

Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data

Auguste Gires, Ioulia Tchiguirinskaia, D Schertzer, A. Schellart, A. Berne, S. Lovejoy

▶ **To cite this version:**

Auguste Gires, Ioulia Tchiguirinskaia, D Schertzer, A. Schellart, A. Berne, et al.. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. Atmospheric Research, 2014, 138, pp.125-138. 10.1016/j.atmosres.2013.11.008 . hal-00913286

HAL Id: hal-00913286
https://enpc.hal.science/hal-00913286

Submitted on 12 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title: Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data

Authors

Auguste Gires[(1)], Ioulia Tchiguirinskaia[(1)], Daniel Schertzer[(1)], Alma Schellart[(2)], Alexis Berne[(3)], Shaun Lovejoy[(4)]

(1) U. Paris-Est, Ecole des Ponts ParisTech, LEESU, Marne-la-Vallée, France

(2) U. of Sheffield, Dept. of Civil and Structural Engineering, United Kingdom

(3) Laboratoire de Télédétection Environnementale, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

(4) McGill U., Physics dept., Montreal, PQ, Canada

Corresponding author : Auguste Gires (auguste.gires@leesu.enpc.fr, +33 1 64 15 36 48)

Abstract:

Rain gauges and weather radars do not measure rainfall at the same scale; roughly 20 cm for the former and 1 km for the latter. This significant scale gap is not taken into account by standard comparison tools (e.g. cumulative depth curves, normalized bias, RMSE) despite the fact that rainfall is recognized to exhibit extreme variability at all scales. In this paper we suggest to revisit the debate of the representativeness of point measurement by explicitly modelling small scale rainfall variability with the help of Universal Multifractals. First the downscaling process is validated with the help of a dense networks of 16 disdrometers (in Lausanne, Switzerland), and one of 16 rain gauges (Bradford, United Kingdom) both located within a 1 km^2 area. Second this downscaling process is used to evaluate the impact of small scale (i.e.: sub - radar pixel) rainfall variability on the standard indicators. This is done with rainfall data from the Seine-Saint-Denis County (France). Although not explaining all the observed differences, it appears that this impact is significant which suggests changing some usual practice.

Keywords: radar - rain gauge comparison, Universal Multifractals, downscaling

1) Introduction

The most commonly used rainfall measurement devices are tipping bucket rain gauges, disdrometers, weather radars and (passive or active) sensors onboard satellites. In this paper we focus on the observation scale gap between the two first devices which are considered here as point measurements and weather radars. A rain gauge typically collects rainfall at ground level over a circular area with a diameter of 20 cm and the sample area of operational disdrometers is roughly 50 cm^2 whereas a radar scans the atmosphere over a volume whose projected area is roughly 1 km^2 (for standard C-band radar operated by most of the western Europe meteorological national services). Hence observation scales differ with a ratio of approximately 10^7 between the two devices. A basic consequence, (e.g. Wilson, 1979), is that direct comparison of the outputs of the two sensors is at least problematic.

Standard comparison between rain gauge and radar rainfall measurements are based on scatter plots, rain rate curves, cumulative rainfall depth curves, and the computation of various scores such as normalized bias, correlation coefficient, root mean square errors, Nash-Sutcliffe coefficient etc. (see e.g., Diss et al., 2009; Emmanuel et al., 2012; Figueras I Ventura et al., 2012; Krajewski et al., 2010; Moreau et al. 2009). Despite usually being mentioned the issue of the representativeness of point measurement (i.e. disdrometer or rain gauge) with regards to average measurements (i.e. radar) is basically not taken into account and its influence on the standard scores is not assessed. Furthermore the authors who addressed it either to separate instrumental errors from representativeness errors (Ciach et al., 1999, Zhang et al., 2007; Moreau et al. 2009), or to introduce an additional score taking into account an estimation of the representativeness error (Emmanuel et al., 2012; Jaffrain and Berne, 2012) all rely on a geostatistical framework which may tend to underestimate rainfall variability and especially the extremes. Indeed this framework assumes that the rainfall field or a transform of it is Gaussian, which does not enable to fully take into account the fact that the extremes of rainfalls exhibit a power law behaviour as it has been shown by various authors (Schertzer et al., 2010; Hubert, 2001; Ladoy et al., 1993; de Lima and Grassman, 1999; Schertzer and Lovejoy, 1992).

