Compiled reply to reviewers' and Editor's comments on the manuscript "Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications" by L.-P. Wang et al.

We would like to thank the reviewers and the Editor for their thorough review and constructive comments and suggestions, which have helped improve the manuscript. Below are the questions (Q) raised by the reviewers and the Editor, followed by their individual answers (R).

Since the manuscript was prepared in Latex, it was not possible to track changes. Instead, we have used RED FONTS to denote all changes made to the original manuscript (these are also described in detail in the present file).

The 'marked up' document in HESS-Discussion format has been included at the end of the present file. Note that we were required to upload the manuscript in HESS –double column– format to the website. Since the comments given by reviewers were based on the HESS-Discussion (HESSD) format (hence line numbers and the like would not match to the new HESS format), we decided to include the marked up document in the original HESSD format as part of the present file.

Reviewer #1 (Rev1) (Dr. Sinclair)

General comments:

The authors present a technique for merging weather radar and rain gauge data which aims to preserve the spatially variable character of the radar rainfall while simultaneously reducing the bias between radar and gauges. They assess the technique by evaluating the estimated rainfall against alternative merging/adjustment methods. In addition they evaluate modelled flows from a small urban catchment. My overall impression is that the paper is well written and presents an interesting and useful method of rainfall data merging which is new to me (although it builds on ideas of separating scales of variability in radar rainfall that have been explored previously). I think that the paper can be published essentially as is, subject to some clarifications in the description of the singularity extraction and recovery procedure in sections 2.2 and 2.3.

(R): We would like to thank the reviewer for taking the time to review this manuscript and for the positive and constructive comments about the proposed data merging method.

(Rev1-Q1) It's unclear to me how the initial unknown constant value $c(\mathrm{x})$ and singularity index $\alpha(\mathrm{x})$ are computed from eqn (3) before applying the iterative procedure described. The authors discuss a spatial-scale range on page 11 (line 6 and following), so I assume that the two unknowns are fitted based on data from "a small number of data samples" from radar grid cells in the neighbourhood of each target grid cell? I think this requires clarification.

(R1): The following explanation has been incorporated in the manuscript and also in the supplementary document (i.e. document with detailed description of the formulation and demonstration of the proposed method).

"A schematic of the estimation of the constant value $c(\mathbf{x})$ and singularity index $\alpha(\mathbf{x})$ from real gridded data is provided in Figure 3 *(figure number assinged according to its position in the manuscript)*. For a given pixel with centre **x** (centre of top plot in Figure 3), the 'mean' rainfall intensities at different spatial scales (centred in **x**) can be calculated (i.e. rainfall intensities ρ_1, ρ_2 and ρ_3, respectively at scales ϵ_1, ϵ_2 and ϵ_3). Then, the logarithms of these mean values and the associated spatial scales are compared (bottom plot in Figure 3). The constant value $c(\mathbf{x})$ and singularity index $\alpha(\mathbf{x})$ of the dataset 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 $(\alpha(\mathbf{x}) - Euclidian\ Dimension)$ and $\log c(\mathbf{x})$. A detailed explanation of the computation of $c(\mathbf{x})$ and $\alpha(\mathbf{x})$ can be found in previous studies (Agterberg, 2012; Chen et al., 2007; Cheng et al., 1994). It is worth mentioning that the estimation of $c(\mathbf{x})$ and $\alpha(\mathbf{x})$ can be trusted only if a good linear relation is observed (i.e. if the scaling behaviour is well followed). In the example shown in Figure 3, three scale ranges (i.e. l_1, l_2 and l_3) are used to derive $c(\mathbf{x})$ and $\alpha(\mathbf{x})$. In real applications, the modeller must choose a suitable number of scale ranges to conduct this analysis, bearing in mind that a small number of scale ranges may result in increased uncertainty in the estimation of $c(\mathbf{x})$ and $\alpha(\mathbf{x})$, but a too large number of scale ranges may result in loss of local features present in geo-fields.

In this study, the scales of interest ranged from 1 to 9 km, which results in a total of five rainfall intensity samples (at scales 1, 3, 5, 7 and 9 km) to fit the regression line at each radar pixel location. Based upon our analyses, a good linear behaviour was generally observed within this scale range. In addition, valuable small-scale structures were indeed preserved in the resulting rainfall product. As such, the selected spatial scale range was deemed to represent a good balance between estimation uncertainty and local feature preservation."

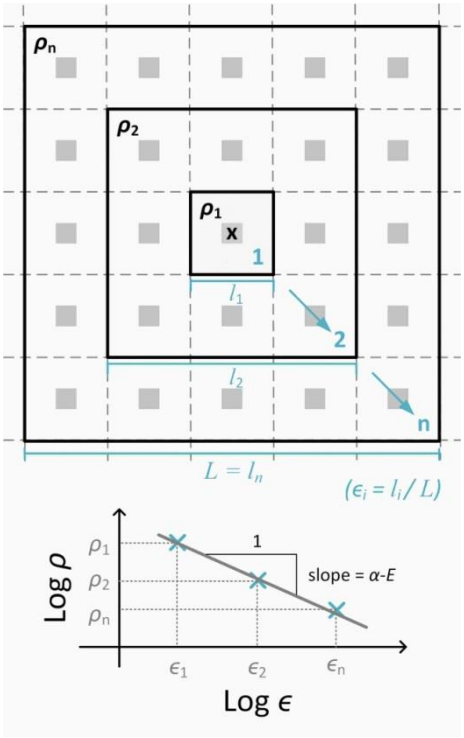

Figure 3: Schematic of the local singularity analysis (Wang et al., 2014).

(Rev1-Q2) Is the singularity map simply the difference between the original radar field and the computed c*(x) field? I think that it should be more explicit if this is the case and would appreciate more detail on how it is proportionally applied back to the NS-BAY field.

(R2): The singularity map is composed of the iteratively-derived singularity exponents $\alpha^*(\mathrm{x})$ (see eqn (6) of the manuscript for the definition). Based upon it, a ratio at each radar pixel location can be derived:

$$r(\mathrm{x}) = \epsilon^{\alpha^*(\mathrm{x})-E},$$

which corresponds to the ratio difference between the original radar field (RD) and the non-singular radar field (NS-RD). After the merging between NS-RD and BK rain gauge fields is conducted, each pixel location of the resulting NS-BAY field is multiplied by the above ratio, thus 'restoring' the original singularity structures present in the RD field.

The description of this aspect has been improved in the revised manuscript and in the supplement, as follows:

"Afterwards, the singularity map is applied back and proportionally to the NS-BAY merged field (i.e. step (ii) in Figure 1(b)), thus yielding a singularity-sensitive merged field (SIN). This is done by multiplying each pixel value of the NS-BAY field by the following ratio:

$$\mathbf{r}(\mathbf{x}) = \boldsymbol{\epsilon}^{\boldsymbol{\alpha}^*(\mathbf{x})-\mathbf{E}}, \qquad (1)$$

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."

Detailed and editorial comments:

(Rev1-Q3) Page 8, line 24 - References error in my version of the PDF.
(R3): This has been corrected in the revised manuscript.

(Rev1-Q4) Page 9, line 5 - Agterberg, 2017?
(R4): It should be Agterberg, 2007. This has been corrected in the revised manuscript.

(Rev1-Q5) Figure 2 - Circled flow gauges FM * don't match the caption text, missing section number at end of caption.
(R5): These should be gauges FM 3, 10, 14 and 19. This has been corrected in the revised manuscript.

(Rev1-Q6) Page 14, line 19 - Only one of the 4 storms chosen in the analysis were not used for calibration of the flow model. Discussion in section 3.3.2 should proceed with this in mind.
R6: As suggested by the reviewer, this has been highlighted in Section 3.3.2. (see document with marked changes). There were indeed some differences between the hydraulic performance in the events used and not used for model calibration. As can be seen in Fig 10, the performance measures for the hydraulic outputs for Storm 4 were in general lower than for the other three events.

However, the general features and relative performance of the different rainfall inputs was similar 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.

It would have been ideal to focus only on events different from those used for model calibration. However, as specified in Section 3.1.4, the measurements used in this paper were from a flow survey which lasted only three months. During this period only a very limited number of 'significant' events occurred and these were selected for model calibration (as these were the only ones which fulfilled the UK standards for calibration and verification of urban drainage models). Since there were not many other significant events during the 3-month period, we had to adopt the 'calibration events' for our analysis.

(Rev1-Q7) Page 16, line 20 - How is the areal rain gauge estimate calculated? Is it different from the block-kriged gauge estimate?

(R7): In this paper, we used the average of the 'point' rain gauge records to represent the areal rain gauge estimates. Due to the fact that the rain gauge network used in this study is of high density, we believe that this point average is a good approximation of the areal rain gauge estimate.

In theory, this point average should be equivalent to the areal averages obtained from the block-kriged rain gauge estimates because (block-) kriging is an unbiased estimator. However, according to the statistics given in Table 2, we can see some minor difference between these two estimates. This could be due to the following two reasons:

1. In order to obtain 'block' estimates using point rain gauge records, it is necessary to conduct 'integration' of the semi-variogram model over the block area (i.e. radar pixel square). In this paper, the Gaussian variogram model was employed and its integration (i.e. the error function) was numerically calculated, so some numerical round-off errors may have appeared during the block-kriging process.
2. In order not to obtain negative kriged estimates, an addition constraint (i.e. the weights of linear combination cannot be negative values) was imposed into the kriging system. This converted the original kriging into a quadratic programming (optimisation) problem (Barnes and Johnson, 1984). Similarly, this optimisation process was done numerically, which also brought about some numerical rounding errors.

(Rev1-Q8) Page 17, lines 8-10 - Is the regression forced to pass through zero?

(R8): The regression was not forced to pass through zero; however, the y-axis intercepts we obtained were usually close to zero.

(Rev1-Q9) Fig 7 - It's tough to read the hydrographs, too small, too busy

(R9): To make this figure less 'busy', it has been split into two separate figures:

Figure 7 shows observed vs. simulated hydrographs resulting from the SIN only rainfall inputs (including all five singularity ranges). This allows better visualisation of the impact of the singularity ranges.

