Environment Agency, Bristol, UK.

Parkinson, J. & Mark, O., 2005. Urban Stormwater Management in Developing Countries. 2nd edition, IWA Publishing, London.

René, J.-R., Djordjevic, S., Butler, D., Madsen, H. & Mark, O., 2014. Assessing the potential for real-time urban flood forecasting based on a worldwide survey on data availability. *Urban Water Journal* 11(7): 573-583.

Sampson, C.C., Fewtrell, T.J., Duncan, A., Shaad, K., Horritt, M.S. & Bates, P.D., 2012. Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in Water Resources* 41: 1–17.

Schmitt, T.G., Thomas, M. & Ettrich, N., 2004. Analysis and modelling of flooding in urban drainage systems. *Journal of Hydrology* 299: 300-311.

Vaes, G., 1999. The influence of rainfall and model simplification on combined sewer system design. PhD Thesis, Department of Civil Engineering, University of Leuven, Belgium.

Vermuyten, E., Van den Zegel, B., Wolfs, V., Meert, P. & Willems, P., 2014. Real-time flood control by means of an improved MPC-GA algorithm and a fast conceptual river model for the Demer basin in Belgium. *Proceedings of the 6th International Conference on Flood Management. ICFM. Sao Paulo, Brazil, September 2014*, pp. 1-8.

Willems, P. & Vrac, M., 2011. Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *Journal of Hydrology* 402: 193–205.

Willems, P., 2013. Revision of urban drainage design rules after assessment of climate change impacts on precipita-tion extremes at Uccle, Belgium. *Journal of Hydrology* 496: 166–177.

Willems, P., Olsson, J., Arnbjerg-Nielsen, K., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen, H. & Nguyen, V.-T.-V. 2012. Impacts of Climate Change on Rainfall Extremes and Urban Drainage. IWA Publishing, 252 p., Paperback Print ISBN 9781780401256; Ebook ISBN 9781780401263.

Wolfs, V. & Willems, P., 2015. Quantification of impact of retention basins on river floods in the Dender catchment in Belgium using computationally efficiënt models. *36th IAHR World Congress, Delft/The Hague, The Netherlands*, 29 June – 3 July 2015.

Wolfs, V., Meert, P. & Willems, P., 2015. Modular conceptual modelling approach and software for river hydraulic simulations. *Environmental Modelling and Software* 71: 60-77.

Wolfs, V. & Willems, P., submitted. Conceptual modelling approach and software for sewer hydraulic simulations. *Water Resources Management*.

Zhang, S. & Pan, B., 2014. An urban storm-inundation simulation method based on GIS. *Journal of Hydrology* 517: 260-268.

# High resolution radar-rain gauge data merging for urban hydrology: current practice and beyond

Susana Ochoa Rodriguez (1,3), Li-Pen Wang (1,2), Andy Bailey (3,5), Patrick Willems (4), and Christian Onof (1)

(1) Department of Civil and Environmental Engineering, Imperial College London, UK (s.ochoa-rodriguez@imperial.ac.uk), (2) RainPlusPlus Ltd., UK, (3) RPS Water, UK, (4) Hydraulics Laboratory, KU Leuven, Belgium, (5) Pennine Water Group, University of Sheffield, UK

In this work a thorough test is conducted of radar-rain gauge merging techniques at urban scales, under different climatological conditions and rain gauge density scenarios. The aim is to provide guidance regarding the suitability and application of merging methods at urban scales, which is lacking at present. The test is conducted based upon two pilot locations, i.e. the cities of Edinburgh (254 km̂2) and Birmingham (431 km̂2), for which a total of 96 and 84 tipping bucket rain gauges were respectively available, alongside radar QPEs, dense runoff records and urban drainage models.

Three merging techniques, namely Mean Field Bias (MFB) adjustment, kriging with external (KED) and Bayesian (BAY) combination, were selected for testing on grounds of performance and common use. They were initially tested as they were originally formulated and as they are reportedly commonly applied using typically available radar and rain gauge data. Afterwards, they were tested in combination with two special treatments which were identified as having the potential to improve merging applicability for urban hydrology: (1) reduction of temporal sampling errors in radar QPEs through temporal interpolation and (2) singularity-based decomposition of radar QPEs prior to merging. These treatments ultimately aim at improving the consistency between radar and rain gauge records, which has been identified as the chief factor affecting merging performance and is particularly challenging at the fine spatial-temporal resolutions required for urban applications. The main findings of this study are the following:

- All merging methods were found to improve the applicability of radar QPEs for urban hydrological applications, but the degree of improvement they provide and the added value of radar information vary for each merging method and are also a function of climatological conditions and rain gauge density scenarios.