In this paper we suggest to revisit how the representativeness issue is taken into account in standard comparison tools between point measurement devices (disdrometers or rain gauges) and radar rainfall measurements by explicitly modelling the small scale rainfall variability with the help of Universal Multifractals (Schertzer and Lovejoy, 1987). They rely on the physically based notion of scale-invariance and on the idea that rainfall is generated through a multiplicative cascade process. They have been extensively used to analyse and simulate geophysical fields extremely variable over wide range of scales (see Schertzer and Lovejoy 2011 for a recent review). The issue of instrumental errors is not addressed in this paper.

The standard comparison tools are first presented and implemented on 4 rainfall events over the Seine-Saint-Denis County for which radar and rain gauges measurements are available (section 2). A downscaling process is then suggested and validated with two dense networks of point measurement devices (disdrometers or rain gauges) (section 3). Finally the influence of small scale rainfall variability on the standard scores is assessed and discussed (section 5).

2) Standard comparison

2.1) Rainfall data in Seine-Saint-Denis (France)

The first data set used in this paper consists in the rainfall measured by 26 tipping bucket rain gauges distributed over the 236 km^2 Seine-Saint-Denis County (North-East of Paris). The rain gauges are operated by the Direction Eau et Assainissement (the local authority in charge of urban drainage). The temporal resolution is 5 minutes. For each rain gauge the data is compared with the corresponding radar pixel of the French radar mosaic of Météo-France whose resolution is 1 km in space and 5 min in time (see Tabary, 2007, for more details about the radar processing). The closest radar is the C-band radar of Trappes, which is located South-West of Seine-Saint-Denis County. The distance between the radar and the rain gauges ranges from 28 Km to 45 Km. See Fig. 1 mapping the location of the rain

gauges and the radar pixels. Four rainfall events whose main features are presented Table 1 are analysed in this study. They were selected (especially the last three) being among the heaviest observed events.

2.2) Scores

The radar and rain gauge measurements of the four studied rainfall events over the Seine-Saint-Denis County are compared with the help of scores commonly used for such tasks (Diss et al., 2009; Emmanuel et al., 2012; Figueras I Ventura et al., 2012; Krajewski et al., 2010; Moreau et al. 2009):

- The Normalized Bias (*NB*) whose optimal value is 0: $NB = \frac{\langle R \rangle}{\langle G \rangle} - 1$ (Eq. 1)

- The correlation coefficient (*corr*) which varies between -1 and 1 and whose optimal value is 1: $corr = \frac{\sum_{\forall i} \left(G_i - \langle G \rangle\right)\left(R_i - \langle R \rangle\right)}{\sqrt{\sum_{\forall i} \left(G_i - \langle G \rangle\right)^2} \sqrt{\sum_{\forall i} \left(R_i - \langle R \rangle\right)^2}}$ (Eq. 2)

- The Nash-Sutcliffe model efficiency coefficient (*Nash*), which varies between $-\infty$ and 1 and whose optimal value is 1: $Nash = 1 - \frac{\sum_{\forall i} \left(R_i - G_i\right)^2}{\sum_{\forall i} \left(G_i - \langle G \rangle\right)^2}$ (Eq. 3)

- The Root mean square error (*RMSE*), which varies between 0 and $+\infty$ and whose optimal value is 0 : $RMSE = \sqrt{\frac{\sum_{\forall i} \left(R_i - G_i\right)^2}{N}}$ (Eq. 4)

- The *Slope* and *Offset* of the orthogonal linear regression. It minimizes the orthogonal distance from the data points to the fitted line, contrary to the ordinary linear regression which minimizes the vertical distance and hence considers one of the data types as reference which is not the case for the orthogonal regression. The optimal values are respectively 1 and 0.