Figure 8 shows observed vs. simulated hydrographs resulting from different rainfall inputs, including rain gauges, MFB, BAY and only SIN3 (SIN3 corresponds to the intermediate singularity range, which

covers most singularity exponents and, as explained in the manuscript (Section 3.3.1), is considered to be an appropriate range).

Besides, it is important to note that one of the reasons why this figure looks too small and busy is because in the HESSD format, pages are shorter than normal. As a result, the size of this particular figure was reduced in order to fit the shorter page. However, in a normal (A4) size page (as is the format of the final papers in HESS), the figure will be much bigger and therefore less busy and clearer.

(Rev1-Q10) Page 30, line 18 - In addition to Wavelets consider data driven techniques like Empirical Mode Decomposition, PCA etc.

(R10): Thank you for this suggestion. Other data-driven techniques, including Principal Component Analysis and Empirical Mode Decomposition, can indeed be used for singularity identification and extraction. These have been included in Section 4 (Conclusions and Future Work).

(Rev1-Q11) What are the BAY merging artifacts in fig 3 (c1) (e1) from the supplement? Very sharp discontinuities in regular blocks.

(R11): The data used to produce the BAY and NSBAY images in Fig 3 (c1) and (e1) from the supplementary files were generated using the RainMusic software tool. Some thresholds (or rules) might be given in the tool in order to ensure numerical stability. These may have caused some artefacts in the images. Because it is a commercial software, we had no access to the source code to find more evidences of this.

Nonetheless, as part of this research, a C/C++ tool has been implemented to replicate the function of the RainMusic software package, with the difference that the new tool has been customised for urban applications. As can be seen from the images produced by the new tool (see Fig 6 in the manuscript), the problem of artifacts have been resolved.

References

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

Barnes, R. and Johnson, T.: Positive Kriging, in Geostatistics for Natural Resources Characterization SE - 14, edited by G. Verly, M. David, A. Journel, and A. Marechal, pp. 231–244, Springer Netherlands., 1984.

Chen, Z., Cheng, Q., Chen, J. and Xie, S.: A novel iterative approach for mapping local singularities from geochemical data, Nonlinear Process Geophys., 14, 317–324, 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.

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), Washington DC., 2014.

Reviewer #2 (Rev2) (Dr. Mazzetti)

General comments:

The authors present a technique for merging weather radar and raingauge data that is suitable for application in small urban catchments. Validation of the proposed technique is performed through the comparison between estimated rainfall, raingauges and alternative merging techniques. Finally, results from urban flow simulations are compared. The paper is well written, concepts and application are clearly described, results are presented in a straightforward way.

(R): We would like to thank the reviewer for taking the time to read the manuscript and for the constructive comments.

The subject of rainfall data merging is not original or unexplored, but the paper has the advantage of providing an example of coupling data merging and urban flow simulations.

(R): Indeed, as the reviewer indicates, rainfall data merging is not an unexplored subject. However, its applications to urban hydrological modelling are rather limited and mainly reduced to the testing of existing merging techniques (that have been widely used in large-scale hydrological applications). This paper therefore focused on developing a merging methodology that is especially suitable for urban-scale applications by integrating existing merging techniques (mostly resting upon Gaussian assumption) with the local singularity (or fractal) analysis (Cheng et al., 1994).

It is worth mentioning that the singularity-sensitive method is not limited to the use of Bayesian data merging (Mazzetti and Todini, 2004; Todini, 2001a, 2001b). In fact, the proposed methodology can also be integrated with other existing geostatistical-based merging techniques, such as Kriging with External Drift, Conditional Merging and Cokriging. The reason we chose the Bayesian data merging is because it has shown to outperform many other merging techniques at urban scales (Ochoa-Rodríguez et al., 2013a, 2013b; Wang et al., 2013).

(Rev2-Q1) One of the weak points in the paper is the fact that, raingauge precipitation estimates are assumed to be a better representation of the "true" rainfall field and used as a reference to evaluate the performances of the different areal gridded rainfall estimates. In this way, rainfall data merging that gets the better results is the one that better represents raingauge estimates, which in fact is not the true rainfall field, because of the 1) raingauge point measurement errors and 2) the extrapolation technique used to extend the point measurement to the grid average. Although raingauge density is an important factor in assessing performances of rainfall estimates, it is not the only one and in general, the availability of a dense raingauge network is necessary, but not sufficient condition to provide a good approximation of the actual rainfall field, particularly during intense events. Moreover, raingauge point measurements are affected by errors, particularly due to wind, which tend to introduce a negative bias. A small discussion on how to deal with this issue would be of interest, particularly for the high intensities of short duration rainstorms that affect urban environment.

(R1): We fully agree with the reviewer in that it is not ideal to 'over-trust' rain gauge rainfall data because they are also affected by a number of error sources (e.g. the effect of wind, evaporation, splashing, discalibration, amongst others). In this work, rain gauge records were obtained from tipping bucket rain gauges. According to the literature (Luyckx and Berlamont, 2001; Molini et al., 2005), the most critical sources of error from this type of rain gauges are wind-induced losses and loss of water during the tipping action, which is strongly dependent on rainfall rate. Errors due to the former may amount to 2-10%, depending on wind speed, weigh of precipitation and gauge construction parameters) and errors due to the later may amount up to 10 % for rainfall intensities of 100 mm/h and 20 % for intensities of 200 mm/h in non-mountainous areas. Because the storm events studied in the work were not extremely heavy and some basic treatment had been applied to ensure the quality of rain gauge measurements (e.g. manual comparison between neighbouring rain gauge data), we worked with the assumption that the errors due to the loss of water during tipping action and other sources were not significant. Therefore, the rain gauge data could be an acceptable reference for the comparison, although it is not perfect. This had already been briefly mentioned in Section 3.2. (Evaluation Methodology), as follows:

"Both evaluation strategies [rainfall analysis and hydrological testing] have advantages and disadvantages. 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)."

However, following the reviewer's comments, this issue has been further discussed in Section 3.2. The following paragraph was added:

"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. In order of general importance, systematic errors common to all rain gauges include errors due to systematic 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 (Habib et al., 2008). The order of magnitude of the two main error sources associated to tipping bucket rain gauges is 2-10 % for wind induced losses (depending on wind speed, weight of precipitation and gauge construction parameters) (Sevruk and Nespor, 1998) and up to 10 % for rainfall intensities of 100 mm/h and 20 % for intensities of 200 mm/h, for water loss 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."

(Rev2-Q2) Regarding the urban flow model, I think it is a valuable tool, but results must be evaluated carefully. In general it is not a good idea to evaluate the performances of the rain field estimates downstream an additional model, because it involves calibration and introduces additional uncertainty elements. The model was calibrated using raingauge data, therefore it has been "instructed" to provide the best results with that type of input information. No wonder that it provides the best results using outputs from the merging technique that was better reproducing raingauges, but still this is not a valid proof that the technique is also the most appropriate to conveniently represent the true rainfall field. As a paradox we might get worse results using the true rainfall field as input to the urban flow model, than the ones obtained using the raingauges based estimate utilized for calibrating the model. Moreover using an urban flow model to evaluate performances of rainfall merging introduces more modelling uncertainty (model approximation, parameter calibrated values, initial and boundary conditions, etc) into the process making it difficult to distinguish rainfall "errors" from model "errors". It can be acknowledged that is difficult to find a way out to the problem of "true" rainfall (unless via extensive simulation with synthetic data) and the authors clearly indicate the limits of their assumptions in the paper. For this reason I think the paper can be published.

R2: It is indeed difficult to find the best way to evaluate rainfall results. As mentioned by the reviewer, the use of urban hydraulic modelling will inevitably introduce additional errors. However, it provides an alternative to compare different rainfall estimates, which can complement the results of other analyses (e.g. comparison against rain gauge measurements) which also entail errors. Another important consideration is that this comparison serves the purpose of identifying the best rainfall estimates for urban drainage applications: some discrepancies in the rainfall may not turn out to be relevant to the generated flows, and we need to know if and when that is the case.
In addition, because the model was calibrated using rain gauge data, merged rainfall estimates (i.e. BAY and SIN, in this work) could be indeed expected to result in better hydraulic output than the original radar rainfall estimates (i.e. RD). However, if we only focus on the hydraulic performance of the BAY and SIN estimates (and assume that the model errors remain similar), we can identify the different features of these two rainfall inputs and this gives useful insights regarding the benefits of employing the proposed methodology.

As the reviewer indicated, the limitations associated to the hydraulic verification have been clearly stated in the paper.

Detailed and editorial comments:

(Rev2-Q3) Page 1868 lines 7: It could be interesting using radar estimates with higher spatial resolution, for example provided by an X-band RADAR.
R3: Indeed, it would be very interesting to conduct tests with higher resolution datasets, especially given that the proposed method aims at improving rainfall estimates in (small-scale) urban areas. This has been mentioned in the 'Conclusions and future work' section.
Work in this direction is already ongoing as part of the project in which we work (Interreg IVB NWE RainGain: http://www.raingain.eu/).

(Rev2-Q4) Page 1870, lines 17-22: Density and coverage of rain-gauges are not the only factor to be considered. Rain-gauges may have errors themselves.

R4: As mentioned by the reviewer previously in the general comments, the rain gauge measurement errors are also very critical apart from the coverage and density of rain gauge network. This has been better discussed in the revised manuscript.

(Rev2-Q5) Page 1873, lines 13-15: Radar should in principle be better than raingauge network in capturing rainfall dynamics. If not, I would investigate also if it might depend on space or time resolution, type of radar, type of corrections.

R5: What we meant to say here is that radar rainfall estimates fail to well capture instantaneous (dynamic) rain fall rates. As explained between page 1857/line 21 and page 1858/line 8, this is due to the way in which radar reflectivity is generally converted to rainfall rate, using a single Z-R relationship. This is insufficient to deal with dynamic changes in storm type and associated changes in raindrop size distribution. It is true that the term 'dynamics' that we used may cause some confusion, so we have rephrased this sentence as follows:

"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."

(Rev2-Q6) Fig. 7: I would suggest to find a better representation for hydrographs. Different signs and colors are not enough to distinguish different lines, even if you improve quality of the picture. Maybe you should consider to make a different figure for each hydrograph.

(R6): A similar concern was raised by Reviewer #1 – see Rev1-Q9. As indicted above, this figure has been split into two separate figures.

References

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.