- Overall, KED displayed the best performance, with BAY being a close second and MFB providing the smallest improvements upon radar QPEs. However, as compared to BAY, KED performance is more sensitive to rain gauge density and to the ability of rain gauges to sample critical features of the rainfall field. By incorporating more information from radar than KED, BAY is less sensitive to rain gauge density and to poor rain gauge predictability and proved able to provide a good representation of convective cells even in cases in which gauges completely missed such structures.

- Based on the findings of this study, it is recommended that KED be used when gauge densities are relatively high (of the order of 30 km2 per gauge or higher) and/or when the quality of radar QPEs is known to be very poor, in which case it is desirable to rely more upon rain gauge records. For low rain gauge density situations and QPEs of reasonable quality (as is the case in most of EU), BAY may be a more appropriate choice. MFB should be the last choice; however, it is better than no correction at all.

- The two special treatments under consideration successfully improved overall merging performance at the spatial-temporal resolutions required for urban hydrology, with benefits being particularly evident at low rain gauge density conditions.

**Title: Development and comparison of two urban pluvial flood modelling approaches for real-time use**

María Bermúdez[1,2], Victor Ntegeka[1], Vincent Wolfs[1], Patrick Willems[1]

[1]Hydraulics Section, Department of Civil Engineering, KU Leuven, Belgium

[2]Environmental and Water Engineering Group, Universidade da Coruña, Spain

**Summary** (700 characters)

1D-2D dual drainage models are a well-established approach to simulate urban pluvial floods. However, despite modelling advances and increasing computer power, this approach remains unsuitable for real time practical applications. We thus propose and test two computationally efficient surrogate models of such a dual drainage model. In the first one we couple a 1D sewer model to a GIS-based overland flood network. In the second one we develop a conceptual sewer and flood model using data-driven and physically based structures, and employ pre-simulated scenarios for mapping. The city of Ghent is used as a test case. The surrogate models can provide comparable results to the original model.

**Introduction**

1D-2D dual drainage models are a well-established approach to simulate urban pluvial flooding, which can provide a realistic description of flood flow conditions. They simulate the coupling between the sewer network, represented by a 1D model, and the surface flow, computed using a 2D hydrodynamic model. However, this approach is generally too computationally intensive for real-time applications such as urban pluvial flood forecasting or for probabilistic approaches that require multiple simulations.

Therefore, there remains a need for developing computationally more efficient surrogates of these models. In this context, two broad surrogate modeling approaches can be considered: lower-fidelity physically based models, which are simplified models of the original system, and data-driven models which emulate the original model responses (Razavi et al., 2012). In this study we propose and test two surrogate models of a 1D-2D dual drainage model which combine both modelling approaches. Their capabilities and limitations for pluvial flood inundation mapping are evaluated and compared.

**Material and Methods**

<u>Case study</u>

The city of Ghent was used as a test case. A total of 8 synthetic storms with return periods from 2 to 100 years, 6 rainfall events obtained with a stochastic rainfall generator (Muñoz et al., 2015) and 2 historical rainfall events (occurred on 28th July 2013 and 30th May 2016) were used for calibration and validation of the models.

<u>Numerical models</u>

A detailed 1D-2D full hydrodynamic model of the sewer network and the surface of the city was implemented in InfoWorks ICM (Fig. 1). The model covered an area of 27.5 km$^2$ and counted in total 6025 conduits, 182 hydraulic structures and 5855 manholes. The resolution of the surface triangular mesh ranged from 3.75 m$^2$ up to 50 m$^2$ in the flood prone areas. The interaction between the 1D underground sewer conduits and the 2D surface was through the manholes, conceptualized as weirs.

*Fig. 1. Definition of the sewer network of the case study area in the 1D-2D original model. The region analyzed in Fig. 2 and Fig. 3 is marked in yellow.*

Two different surrogate models of the above model were developed in this work. The first one consists of a 1D representation of the sewer network in which the flood volumes are stored in virtual reservoirs on top of the manholes. A conical flood storage volume was defined for each manhole, being two alternative definitions implemented: a default definition based on the size of the contributing areas draining to each node and an enhanced definition based on the floodable areas estimated from the surface topography. A volume spreading algorithm, which relies on a GIS analysis of the surface topography, was subsequently used to translate the flood water from surcharged manholes into a flood depth map.