- The percentage ($\%_{1.5}$) of radar time steps (R_i) contained in the interval $\left[G_i / 1.5; 1.5 G_i\right]$ (it should be mentioned that this score is less commonly used than the others, 1.2 and 2 instead of 1.5 were also tested and yield similar results which are not presented here).

Where *R* and *G* correspond respectively to radar and rain gauge data. <> denotes the average. Time steps (index *i* in the previous formulas) of either a single event or all of them are used in the sum for each indicator. The time step usually considered by meteorologist is one hour. As in Emmanuel et al. (2012) and Diss et al. (2012), in addition we will consider time steps of 5 and 15 min which are particularly important for various practical applications in urban hydrology (Berne et al., 2004; Gires et al., 2012b). In this paper in order to limit the influence of time steps with low rain rate and especially the zeros rainfall time steps on the indicators, we only take into account the time steps for which the average rain rate measured by either the radar or the rain gauges is greater than 1 mm/h (Figueras i Ventura et al., 2012). The results are presented for an identical threshold for the various time steps. However, conclusions on the relation between the scores for different thresholds remain similar but with

slightly different score values. The scatter plot for all the events is visible in Fig. 2. The values of the scores are displayed in Fig. 3 for the 4 events, while the summary for all the events is given in Tab. 2. Whatever the case, all the rain gauges are considered at once, implying that the influence of the distance between the rain gauges and the radar is not analysed here. This choice was made because this distance ranges from 28 Km 45 Km according to the rain gauge which is not significant enough to enable a proper analysis of this issue (see Emmanuel et al. 2012 for a more precise study of this effect).

Overall it appears that there are great disparities between the events with much better scores for the 15 Aug. 2010 and 15 Dec. 2011 events than for the other two events. As it was observed in previous studies (Emmanuel et al., 2012; Diss et al., 2009) the scores tend to improve with increasing time steps. We find that for a given score, the ranking between the events remains the same for all the time steps which highlights the interests of performing analysis through scales rather than multiplying analysis at a given scale which is commonly done. More precisely it appears that the ranking between the events varies according to the selected score. For instance the 9 Feb. 2009 event is the worst event for *Nash*, *corr* and *NB*, whereas this is the case for $\%_{1.5}$, *slope*, *offset* and *RMSE* on 14 Jul. 2010. Furthermore for some scores, the estimated value is significantly different for an event with regards to the other ones. This is the case for *RMSE* for the 14 Jul. 2012 event (which strongly affects the value of this score when all the events are considered) or for *Nash* for the 9 Feb. 2009 event (here the score for all the events is not too affected). These differences make it harder to interpret precisely the consistency between rain gauges and radar measurements according to the event.

3) Bridging the scale gap

3.1) Methodology

In this paper we suggest to bridge the scale gap between radar and point (disdrometer or rain gauge) measurements with the help of a downscaling process based on the framework of Universal Multifractals (UM) (Schertzer and Lovejoy, 1987; Schertzer and Lovejoy 2011 for a recent review). The UM framework is indeed convenient to achieve this, because its basic assumption is that rainfall is generated through a space-time cascade process meaning that the downscaling simply consists in extending stochastically the underlying multiplicative cascade process over smaller scales (Biaou, 2005). The underlying multiplicative cascade process fully characterizes the spatio-temporal structure, especially the long range correlation and the variability through scales of the field. In the UM framework the conservative process (e.g., rainfall) is characterized with the help of only two parameters; C_1 being the mean intermittency (which measures the clustering of the average intensity at smaller and smaller scales with $C_1=0$ for an homogeneous field) and α being the multifractality index (which measures the clustering variability with regards to intensity level, $0 \le \alpha \le 2$). The UM parameters used here are $\alpha=1.8$ and $C_1=0.1$ which are in the range of those found by various authors who focused their analysis on the rainy portion of the rainfall field (de Montera et al, 2009; Mandapaka et al., 2009; Verrier et al., 2010, Gires et al., 2013). In this framework the statistical properties (such as the moment of order q) of the rainfall intensity field (R_λ) at a resolution λ ($\lambda=L/l$ the ratio between the outer scale L and the observation scale l) are power law related to λ:

$$\left\langle R_\lambda{}^q \right\rangle \approx \lambda^{K(q)} \qquad \text{(Eq. 5)}$$

With

$$K(q) = \frac{C_1}{\alpha - 1}\left(q^\alpha - q\right) \qquad \text{(Eq. 6)}$$

being the scaling moment function which fully characterizes the rainfall structure and variability not only at a single scale but through scales.