Luyckx, G. and Berlamont, J.: Simplified method to correct rainfall measurements from tipping bucket rain gauges, in Urban Drainage Modeling, pp. 767–776., 2001.

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.

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(5), 1073–1088, doi:10.1002/hyp.5646, 2005.

Ochoa-Rodríguez, S., Rico-Ramirez, M., Jewell, S. A., Schellart, A. N. A., Wang, L., Onof, C. and Maksimović, Č.: Improving rainfall nowcasting and urban runoff forecasting through dynamic radar-raingauge rainfall adjustment, 7th Int. Conf. Sewer Process. Networks, 2013a.

Ochoa-Rodríguez, S., Wang, L., Simoes, N., Onof, C. and Maksimović, Č.: On the possibility of calibrating urban storm-water drainage models using gauge-based adjusted radar rainfall estimates, 7th Int. Conf. Sewer Process. Networks, 2013b.

Todini, E.: A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, Hydrol. Earth Syst. Sci., 5(2), 187–199, 2001a.

Todini, E.: Influence of parameter estimation uncertainty in Kriging: Part 1 - Theoretical Development, Hydrol. Earth Syst. Sci., 5(2), 215–223, 2001b.

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(4), 737, doi:10.2166/wst.2013.300, 2013.

Editor (Dr. Ehret)

(Q1) The issue was raised about whether the raingauge observations are the best benchmark to compare the other methods against. Although the raingauge density in your study is quite high (~ 1 per 5 km²), this is true due to the points Dr. Mazzetti raised. Actually, you can reduce the problem of not knowing the true rainfall field by cross-validation: Split your gauge data set (even one station per 10 km² is still a good coverage), use one part to fuel the gauge-based estimation techniques, evaluate all methods on the remaining gauges or products derived thereof. This way you can rank all methods against a non-used data set. This does not remove the problem of a general bias of the raingauges (undercatch in case of wind etc.), but does not pre-favor gauge-based methods over the others. I realize that this requires substantial additional work, but this is important to make your study technically sound.

(R1) The issue of the methodology that is used to assess the performance of the merged QPEs is indeed a challenging one. In order to properly address the comment made by the Editor, it is important to make a clear distinction between two separate, yet highly related issues:

(1) **The use of rain gauge (RG) estimates as reference in the evaluation of other QPEs**. Doing so has inherent limitations arising from (1) the fact that RG records may contain errors (due to wind, splashing, miscalibration, etc.) and (2) the difference in the spatial and temporal resolutions of RG records and gridded QPEs.

(2) **The way in which QPEs under consideration (in our case, radar and merged products) are quantitatively compared against the reference rainfall estimates (in this case, RG)**. There are different ways of conducting such comparison, including comparison of areal average rainfall estimates and evaluation at point RG location. This can be done using all available RG in the generation of merged QPEs as well as in their evaluation, using an independent network of RGs (i.e. setting some gauges aside which are not used in the merging, but are used in the evaluation of QPEs - as suggested by the reviewer), or through cross validation (i.e. 'leave one out' strategy': the RG data at each site in turn is omitted from the calculations and the estimated value of rainfall at the ''hidden data" location is interpolated/estimated using the remaining data. The cross-validation error, defined as the difference between the estimated and known (but not used) value, can be then computed). Each of these strategies has advantages and disadvantages which will be discussed next. However, **what they all have in common is that they cannot overcome the limitations associated with the previous issue** (i.e. use of RG estimates as reference for evaluation). The fact that we leave some RGs out of the merging and use them for an independent assessment does not mean that the 'independent RG records' will be error free, nor does it resolve the issue of the differences in spatial and temporal resolution of the different rainfall estimates (i.e. point-area comparison, RG temporally-aggregated measurements vs. radar scans).

As recommended by Dr. Mazzetti, we have discussed in more detail the errors associated to rain gauge measurements and the limitations that this imposes on the testing that we conducted. However, as she pointed out and as explained above, these limitations are nearly impossible to overcome as the true rainfall field will remain unknown, regardless of the comparison strategy that is adopted.

With regard to the 'quantitative comparison strategies', in the original version of the manuscript we employed two strategies:

- ***Comparison of areal average rainfall estimates*:** Given the density of the RG network employed in our investigation, areal average RG rainfall estimates can be 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)), thus making this evaluation strategy robust (although the problem with errors in RG estimates, such as those caused by wind, will still be present in this comparison). While the evaluation of areal average rainfall estimates provides useful insights, it fails to provide sufficient information about small scale features, which are of outmost important in our investigation. Such localised information can be obtained from comparison at point RG locations.
- ***Comparison at point RG locations:*** the evaluation of rainfall estimates at point RG locations has the advantage of providing more information about small-scale rainfall features; however, it entails more uncertainty than areal average comparisons, given the well-known problem of point-areal differences as well as the higher sampling error (whereas in the case of areal average estimates the averaging of several point or grid records can statistically reduce the sampling error). In our case, we used all available RGs in the interpolation (block-kriging) and merging of rainfall estimates. Afterwards, we conducted a comparison between each pair of co-located RG and pixel of a given gridded rainfall estimate (so the same gauges used in the estimation of merged QPEs were used in the assessment). Given that all RGs were included in the generation of merged QPEs, this comparison at point RG locations will of course favour RG-resembling QPEs. **Alternatives to somehow overcome this problem include the use of an *independent network of rain gauges* (as suggested by Dr Ehret) or application of *cross-validation* (i.e. 'leave-one-out' strategy)**. The ***independent network*** strategy has the problem of resulting in a significantly reduced number of gauges to conduct the merging. As the Editor points out, leaving half of the RGs out of the merging would result in a density of ~1 RG / 10 km^2, which is very low for urban hydrological applications (RG densities between 1 RG / 4 km^2 to 1 RG / 1 km^2 are recommended for urban areas, depending on topography and other catchment characteristics; see Ochoa-Rodriguez et al. (2015), Schilling (1991) and WaPUG (2002)). **For the present paper, which focuses on demonstrating a new methodology, we prefer to use the best dataset that we can / that is available. Moreover, the objective of the present study is not so much to 'rank' the performance of different merging methods, but rather to analyse the relative performance of the proposed method (i.e. its special features, in relation to the original BAY method)**. An investigation into the impact of RG density on the SIN method and on other popular merging methods (e.g. BAY, MFB, KED, KRE) is envisaged as part of a new study for which we are sourcing larger rainfall datasets (both in terms of spatial and temporal extents). The ***cross-validation*** strategy, on the other hand, has the advantage of leaving out only one RG at the time, thus avoiding significant reduction in gauge density, while allowing an 'independent' assessment. The disadvantage of the cross-validation strategy is that the RG network varies at each iteration (Goudenhoofdt and Delobbe, 2009). **In spite of this limitation and motivated by the Editor's comments, we decided to apply the cross-validation strategy and compare the results against those we had previously obtained with our original point RG comparison strategy**. The results of the cross-validation and original point RG comparison (including all RGs) are shown in Figure 1 and Figure 2, respectively (for a description of the performance measures

please refer to the manuscript). As can be seen, **there are some differences in the results; however, the general trends, the relative performance of the different merging techniques and the conclusions that can be drawn are very similar for both evaluation strategies**. The main difference is that in the case of cross-validation, in which the gauge used for assessment has not been used in the interpolation and merging, RG-emulating methods perform less well than they do when all gauges are included in the estimation and evaluation. This is particularly the case for BK and BAY estimates. Not surprisingly, the performance of the MFB is barely affected by the evaluation strategy. Noteworthy is the fact that the performance of the SIN estimates shows very little variation between the cross-validation and the original evaluation strategy; this confirms that the SIN methodology can better preserve spatial radar features and is less sensitive to variations in the number of gauges and configuration of the RG network.

Based upon the results of the comparison between the cross-validation and original results, we can conclude that both evaluation strategies are valid, though none of them is perfect. However, we do appreciate that having a somehow independent assessment is desirable. Therefore, we have decided to replace the point-RG evaluation results included in the original manuscript by the cross-validation ones. We have changed the 'Evaluation methodology' accordingly (we have described the cross-validation method there) and we also made minor changes in the analysis of results, although not many changes were needed given that the general trends and conclusions remained similar.