The second surrogate model involves a conceptual sewer and flood model which uses both data-driven and physically based model structures. It first emulates the volumes in the sewer system using a conceptual network topology, as described in Wolfs & Willems (2017). It subsequently estimates the surface flood volumes, aggregated in pre-defined regions of the study area (with sizes between 0.25 and 0.5 km$^2$), by means of a serial connection of two artificial neural networks and an infiltration model. Pre-simulated scenarios are used for flood mapping in the different regions.

**Results and Discussion**

The results show that the first surrogate model is sensitive to the definition of the virtual storage reservoirs in the 1D sewer model. When this definition is based on the floodable areas estimated from the surface topography, the maximum flood volumes can be correctly estimated (Fig. 2). This surrogate model is however unable to emulate the behavior of the surface flow, and thus cannot be used to estimate flood duration or to predict the evolution of the volumes in the surface (Tab. 1).

*Fig. 2. Surface flood volume in the region indicated in Fig. 1 for the May 2016 event.*

As illustrated in Fig. 2, the second surrogate model emulates the results of the original model accurately, in terms of flood volumes. It can provide accurate estimations not only of the maximum volume, but also of its evolution during the flood event (Tab. 1). Given that this surrogate model aggregates these volumes into pre-defined regions, it is not possible to identify the surcharged manholes. However, the use of pre-simulated maps based on the above flood volumes can provide a reasonable characterization of the flood depth field at the street level (Fig. 3).

*Fig. 3. Depth fields in the region indicated in Fig. 1 at the time of the peak flood volume for the May 2016 event.*

*Tab. 1. Model capabilities.*

**Conclusions**

The first surrogate model can identify overflow locations and give an estimation of the maximum flood volume. These data can then be used to derive a reasonably accurate maximum flood depth map based on the surface topography. The second surrogate model can emulate the evolution of the flood volumes accurately and can thus be used to predict the dynamics of the flood event. Both models require shorter calculation times (by more than 100 times) than the original model.

**Acknowledgments**

**References**

Muñoz C., Wang L.P. & Willems P. 2015. Towards a high resolution stochastic rainfall generator for urban applications. In Molnar, Peter & Peleg, Nadav (Eds.), Rainfall in urban and natural systems. Proceedings of the 10th International Workshop on Precipitation in Urban Areas.

Razavi S., Tolson B.A. & Burn D.H. 2012. Review of surrogate modeling in water resources. *Water Resources Research*, 48(7), W07401