In this paper discrete cascades are implemented, meaning that the rainfall over a large scale structure is distributed in space and time step by step. At each step the "parent structure" is divided into several "child structures" and the intensity affected to a child structure is equal to its parent's one multiplied by a random increment. In order to ensure the validity of Eq. 5 and 6 the random multiplicative increment must be chosen as $\exp\left[\left(\frac{C_1 \ln(\lambda_0)}{|\alpha - 1|}\right)^{1/\alpha} L(\alpha)\right] / \lambda_0{}^{\frac{C_1}{\alpha - 1}}$, where λ_0 is the scale ratio between two consecutive time steps. $L(\alpha)$ is an extremal Lévy-stable random variable of Lévy stability index α (i.e. $\left\langle \exp\left(qL(\alpha)\right)\right\rangle = \exp\left(q^\alpha\right)$), which corresponds to a mathematical definition of the multifractality index. The algorithm presented in Chambers et al. (1976) was used to generate it. To be consistent with the scaling of life-time vs. the structure size in the framework of the Kolomogorov picture of turbulence (Kolmogorov, 1962) the scale of the structure is divided by 3 in space and 2 in time at each step of the cascade process (Marsan et al., 1996; Biaou et al., 2005; Gires et al., 2011), which leads to 18 child structures. A new seed is chosen at the beginning of each new realizations of a downscaled rainfall field. Finally it should be mentioned that in this paper we are focusing the analysis on selected rainfall episodes, which means that the zeros of the rainfall (on - off intermittency) do not play a significant role. Hence we did not include any process to generate additional zero values other than the small values spontaneously obtained with the help of the multiplicative cascade process itself (see Gires et al. 2013 for some suggestions on how to proceed to include additional zeroes for longer series). However we used UM parameters obtained on focusing on the rainfall episodes of the rainfall fields since analysis on large areas or long period which include many zeros lead to significantly biased estimates (de Montera et al. 2009, Gires et al. 2012a).

3.2) Rainfall data from dense networks of point measurements

- EPFL network of disdrometers in Lausane (Switzerland)

A network of 16 autonomous optical disdrometers (first-generation Parsivel, OTT) was deployed over EPFL campus from March 2009 to July 2010 (see Jaffrain et al., 2011, for more detailed information). The minimum distance between 2 disdrometers was about 8 m, the maximum one about 800 m. The measured spectra of raindrop size distribution (DSD) have been used to derive the rain rate at a 1-min temporal resolution. The processing of the DSD data is described in Jaffrain et al., 2011. We selected a set of 36 rainfall events for which the bias between a disdrometer and a collocated rain gauge was below 10% over the total rainfall amount (see Jaffrain and Berne, 2012 for details). Out of these 36 events, we selected six having the largest rainfall amounts for the present study.

The main features of the six studied rainfall events are displayed Table 3.

- Bradford U. network of rain gauges in Bradford (United-Kingdom)