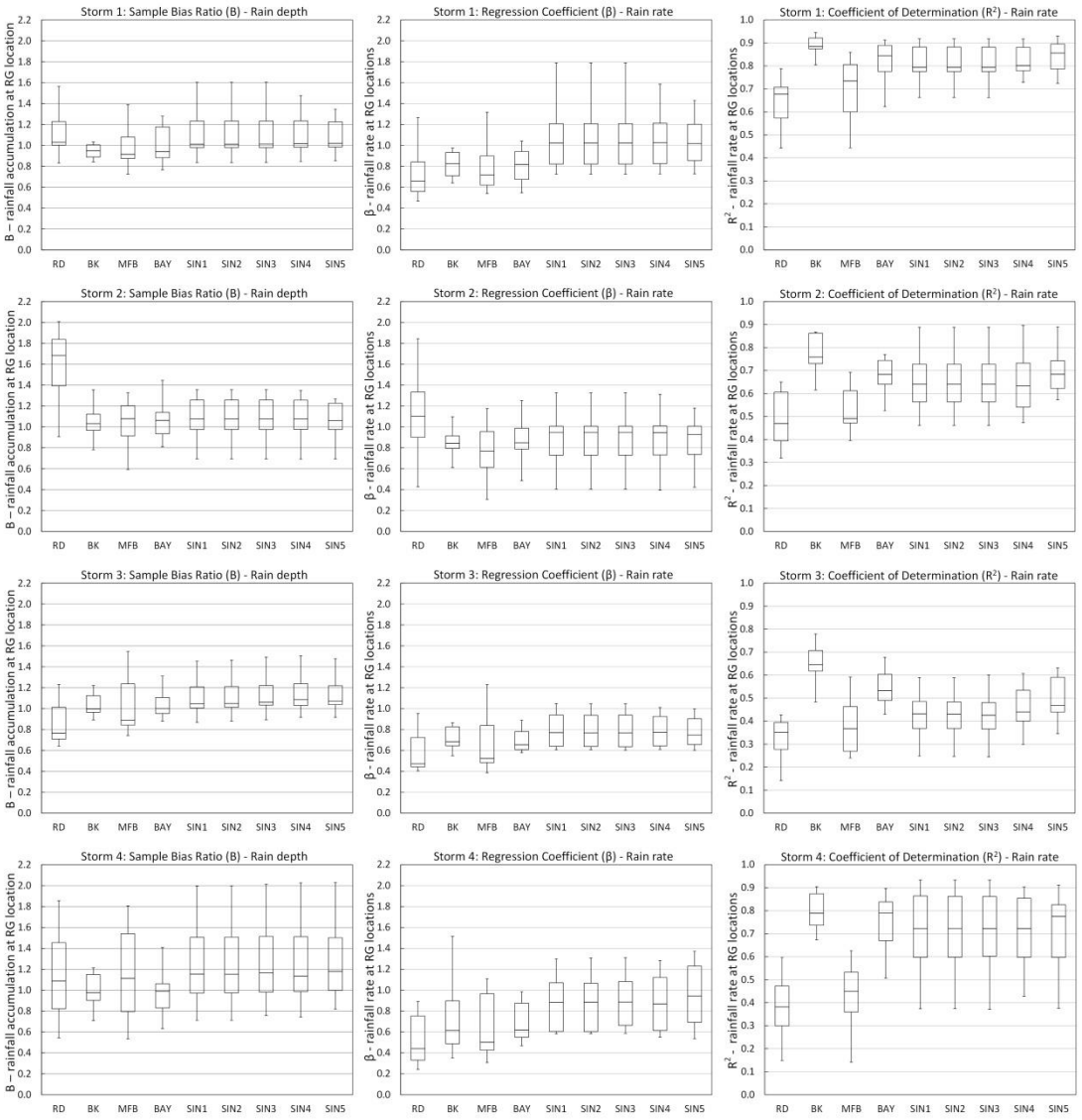

Figure 1: Boxplots displaying the result of **cross-validation**: distribution of sample bias ratio (B) (left column), regression coefficient (β) (middle column) and coefficient of determination (R^2) (right column) for regressions of the different gridded rainfall estimates at individual rain gauge locations against the RG rainfall.

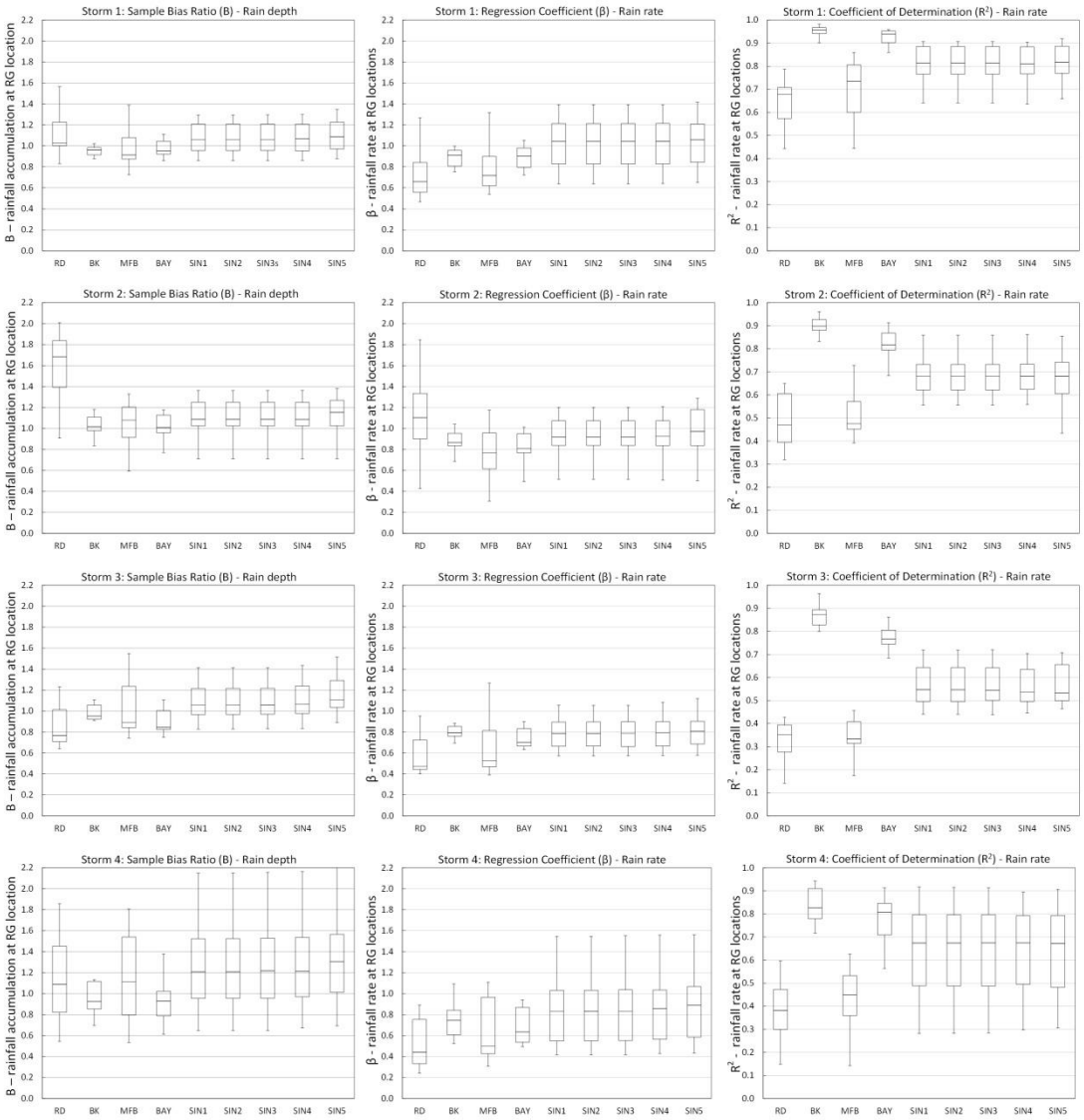

Figure 2: Boxplots displaying the result of the **original evaluation strategy** (i.e. including all rain gauges in the merging): distribution of **sample bias ratio (B)** (left column), **regression coefficient (β)** (middle column) and **coefficient of determination (R^2)** (right column) for regressions of the different gridded rainfall estimates at individual rain gauge locations against the RG rainfall.

(Q2) Generally, you explore and present an interesting new avenue for multi-sensor based estimation of a true rainfall field for small spatial scales and high spatio-temporal resolution by aiming to overcome major limitations of existing merging techniques, which generally rely on Gaussian data distributions and smoothing/interpolation approaches. Your singularity-based approach takes singularities into account, but is relatively 'complex'. What I would really like to see, in the sense of a parsimonious approach to solve a problem, is an evaluation of how a standard merging technique (here the original Bayesian merging) performs with data that have been brought closer to a Gaussian distribution beforehand, e.g. by log-transform. This way, the major problem of smoothing away singularities should be mitigated without much effort.

(R2) We appreciate that the Editor valued the idea of the proposed methodology and at the same time raised some interesting issues, with regard to the local singularity analysis and the normalisation of rainfall fields.

First, we would like to argue that the local singularity analysis used here is not a complex model, in terms of the concept and the techniques used (please refer to the detailed description of the local singularity analysis which has been added to the manuscript). As mentioned in the manuscript, the local singularity analysis was developed based upon the definition of the (coarse) Hölder exponent. This quantifies how the areal mean behaves with varying size of the area. The Hölder exponent has been widely used in multifractal analysis in hydrology to identify singular structures (although not necessarily under that name, but simply described as a scaling exponent); and similar concepts have been applied operationally to many fields, including signal processing, geo-chemicals and meteorology, to detect local abnormalities (Agterberg, 2012b; Cheng and Zhao, 2011; Cheng, 2012; Cheng et al., 1994; Lovejoy and Mandelbrot, 1985; Mallat and Hwang, 1992; Robertson et al., 2003; Schertzer and Lovejoy, 1987). In the field of meteorology, for example, a similar concept (i.e. comparing the rainfall rate at a specific pixel with those at the neighbouring pixels) has been employed by many national meteorological services to identify convective storm cells (Dixon and Seed, 2014; Goudenhoofdt and Delobbe, 2013; Steiner et al., 1995).

In terms of the techniques used, the local singularity analysis requires merely two parameters (alpha and C), which can be obtained via a simple linear regression. In fact, a comparison has been conducted in the literature between the local singularity analysis and a number of rather 'complex' multifractal-based techniques. The associated results have demonstrated that the local singularity analysis can provide almost the same information than those complex methods (Agterberg, 2012a; Cheng, 2012; Cheng et al., 1994).

Second, with regard to the creation of a field of higher normality, many data transformation methods (e.g. log- and Box-Cox transformation) indeed exist in the literature and have proven to be useful to improve the normality of rainfall data and consequently showed the potential to improve data merging/fusion (Berndt et al., 2014; Sinclair and Pegram, 2005). However, we would like to explain the reasons why we did not include them in our investigation:

1. 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 physically associated with abnormal energy concentration, such as 'convective' cells. So the logic behind is different.
2. Previous research suggests that excessive transformation (i.e. log-transformation) will cause positive bias in estimation, and the Box-Cox (trans-Gaussian) may be a better way to go for, but the transformation parameter (lambda) needs to be subjectively tuned or calibrated (Ciach et al., 2007; Erdin et al., 2012; Schuurmans et al., 2007).
3. Kriging is a critical technique in Todini's Bayesian data merging technique; however, it is no longer an optimal estimator while applied to the log transforms of the rainfall estimates. That is, the kriging estimator here will be the unbiased estimator which minimises the means square prediction error in the logarithms of the rainfall. But that does not mean that the variance of the error in the rainfall depths is thereby minimal (Nour et al., 2006).