Wolfs V. & Willems P. 2017 Modular Conceptual Modelling Approach and Software for Sewer Hydraulic Computations. *Water Resources Management*, 31(1), 283-298.
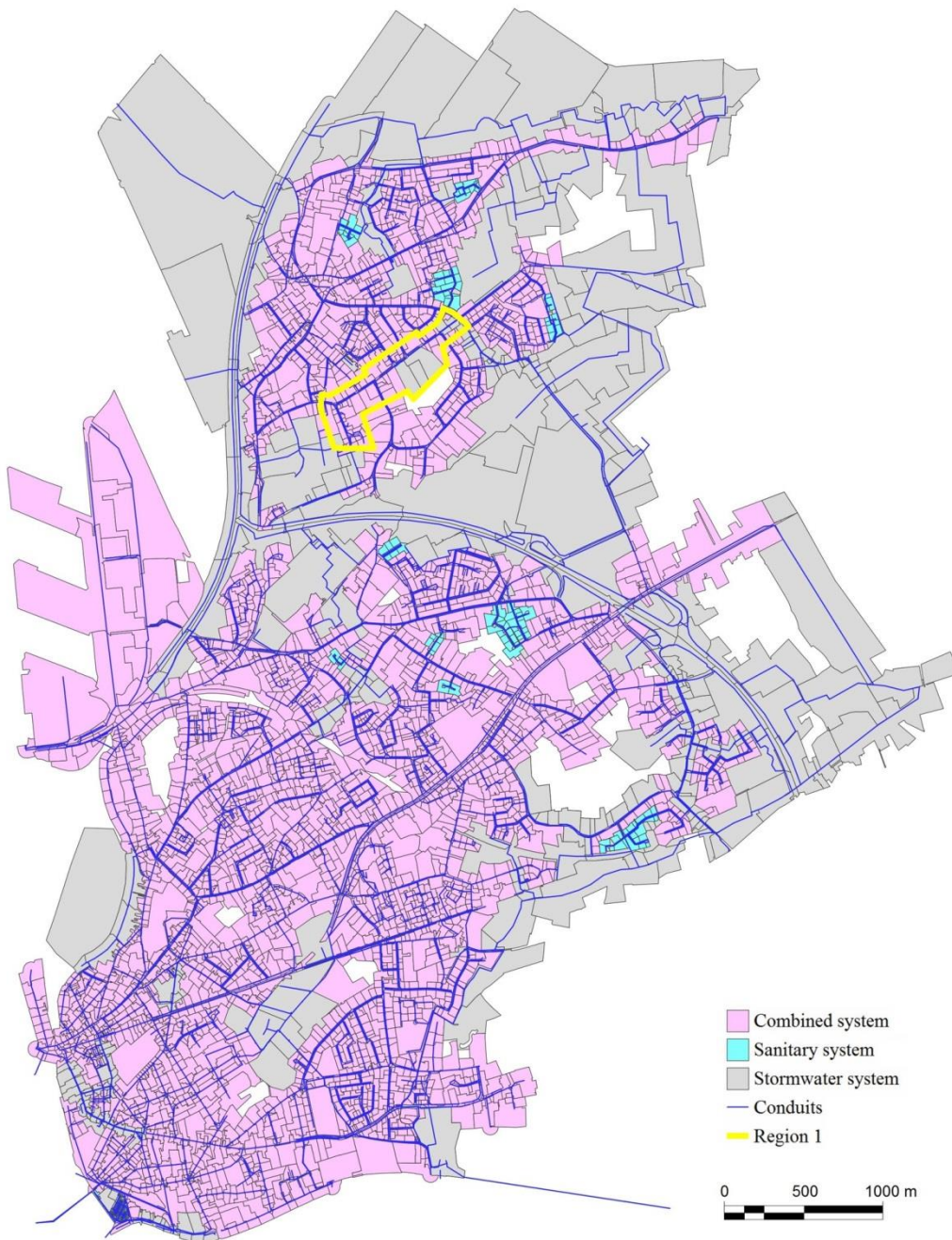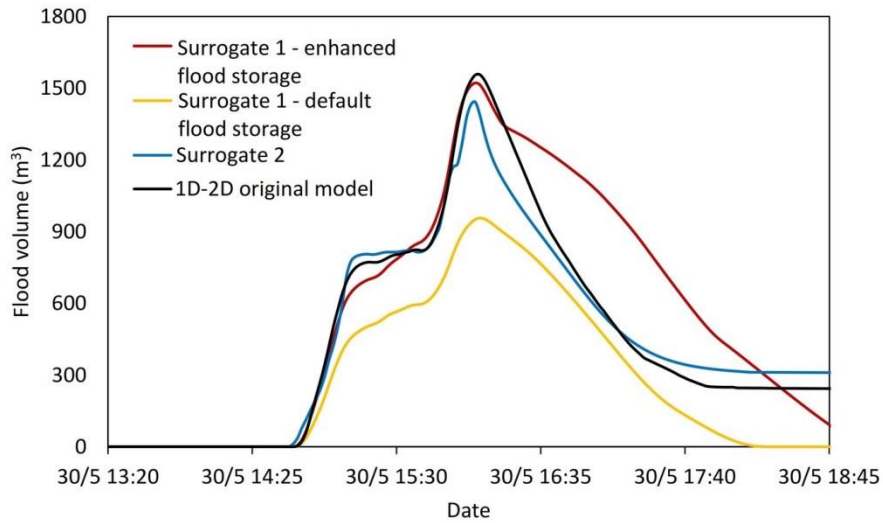
*FIGURE 1*

FIGURE 2



(a) 1D-2D original model

(b) Surrogate 2, pre-simulated scenario

FIGURE 3

TABLE 1

|  |  | Surrogate model 1 | Surrogate model 2 |
|---|---|---|---|
| Flood identification (yes/no) |  | Yes | Yes |
| Flood magnitude (total flood volume) |  | Yes | Yes |
| Spatial flood prediction | Overflow location | Yes | No |
|  | Maximum flood extent | Yes | Yes |
|  | Maximum flood depth | Yes | Yes |
| Flood dynamics | Flood start time | Yes | Yes |
|  | Flood duration | No | Yes |
|  | Flood evolution | No | Yes |
| Computational time (speed gain) |  | $10^2$ | $10^6$ |