The second data set used in this paper consists in the rainfall measured by 16 tipping bucket rain gauges installed over the campus of Bradford University (United-Kingdom). Eight measuring locations with 2 co-located rain gauges are installed on the roofs of the campus, this has been done to help find random rain gauge errors as described in Ciach and Krajewski (2006). The rain gauges installed at Bradford University are type ARG100, commonly used in the UK and described in Vuerich et al. (2009), the 'WMO field intercomparison of rainfall intensity gauges' report. The ARG100 rain gauges are supplied with a calibration factor between 0.197 and 0.203 mm per tip. If the calibration would be accurate, a pair of co-located rain gauges should give near identical readings when no random errors such as blockages have occurred. A dynamic re-calibration of all rain gauges, similar to the description in the manufacturers' documentation has therefore been carried out in the laboratory. A peristaltic pump was set-up to drip 1 litre of water in the rain gauge for over 60 minutes, simulating 20 mm/hr intensity rainfall. During this re-calibration it was found that two of the purchased rain gauges lay outside the accepted range of 0.197 and 0.203 mm per tip, these rain gauges were sent back to the manufacturer to be recalibrated. For the other rain gauges it proved difficult to confirm exactly the same calibration factor in the laboratory. Repeated calibration of a single gauge could deliver a calibration factor between for example 0.199 and 0.201, whereas the factory calibration provided could be outside this interval, for example 0.198. The rain gauge data were therefore derived using the average value of calibration factor from the re-calibration carried out in Bradford. Given this information, it was deemed that the maximum difference between 2 co-located rain gauges due to potential errors in the calibration factor would be (0.204/0.196)/0.196*100% = 4.1 %, n.b. as worst case scenario a slightly wider range of 0.196 to 0.204 mm per tip was used. This 4.1 % was used as cut-off point, i.e. if a pair of co-located rain gauges shows an absolute difference, |(RG1-RG2)/RG2*100%| or |(RG2-RG1)/RG1*100%|, that is larger than 4.1% , the pair was removed from the dataset as it is likely that one of the rain gauges suffered from random errors, such as temporary blockages etc. The rain gauges were visited approximately every 5 weeks, when the gauge funnel and tipping bucket were cleaned of any debris, and notes made of any blockages. The maximum distance between two rain gauges is 404 m and the time resolution 1 min. Three rainfall events were analysed (see Table 4).

3.3) Validation, Results and discussion

The measurement devices of both Bradford U. and EPFL data sets are located within a $1km^2$ area. Hence it is possible to use them to test the suggested spatio-temporal downscaling process. This is achieved by implementing the following methodology.

First the average rain rate over the surrounding 1 km^2 area with a 5 min resolution is estimated by simply taking the arithmetic mean of the rain rates computed by the available devices over 5 min.

Then the obtained field is downscaled with the help of the process described in the section 3.1. More precisely, seven steps of discrete cascade process are implemented leading to a spatial resolution of 46 cm and a temporal one of 2.3 s. The field is then re-aggregated in time to obtain a final temporal resolution of 1 min equal to the one of the two measuring devices. The output of the process consists in a realistic (if the downscaling process is correct!) rainfall estimate for 2187x2187 virtual disdrometers (or rain gauges) located within the 1 km^2 area.

Fig. 4 displays the 5 consecutive time steps of the simulated rain rate at a resolution of 1 min in time and 46 cm in space starting with a uniform rain rate equal to 1 mm/h at the initial resolution of 5 min in time and 1 km in space. It should be mentioned that the straight lines which remind the pixelisation associated with radar data are due to the use of discrete cascades. The use of the more complex continuous cascades (see Lovejoy et Schertzer 2010 for more details on how to simulate them) would avoid this unrealistic feature in the spatio-temporal structure of the field but would not change the retrieved statistics.

Third observed and simulated data are compared with the help of the temporal evolution of rain rates simulated for the various virtual disdrometer and quantile plots. More precisely the temporal evolution of the rain rate and the cumulative rainfall depth are computed for each of the virtual disdrometer (or rain gauge). Then, instead of plotting the 2187 x 2187 curves which leads to unclear graph, for each time step the 5, 25, 75 and 95 % quantiles among the virtual disdrometers are evaluated. The corresponding envelop curves (R_5(t), R_{25}(t), R_{75}(t), R_{95}(t) for rain rate, and C_5(t), C_{25}(t), C_{75}(t), C_{95}(t) for cumulative depth) are then plotted with the recorded measurements on the same graph. The corresponding curves for the EPFL data set and the Bradford U. one are displayed respectively Fig. 5.a and Fig. 5.b. For some events the graph of the rain rates is zoomed on portion of the event to enable the reader to see the details of the curves which is not always possible if the whole event is shown on a single graph. With regards to the quantile plots, for each location all the measured data (i.e. all the stations and all the available time steps, corresponding to respectively 36 495 and 29 000 values for the EPFL and Bradford data set) is considered at once and compared with a random selection of the same number of virtual point measurements. Note that for the Bradford data set the random selection of virtual point measurements mimics the fact that the network is made of pairs of collocated rain gauges by selecting always two adjacent virtual rain gauges. Fig. 6.a and Fig 6.b provide an example of obtained quantile plots for respectively the EPFL and the Bradford data set. Similar plots are obtained for other realizations of the random selection of virtual point measurements within the square km.