4. The suggestion of an easier way of making the rainfall Gaussian partly misses the point (as explained in the first bullet point above). The issue is not only/primarily that the marginal distributions are not normal. If it were just a question of correcting the marginal distribution, something like the approach in (Allcroft and Glasbey, 2003) could be used. But that leaves the problem of the **spatial structure**. The much more important issue that many analyses suggest is that the spatial structure is not Gaussian. This is what emerges from the multifractal analyses of rainfall in space. And this cannot be corrected by making the marginal distributions normal. Indeed, one can have variables that are marginally normally distributed and such that the joint distribution is not normal. An easy example is that of two variables (can easily be extended to more) X and Y, where X has a normal distribution with expected value 0 and variance 1, and $Y = X$ if $|X| > c$ and $Y = -X$ if $|X| < c$, for some $c > 0$. Y will have the same normal distribution, but when you look at the joint density, it is not Gaussian. This is a simple textbook example, but is enough to demonstrate that the proposed transformation is not guaranteed to achieve Gaussian behaviour. And for grounds for believing that the spatial structure is not Gaussian, we would refer either to some multi-fractal work (the key point here is that for a non-Gaussian structure, moments beyond the second order moment are important, and each brings new information), or to the successful use of non-Gaussian copulas to represent spatial dependence, e.g. see Serinaldi (2009).

(Q3) In your approach, you consider singularities in the radar image. However, singularities will also be present in your raingauge observations, e.g. if one gauge is exactly below a local convective cell or in a rainfall 'hole'. The singularity is of course hard to detect due to the sparse spatial sampling, but can still be there. Would it make sense to apply the singularity-removal also to the gauge data beforehand?

(R3): What the editor mentions is indeed possible and is something we have somehow accounted for, though it was not explained in the manuscript.

While block-kriged (BK) fields are generally highly smooth (due to the nature of the interpolation), 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 partially be preserved in the Non-Singular Bayesian merged field (NS-BAY). 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 (i.e. once from the rain gauge data, and once from the radar data when singularity exponents are applied back to the Bayesian merged fields). This can ultimately result in a merged (SIN) product with more singularities than those originally observed in the radar image. In order to avoid this problem, 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.

This solution has proved generally useful and has been detailed in Section 2.3 of the revised manuscript. However, prompted by the Editor's comment and after reading related papers (e.g. Jewell and Gaussiat (2015) and Laurantin (2013)), during the revision period we explored other

alternatives for dealing with potential singularity structures in rain gauge data. We tested mainly two methods:

1. Extraction of singularities from the point rain gauge records using the singularity exponents derived from the co-located radar pixels:
   In this method, we assumed that the rain gauge and the co-located radar pixel data shared the same singularity information. So the singularity exponents obtained from the radar pixel values could be used to decompose the coincidental rain gauge rainfall data into the singular and non-singular components. This was done by dividing the rain gauge values by the ratios derived from Eq. 7 (page 11, line 25) in the revised manuscript. The non-singular rain gauge data were then used as input into the Bayesian data merging process. That is, the non-singular rain gauge data were interpolated and then merged with the non-singular radar rainfall estimates. However, there were many uncertainties in this method which largely affected the quality of the merged product. The very critical one was the accuracy of the mapping between rain gauge locations and the associated radar pixels, which is not always reliable and could be affected by, for example, the storm advection speed. The consequence is that the wrong singularity components would be extracted from the rain gauge data, and this would consequently decrease the quality of the merged product.
2. Extraction of singularities from the block-kriged rain gauge rainfall (BK) fields using local singularity analysis:
   This was done by applying the local singularity analysis to the BK fields before they were further merged with the non-singular radar rainfall fields. The effect of this method was somewhat similar to that of the moving-window treatment. However, numerically it was relatively unstable, so decided to stick to the original moving-window method.

As can be seen and as mentioned by the Editor, the issue of identifying, let alone removing, singularity structures from rain gauge data is very challenging due to the sparse and irregular distribution of rain gauge sites. While we are generally satisfied with the strategy that we have implemented at the moment, we believe that further work on this particular topic would be beneficial. This has been mentioned in the 'Conclusions and future work' session.

(Q4) The quality of the radar data play a major role in your merging approaches. It is therefore necessary to include a much more in-depth description of the radar data processing chain, operational corrections performed etc. For example, some of the low radar quality during high rainfall intensities could be attributed to path attenuation, if no attenuation correction is made.

(R4): The radar rainfall data used in this paper was the C-band radar composite (Nimrod) product provided by the UK Met Office (Golding, 1998). This product has been consistently quality-controlled by the Radarnet system of the UK Met Office. The main quality control and correction processes that have been applied include the identification and removal of anomalous propagation (e.g. beam blockage and clutter interference), attenuation correction and the vertical profile correction. Because these processes were not conducted by us and also a detailed description of them can be found in the literature (Harrison et al., 2000, 2009), we prefer not to include a detailed description of them in our manuscript. A summary of these correction processes has been added in the revised manuscript and a more clear reference has been made to Harrison et al. (2009, 2000); see new text:

"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 20 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))"

(Q5) Having only four storms and using three of them for the calibration of the hydrological/hydrodynamic model, and doing calibration such that it favors all raingauge-derived rainfall estimation methods is a problem, but I can see the reason for this limitation from the limited duration of the observation period. Still I think things can be learned from the hydrograph-based discussion. I suggest to keep it in, but stress its limitations, and also reduce its extent in the paper and impact on the final conclusions.

(R5) We agree with the Editor's comment in this regard. The limitations of the hydrological/hydraulic test and in particular the fact that three of the storms under consideration were used for model calibration were further emphasised in the Evaluation Methodology section (Section 3.2) and in the Hydraulic Modelling Results section (Section 3.3.2 Hydraulic modelling results). The text below was added to these sections in order to address this aspect. Regarding the extent of the hydrograph-based discussion, we think it is already concise and we found it difficult to shorten it without losing relevant information which may help the reader better understand the results.

NEW TEXT ADDED TO EMPHASISE LIMITATIONS OF HYDROLOGICAL/HYDRAULIC EVALUATION:

In methodology section, with regards to the limitations of the hydraulic evaluation strategy:

"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-resembling) rainfall estimates. Moreover, the relatively coarse spatial resolution of rainfall inputs (in this case ~ 1 RG / 4.4 km^2) may have led to further biases in the model (Kavetski et al., 2006). It would be desirable to recalibrate 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.

In results section:

"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. 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 to 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."

(Q6) Please also include some performance statistics of the calibrated models (Nash etc) in 3.1.1.

(R6) The performance statistics of the calibrated model, including NSE, RMSE, R^2 and β, have been included in Section 3.1.1. It is important to mention that the model used in this study is an operational model, used by the water utility of the area. Given its size and the number of its components, carrying out an automatic uncertainty-based calibration is a complicated task. As explained in the manuscript, the model was manually calibrated following UK standards (cited in the manuscript).

(Q7) Also, please state in which temporal resolution the rainfall data were fed into the model. 2-min? 5-min? Was it different for the raingauge- and radar-based rainfall estimates?

(R7) The local rain gauge records were originally at 2-min temporal resolution. Because the available radar data were at 5-min resolution, we (linearly) interpolated 2-min rain gauge data into 5-min before carrying out the data merging. In this way, all gridded rainfall estimates (i.e. BK, RD, MFB, BAY and SIN estimates) were available at a temporal resolution of 5-min. In addition, to ensure consistency and avoid differences in results due to disparity in the temporal resolution of rainfall inputs, we used the interpolated 5-min rain gauge data (rather than the original 2-min records) as input into the hydraulic model. This has been explained in the 'Case study' section of the manuscript.

Some technical aspects:

(Q8) Like Dr. Sinclair, I found it hard to fully understand the singularity approach (extracting and imposing the singularities) from the text. So I encourage you to go into more detail here as indicated in your response to Dr. Sinclair.
(R8) This has been improved in the revised manuscript. Please see marked manuscript and refer to our answer to Q1 by Dr. Sinclair.

(Q9) Figure 7 is indeed too full. Consider distributing the timeseries over several plots (e.g. all SIN plots of an event into a separate plot), and/or to zoom into the most interesting parts of the hydrographs (peaks).
(R9) This recommendation was followed. The figure was indeed split into two figures, one including all SIN plots and another one including all other rainfall estimates and only one SIN estimate (i.e. SIN3). We think it is now much clearer. Besides, in the final format of the paper, this figure should appear bigger than it does in the HESS-Discussion format.

(Q10) Throughout the paper I found it hard to understand in which temporal aggregations the data were used and why: How were the 2-min raingauge and 5-min radar data attuned for combination or comparison of the instantaneous rain rates?
(R10) Please refer to Q7 above.

(Q11) Please include the location of the radar in your overview map in Fig. 2.

(R11) This was done. Please note that this figure is now Fig. 3, as a new figure (Fig. 2: Schematic of Local Singularity Analysis) was introduced.

References:

Agterberg, F. P.: Multifractals and geostatistics, J. Geochemical 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, Nonlinear Process. Geophys., 19(1), 23–44, doi:10.5194/npg-19-23-2012, 2012b.

Allcroft, D. J. and Glasbey, C. A.: A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation, J. R. Stat. Soc. Ser. C (Applied Stat., 52(4), 487–498, doi:10.1111/1467-9876.00419, 2003.

Berndt, C., Rabiei, E. and Haberlandt, U.: 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, 2014.

Cheng, Q.: Singularity theory and methods for mapping geochemical anomalies caused by buried sources and for predicting undiscovered mineral deposits in covered areas, J. Geochemical Explor., 122, 55–70, doi:10.1016/j.gexplo.2012.07.007, 2012.

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.

Ciach, G. J., Krajewski, W. F. and Villarini, G.: Product-Error-Driven Uncertainty Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD Data, J. Hydrometeorol., 8(6), 1325–1347, doi:10.1175/2007JHM814.1, 2007.

Dixon, M. and Seed, A. W.: Developments in echo tracking - enhancing TITAN, in The Eighth European Concference on Radar in Meteorology and Hydrology., 2014.

Erdin, R., Frei, C. and Künsch, H. R.: Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges, J. Hydrometeorol., 13(4), 1332–1346, doi:10.1175/JHM-D-11-096.1, 2012.

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.

Goudenhoofdt, E. and Delobbe, L.: Statistical Characteristics of Convective Storms in Belgium Derived from Volumetric Weather Radar Observations, J. Appl. Meteorol. Climatol., 52(4), 918–934, doi:10.1175/JAMC-D-12-079.1, 2013.

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.

Harrison, D. L., Scovell, R. W. and Kitchen, M.: High-resolution precipitation estimates for hydrological uses, Proc. Inst. Civ. Eng. Water Manag., 162(2), 125–135, 2009.