Concerning the 6 June 2009 event of the EPFL data set, it appears that the disparities among the temporal evolution of the rain rate of the various disdrometers are within the uncertainty interval predicted by the theoretical model. Indeed the empirical curves are all between R_5(t) and R_{95}(t) and some are greater than R_{75}(t) or lower than R_{25}(t) for some time steps. It should be noted that for a given disdrometer the position of the measured rain rate varies within the uncertainty interval according to the time step (i.e. not always greater than R_{75}(t) for instance), which is expected if the theoretical framework is correct. Concerning the temporal evolution of the cumulative rainfall depth, the measured curves are all within the [C_5(t);C_{95}(t)] uncertainty interval except for two disdrometers. There is furthermore a significant proportion (8 out of 15) of disdrometers within the [C_{25}(t);C_{75}(t)] interval which is expected. Hence for this specific event the downscaling model can be validated overall. Similar comments can be made for the other rainfall events with may be a tendency to slightly overestimate of the uncertainty interval for the rain rate. The quantile plot for a random selection of virtual disdrometer (Fig. 6.a) confirms the overall validity of the downscaling model for rain rates lower than 60 – 70 mm/h since it follows rather well the first bisector. For the extreme values (rain rates greater than 60 – 70 mm/h which corresponds to probability of occurrence roughly lower than 10^{-3}) some discrepancies are visible and the simulated quantiles tend to be significantly greater than the observed ones. Given the validity of the UM model for rest of the curve, an interpretation of this feature could be that the measurement devices have troubles in the estimation of extreme time steps and tend to underestimate them. More extreme events should be analysed to properly confirm this. This nevertheless hints at a possible practical application of this downscaling process; generating realistic rainfall

quantiles at "point" scale. Indeed they do not seem accessible to direct observation because of both limitations in the accurate measurement of extreme rainfall and sparseness of point measurement network.

Before going on with the Bradford U. data set, let us test the sensitivity of the obtained results to the choice of UM parameters which have been set to α=1.8 and C_1=0.1 for all the events which correspond to values commonly estimated on the rainy portions of the rainfall fields. The various parameter sets tested are shown in Table 5, along with γ_s, which is a scale invariant parameter consisting of a combination of both α and C_1 and characterizing the maximum probable value that one can expect to observe a single realization of a phenomenon. It has been commonly used to assess the extremes in the multifractal framework (Hubert et al. 1993, Douglas and Barros 2003, Royer et al. 2008, Gires et al. 2011). The simulated quantiles (not shown here) are roughly similar to the ones found for α =1.8 and C_1=0.1 for all the other UM parameter sets (with a tendency to generate slightly lower ones in the range 20 – 50 mm/h) except for the α =1.8 and C_1=0.2 which generates significantly greater quantiles which are not compatible with the measured ones. For the 6th June events the spread in the simulated cumulative curves was also quantified for the various UM parameter set with the help of CV'_{95} defined as:

$$CV'_{95} = \frac{C_{95}(t_{end}) - C_5(t_{end})}{2 * C_{mean}(t_{end})}$$

Where t_{end} is the last time step of the event and $C_{mean}(t)$ the average temporal evolution of the cumulative rainfall depth over the 16 disdrometers of the network. The values are reported in Table 5, and should be compared with the ratio between the maximum observed depth minus the minimum one divided by twice the average one which is equal to 18% for this event (since we have 16 disdrometers values slightly lower than this one are expected). When α is fixed it appears that CV'_{95} increases with C_1, which was expected since it corresponds to stronger extremes (γ_s also increases). The same is observed when C_1 is fixed and α increases although is appears that the influence of variations of α have a much less significant impact on the computed CV'_{95}. It should be noted that the observed CV'_{95} cannot be interpreted only with the help of γ_s (indeed for α=1.6 and C_1=0.1 we have γ_s = 0.22 and CV'_{95}=11% whereas for α =1.8 and C_1=0.05 we have greater γ_s and lower CV'_{95}) which means that both parameters are needed. Similar results are found for the other rainfall events. Although likely to be oversimplifying the choice of constant UM parameters set to α=1.8 and C_1=0.1 for all the events appears to be acceptable.

The results for the Bradford data set (Fig. 5.b) have a less straightforward interpretation. Indeed the discrete nature of the measurement with tipping bucket rain gauges makes is hard to analyse the results for the rain rates at a 1 min resolution. For example during the 22 June event the rain rate seldom exceeds the one corresponding to one tip in a min (i.e.: 12 mm/h) suggesting that the 1 min resolution is not adequate for this study which is why the temporal evolution of the rain rate was also plotted with a 5 min resolution. The other two event exhibit greater rain rates and the effects of the discretisation are dampened, suggesting that there is no need to analyse the rain rates with a 5 min resolution. Overall it seems that the downscaling model reproduces rather well the observed disparities between the rain gauges. With regards to the cumulative rainfall depth the disparities between the rain gauges are consistent with the theoretical expectations for the 22 June event (except for two rain gauges), and smaller for the other two events. This behaviour is quite different from the one observed with the EPFL data set. It is not clear whether this difference is due to the fact that two measurement devices are

used (suggesting either that the rain gauges artificially dampens the actual disparities or that the disdrometers artificially strengthen it because of instrumental errors) or because the downscaling model is less adapted for two of the Bradford events (6 July and 22 of August). The quantile plot (Fig. 6.b) is harder to interpret because of the discrete nature of the measurements. The seven horizontal segments correspond to measurements of 1 to 7 tips in a minute (there are several point on each segment because not all the rain gauges have the same calibration factor). One can only note that it seems that the simulated quantiles start to tend to be greater than the measured ones for rain rates smaller (20 - 30 mm/h) than with the disdrometers. Following the interpretation given for the EPFL data set, it would mean that the rain gauges start to underestimates rain rates for lower values.

Finally let us remind that the tested downscaling is a very simple and parsimonious one consisting in stochastically continuing an under-lying multifractal process defined with the help of only two parameters which are furthermore considered identical for all the events and locations. The fact that the observed disparities between point measurement for very dense networks of either disdrometers or rain gauges are in overall agreement with the theoretical expectations is a great achievement. It might be possible to refine the model by using different UM parameters according to the event, but the underlying rainfall theoretical representation should be improved first. As a conclusion it appears that although not perfect this very simple and parsimonious model is robust and it is relevant to use it for the purpose of this paper which is to revisit the representativeness issue on standard comparison scores between point and areal measurements.

4) Impact of small scale rainfall variability on the standard indicators

4.1) Methodology

The aim of this section is to estimate the expected values of the scores if neither radar nor rain gauges were affected by instrumental error, and the deviations from the optimum values were only due to the small scale rainfall variability. We will also investigate the related issue of the variations of the scores depending on where the rain gauges are located within their respective radar pixel. We remind that the studied data set is made of the rainfall output of 26 rain gauges and their corresponding radar pixels for four events. In order to achieve this we implement the following methodology:

(i) Downscaling the radar data for each radar pixels to a resolution of 46 cm in space and 5 min in time which is similar to the rain gauge resolution. This is done by implementing 7 steps of the spatio-temporal downscaling process validated in the previous section and re-aggregating it in time. This yields the outputs of 2187×2187 "virtual rain gauges" for each of the 26 radar pixels.

(ii) Randomly selecting a "virtual rain gauge" for each radar pixel and computing the corresponding scores. In order to generate a distribution of possible values for each score, 1000 sets of 26 virtual rain gauges locations (one per radar pixel) are tested

4.2) Results and discussion

The distributions of the scores obtained for the 1000 samples of virtual rain gauges set are displayed on Fig. 3 for time steps of 5, 15 and 60 min, with the numerical value of the 5, 50 and 95% quantile. An example of scatter plot with a set of "virtual" rain gauges is visible on Fig. 2.