Jewell, S. A., & Gaussiat, N. (2015). An assessment of kriging-based rain-gauge-radar merging techniques. Quarterly Journal of the Royal Meteorological Society, doi: 10.1002/qj.2522.

Laurantin, O. (2013). ANTILOPE: hourly rainfall analysis over France merging radar and rain gauge data. In 11th International Precipitation Conference, Ede-Wageningen, The Netherlands.

Lovejoy, S. and Mandelbrot, B. B.: Fractal Properties of Rain, and a Fractal Model, Tellus, 37A, 209–232, 1985.

Mallat, S. and Hwang, W.-L.: Singularity detection and processing with wavelets, Inf. Theory, IEEE Trans., 38(2), 617–643, doi:10.1109/18.119727, 1992.

Nour, M. H., Smit, D. W. and Gamal El-Din, M.: Geostatistical mapping of precipitation: implications for rain gauge network design, Water Sci. Technol., 53(10), 101–110, 2006.

Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., Assel, J. van, Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P. and ten Veldhuis, M.-C.: Impact of Spatial and Temporal Resolution of Rainfall Inputs on Urban Hydrodynamic Modelling Outputs: A Multi-Catchment Investigation, J. Hydrol., doi:10.1016/j.jhydrol.2015.05.035, 2015.

Robertson, A. N., Farrar, C. R. and Sohn, H.: Singularity Detection for Structral Health Monitoring Using Holder Exponents, Mech. Syst. Signal Process., 17(6), 1163–1184, doi:http://dx.doi.org/10.1006/mssp.2002.1569, 2003.

Schertzer, D. and Lovejoy, S.: Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Processes, J. Geophys. Res., 92, 1987.

Schilling, W.: Rainfall data for urban hydrology: what do we need?, Atmos. Res., 27(1–3), 5–21, doi:10.1016/0169-8095(91)90003-F, 1991.

Schuurmans, J. M., Bierkens, M. F. P., Pebesma, E. J. and Uijlenhoet, R.: Automatic Prediction of High-Resolution Daily Rainfall Fields for Multiple Extents: The Potential of Operational Radar, J. Hydrometeorol., 8(6), 1204–1224, doi:10.1175/2007JHM792.1, 2007.

Serinaldi, F.: A multisite daily rainfall generator driven by bivariate copula-based mixed distributions, J. Geophys. Res., 114(D10), D10103, doi:10.1029/2008JD011258, 2009.

Sinclair, S. and Pegram, G.: Combining radar and rain gauge rainfall estimates using conditional merging, Atmos. Sci. Lett., 6(1), 19–22, doi:10.1002/asl.85, 2005.

Steiner, M., Houze, R. A. and Yuter, S. E.: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data, J. Appl. Meteorol., 34(9), 1978–2007, doi:10.1175/1520-0450(1995)034<1978:CCOTDS>2.0.CO;2, 1995.

WaPUG: Code of practice for the hydraulic modelling of sewer systems, Wastewater Planning Users Group WaPUG, UK., 2002.

Manuscript prepared for Hydrol. Earth Syst. Sci. Discuss.
with version 4.1 of the LaTeX class copernicus_discussions.cls.
Date: 3 August 2015

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, B-3001 Heverlee (Leuven), Belgium
[2]Department of Civil and Environmental Engineering, Imperial College 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, $Z = aR^b$ (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 can-

not effectively reflect the short-term dynamics of true rainfall intensities and may statistically compromise to 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 (1st and/or 2nd statistical moments) approximations (Goudenhoofdt and Delobbe, 2009; Krajewski, 1987; Todini, 2001), cannot properly cope with the so-called "singularities" (which implies non-normality and often corresponds 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. Cokriging (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 2nd-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 network 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 numer-

ical strategies. In Sect. 3 we present the case study, including a description of the study area and dataset, 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 supplementary document "Formulation and demonstration of the singularity-sensitive Bayesian data merging method".

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 ($V_{\varepsilon_t}^{\mathrm{RG}}$), 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: $\varepsilon_t^{\mathrm{RD}} = y_t^{\mathrm{RD}} -$

y_t^{RG}) is obtained empirically. Assuming that areal rain gauge estimates are unbiased, the expectation value ($\mu_{\varepsilon_t^{\mathrm{RD}}}$) and the co-variance function ($V_{\varepsilon_t}^{\mathrm{RD}}$) 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 (2nd-order) co-variance function that these

techniques employ to characterise radar-rain gauge errors cannot well capture local singularity structures. Instead, in 2nd-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 a singularity and it is often associated to structures depicting fractality or multifractality (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 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 ϵ ($\epsilon = l/L$, 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 $\rho_1, \rho_2, \ldots, \rho_n$, respectively at scales $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$). 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 dataset 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, Eg. 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 ϵ 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 ϵ changes, according to the local singularity index $\alpha(x)$. When $\alpha(x) < 2$, the rainfall magnitude strikingly increases as the measuring scale ϵ decreases (namely local enrichment); this corresponds to a "peak" singularity. In contrast, when $\alpha(x) > 2$, the rainfall magnitude decreases as ϵ 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 con-

dition (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

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

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 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 south-western region. The storm water drainage system is predominantly combined with some separate sewers and drains from the south-west to the north-east (towards the sea).

3.1 Available models and datasets

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 setup 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 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 $0.63\,\mathrm{m\,m^{-1}}$, with a mean slope of $0.031\,\mathrm{m\,m^{-1}}$. 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 $0.25\,\mathrm{m\,m^{-1}}$, 55 % have a gradient between 0.01 and $0.1\,\mathrm{m\,m^{-1}}$, and 43 % have a gradient $< 0.01\,\mathrm{m\,m^{-1}}$.

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; coefficient of determination (R^2) of 0.7756; and regression coefficient (β), 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 $> 5\,\mathrm{mm\,h^{-1}}$ 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 duration. 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 $B_t = \sum \mathrm{RG}_{\text{catchment}} / \sum \mathrm{RD}_{\text{co-located}}$, where $\mathrm{RG}_{\text{catchment}}$ and $\mathrm{RD}_{\text{co-located}}$ 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, $\mathrm{MFB} = B_t \cdot \mathrm{RD}$.

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. 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 (Habib et al., 2008). The order of magnitude of the two main error sources associated to 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 mm/h and 20 % for intensities of 200 mm/h, 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 ~ 1 RG$/4.4$ 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.

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 β (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 "modelled" (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, β, 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 to the gridded estimate are higher in the mean (by a factor of β) 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 β statistics obtained from the linear regression analysis. In addition, the "weighted" coefficient of determination (R^2_{w}) is employed to quantify the joint performance of hydrological efficiency. This measure is defined as (Krause et al., 2005):

$$R^2_{\mathrm{w}} = \begin{cases} |\beta| \cdot R^2 & \text{for } \beta \leq 1 \\ |\beta|^{-1} \cdot R^2 & \text{for } \beta > 1 \end{cases} \quad (8)$$

where higher R^2_w 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 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 evident 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 α values different from 2 (Fig. 4), were observed. This tendency to underestimate can be attributed to a combined effect of the BAY method "overtrusting" 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 (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 non-linear fashion) with the singularity structures preserved throughout the BAY merging, thus leading to an overestimation 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.

Fig. 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 1st-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 2nd- 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 (β) 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 β 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 β 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 regards 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 β 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 β 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 β 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 2nd-order approximation (which smooths off some local extreme magnitudes).

With regards 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 β, it can be seen that although the median values are usually below one, they are generally much closer to one than other rainfall estimates. Moreover, their distribution is more variable than that associated to BK and BAY estimates (which display a systematic underestimation). In line with the results 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 β 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 4 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 mentioned 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 regards 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 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

Fig. 8 and Fig. 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. β, R^2 and R^2_w) conducted at each flow gauging station for each storm event.

From the hydrographs in Figures 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 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 ob-

served flow peaks than RG associated outputs (which were used as input for the calibration of the model). With regards 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 β 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 β values associated with all rainfall inputs are mostly below 1. With regards 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 β values are significantly lower than 1 and lower than the β 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 β values well below one and much lower than those obtained for RG outputs). Moreover, they display relatively low R^2 and associated R^2_{w} 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 β 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 β values, whereas the SIN estimates managed to keep β 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 β values consistently lower than those of RG outputs). This confirms the smoothing of rainfall peaks that occurs when 2nd-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 β values closer to 1, higher R^2 values (sometimes even higher than those of RG outputs) and consequently higher R^2_{w} values. The better performance of SIN hydraulic outputs over BAY ones, particularly in terms of β, 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. Same as was observed in the hydrographs in Fig. 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 to 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 ($53\,\mathrm{km}^2$) (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, rain-

fall 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 aspect on which further work is recommended are the following:

- 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.

- 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.

Supplementary material related to this article is available online at: http://www.hydrol-earth-syst-sci-discuss.net/0/1/2015/hessd-0-1-2015-supplement.pdf.

Acknowledgements. The authors would like to acknowledge the support of the Interreg IVB NWE RainGain project, the Research Foundation–Flanders (FWO) and the PLURISK project for the Belgian Science Policy Office of which this research is part. Special thanks go to Alex Grist and Richard Allitt, from Richard Allitt Associates, for providing the rain gauge data and the hydraulic model, and for their constant support with the hydraulic simulations. Thanks are also due to the UK Met Office and the BADC (British Atmospheric Data Centre) for providing Nimrod (radar) data, to Innovyze for providing the InfoWorks CS software, and to Dr. Cinzia Mazzetti and Prof. Ezio Todini for making freely available to us the RAINMUSIC software package for meteorological data processing. Lastly, the authors would like to thank the reviewers, Dr. Scott Sinclair and Dr. Cinzia Mazzetti, and the Editor, Dr. Uwe Ehret, for their insightful and constructive comments which helped improved the manuscript significantly.

References

Agterberg, F. P.: Mixtures of multiplicative cascade models in geochemistry, Nonlinear Proc. Geoph., 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 Meteor. 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. Technol., 20, 752–759, 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., 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.

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.

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 Meteor. 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 multifractal 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), 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, doi:10.2166/wpt.2007.042, 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.

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, 767–776., 2001.

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.

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 Manag., 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., 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., 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, Nottingham, UK, 13–15 November 2013, 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 re-construction based upon Kriging and local singularity analysis, in: Hydrofractals '13, Kos Island, Greece, 17–19 October 2013, HF-10, 2013.

Wang, L.-P., Onof, C., Ochoa-Rodríguez, S., and Simões, N.: Analysis of kriged rainfields using multi-fractals, in: 9th International Workshop on Precipitation in Urban Areas: Urban Challenges in Rainfall Analysis, St. Moritz, Switzerland, 6–9 December 2012, 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), Washington, DC, 7–9 April 2014, 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., Qin, Z.: Effective image fusion rules Of multi-scale image decomposition. in: 5th International Symposium on Image and Signal Processing and Analysis,Istanbul, Turkey, 27–29 September 2007, 362–366, doi:10.1109/ISPA.2007.4383720, 2007.

Table 1. Selected rainfall events over the Portobello catchment.

Event	Date	Duration (h)	RG Total (mm)	RG Peak Intensity ($\mathrm{mm\,h^{-1}}$)	RD Total (mm)	RD Peak Intensity ($\mathrm{mm\,h^{-1}}$)
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).

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 accumulation (mm)	RG	9.25	7.70	32.96	5.03
	RD	9.80	11.02	26.21	4.69
	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 intensity over 5 min ($\mathrm{mm\,h^{-1}}$)	RG	11.21	5.03	8.46	7.11
	RD	7.93	4.10	3.33	2.41
	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

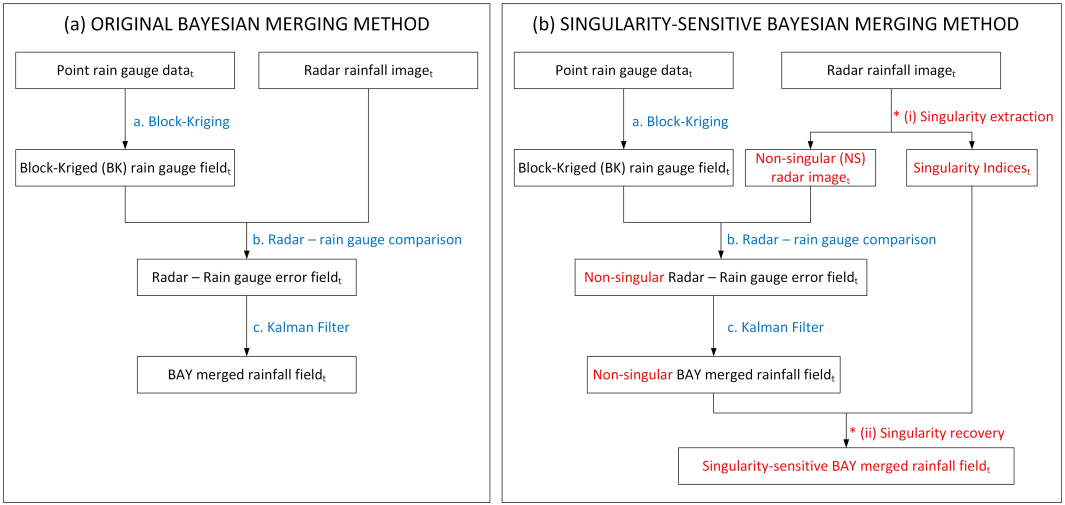

Fig. 1. Schematics of **(a)** Original Bayesian merging method (adapted from Fig. 2 in Mazzetti (2012)) and **(b)** Singularity-Sensitive Bayesian merging method.

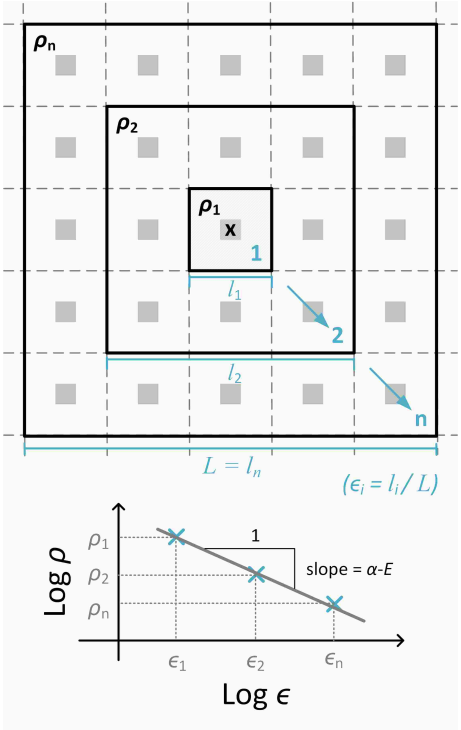

Fig. 2. Schematic of the local singularity analysis (adapted from Wang et al. (2014)).

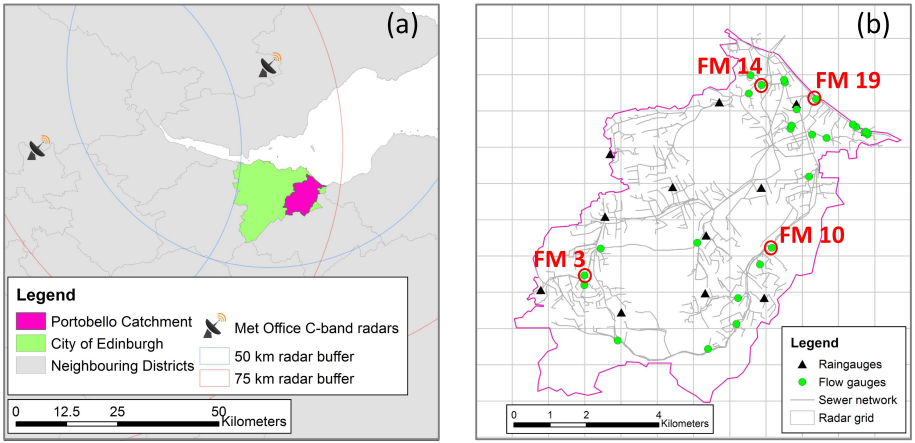

Fig. 3. Portobello catchment **(a)** general location; **(b)** sensor location, sewer network and radar grid over the catchment. Flow gauges FM 3, 10, 14 and 19 (the circled round markers in **(b)**, respectively located at the up-, mid- and downstream parts of the catchment) are particularly selected for visual inspection in Sect. 3.3.2.

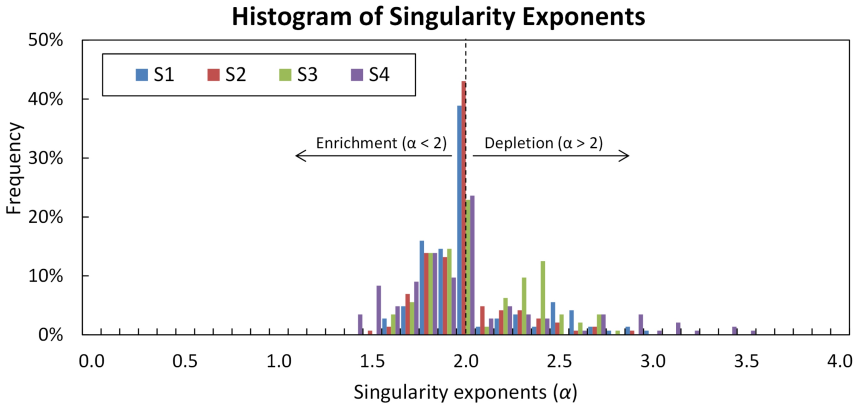

Fig. 4. Histogram of singularity exponents (α) at the time of areal peak intensity for the four selected storm events.

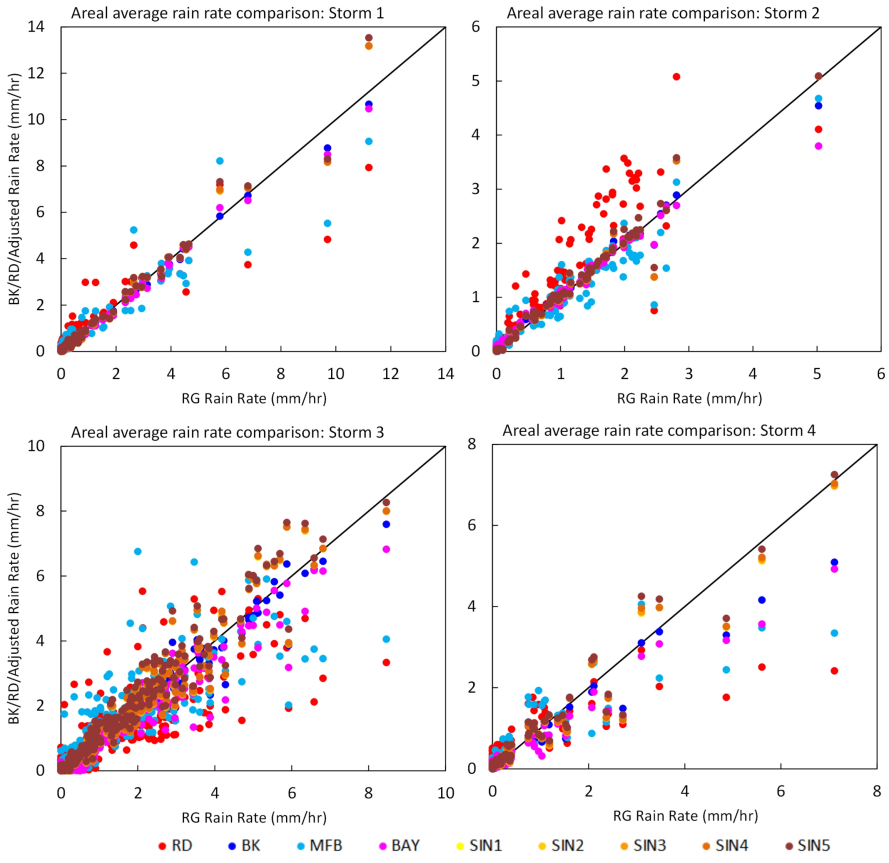

Fig. 5. Scatterplots of instantaneous areal average RG vs. RD/BK/MFB/BAY/SIN rainfall rates over the Portobello catchment for the four selected events, where SIN1–SIN5 represents the SIN estimates with different "truncated" singularity ranges (from widest to narrowest).

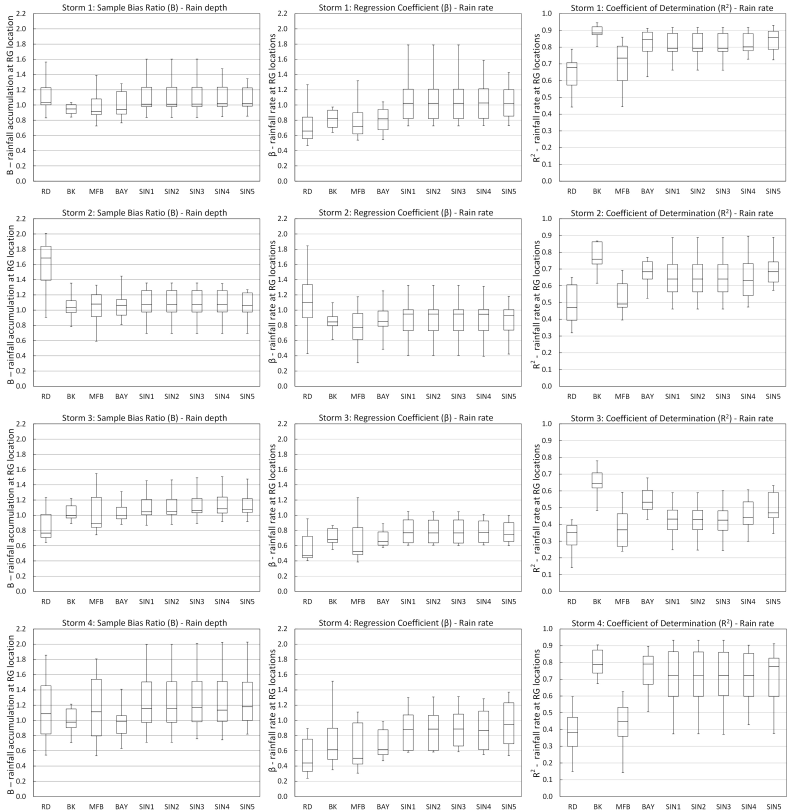

Fig. 6. Boxplots displaying the distribution of sample bias ratio (B) (left column), regression coefficient (β) (middle column) and coefficient of determination (R^2) (right column) estimated between the different gridded rainfall estimates and RG records at individual rain gauge locations following a cross-validation approach.

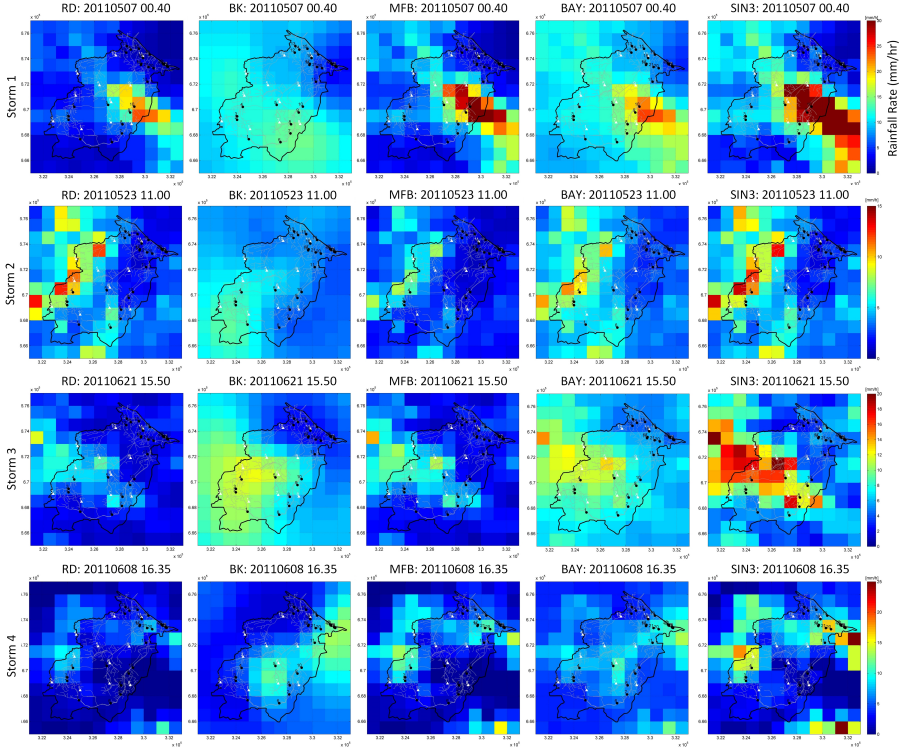

Fig. 7. Snapshot images of the different spatial rainfall products at the time of peak areal intensity for Storms 1 (top) to 4 (bottom) over the Portobello catchment. From left to right: RD, BK, MFB, BAY and SIN3 (with singularity range $[1, 3]$) estimates. The black polygon indicates the boundary of the Portobello catchment, and the black and white markers respectively represent the location of flow and rain gauges.

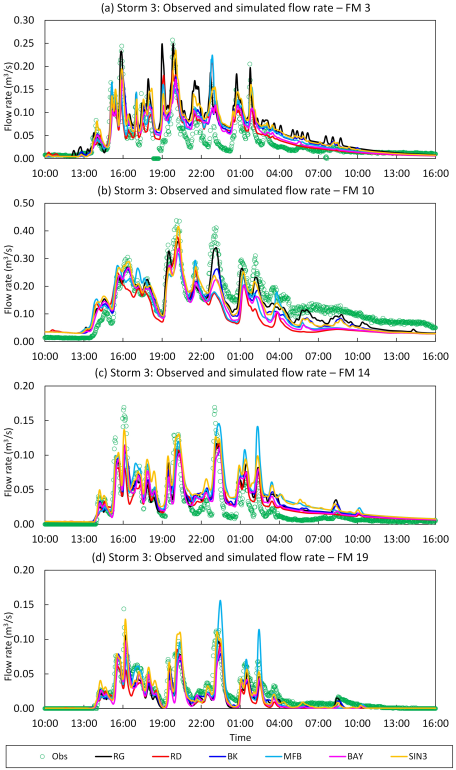

Fig. 8. Observed flows vs. simulated flows with RG, RD, BK, BAY and SIN3 rainfall inputs at selected flow gauging sites of Portobello catchment during Storm 3. Selected gauging sites: **(top)** FM3: upstream end of the catchment; **(bottom)** FM10: mid-stream area. The location of the selected monitoring sites is shown in Fig. 3.

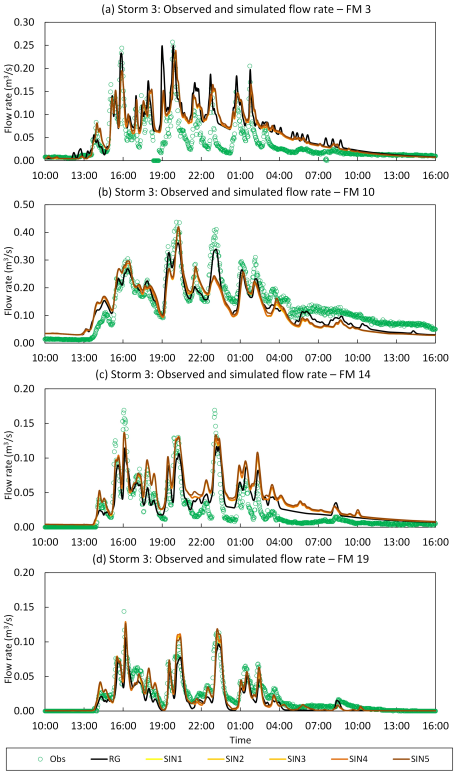

Fig. 9. Observed flows vs. simulated flows with SIN1-SIN5 rainfall inputs at selected flow gauging sites of Portobello catchment during Storm 3. Selected gauging sites: **(top)** FM14: upstream end of a small branch of the sewer system; **(bottom)** FM19: downstream end of the catchment. The location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

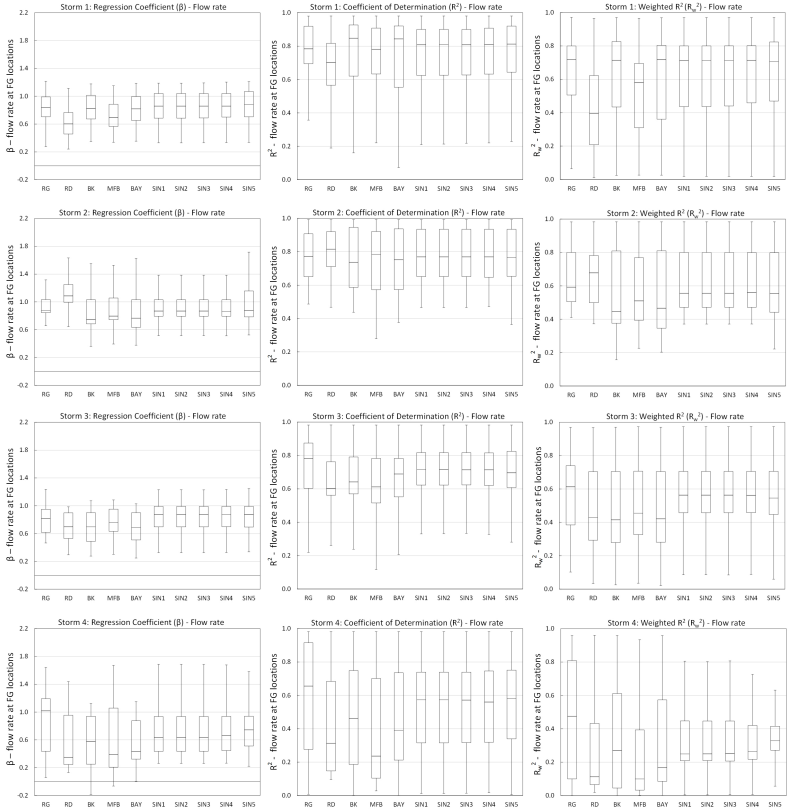

Fig. 10. Boxplots displaying the distribution of regression coefficient (β) (left column), coefficient of determination (R^2) (middle column) and weighted coefficient of determination (R^2_{w}) (right column) statistics derived from the linear regression analysis conducted for each pair of recorded and simulated flow time series at each gauging location.