The 50% quantile for each score provides an estimation of the expected value if neither radar nor rain gauges are affected by instrumental errors. The differences with regards to the optimal values of scores are simply due to the fact that rainfall exhibits variability at small scales (i.e. below the observation scale of C-band radar in this paper) and that radar and rain gauge do not capture this field at the same scale. Practically it means that when a score is computed with real data (i.e. affected by instrumental errors), its value should not be compared with the theoretical optimal values but to the ones displayed on Fig. 3, which is never done. The extent of the distributions, which can be characterized with the help of the difference between the 5 and 95%, reflects the uncertainty on the scores associated with the position of the rain gauges in their corresponding radar pixel. Practically it means that when comparisons of scores are carried out with real data, as it is commonly done to compare the accuracy of the outputs of various radar quantitative precipitation estimation algorithm for example, the observed differences in the scores should be compared with this uncertainty to check whether they are significant or not. This is never done and could lead to qualifying the conclusions of some comparisons.

The values that should be used as reference (i.e. 50% quantile found considering only consequences of small scale rainfall variability) are displayed Fig. 3. Some of them are significantly different from the optimal values and as expected the difference is greater for small time steps which are more sensitive to small scale rainfall variability. For instance for 15 min the 50% quantile is equal 0.91, 0.81 and 79 for respectively the *corr*, *Nash* and $\%_{1.5}$ scores. The values for the *slope* are also smaller than one (0.82, 0.90 and 0.96 for respectively 5, 15 and 60 min time steps), which was not necessarily expected.

With regards to the scores computed for the 4 studied events over the Seine-Saint-Denis County, it appears that independently of the event and time step the scores found for *Nash*, $\%_{1.5}$ and *corr* are not consistent with the idea that they are only due to small scale rainfall variability, meaning that instrumental errors affected the measurement. For *RMSE* the scores found are explained by small scale rainfall variability for 15 Aug. 2010 and 15 Dec. 2011 and almost for 14 Jul. 2010. For the event of 9 Feb. 2009 the observed RMSE is even lower than the values of the distribution of the "virtual gauges". This is quite surprising since this distribution is a lower limit (for instance instrumental errors are not taken into account), which suggest some error compensation for this specific case. For *NB* and *slope* we find that the computed scores reflect instrumental errors for 9 Feb. 2009 and 14 Jul. 2010. It is also the case for the other two events with a time step of 5 min, but not with a 1 h time step.

In the methodology developed previously two steps are random and it is therefore important to check the sensitivity of the results to them. The first one is the downscaling of the radar pixels where the rain gauges are located. The sensitivity is tested by simulating a second set of downscaled rainfall fields. The second one is the selection of the virtual rain gauges that are used to compute the distributions displayed Fig. 3. Indeed in the downscaling process 2187×2187 virtual rain gauges are generated for each of the studied 26 radar pixels,

leading to $(2187\times2187)^{26}$ ($\approx 4.7\times10^{173}$) possible combinations. A set of 1000 combinations is used to generate the studied distributions. To test the sensitivity, the distributions are assessed for two more sets for each of the two downscaled rainfall fields. Hence a total of 6 samples for each score are tested to analyse this sensitivity issue. Figure 7 displays the cumulative probability function for the *NB* with a 15 min time steps which is representative of other time steps and scores. Visually it seems that the obtained distributions are very similar. It is possible to confirm this assertion more quantitatively with the help of a two samples Kolmogorov-Smirnov test (Massey, 1951) which is commonly used to check whether two samples are generated with the help of the same distribution. The null hypothesis that the samples are from the same underlying distribution is tested for the 6 samples two by two. It is rejected with a 95% confidence interval only 27 times out of 378 tests (378 = 14 tests for the 6 samples two by two x 9 scores x 3 time steps). This confirms the first impression that the 6 samples reflect similar distributions, and that the results previously discussed are robust and not sensitive to the random steps of the downscaling process and the selection process of the virtual rain gauges.

The sensitivity of the results to the choice of the UM parameter α=1.8 and C_1=0.1 was also tested as in the section 3.3. Fig. 9 displays the cumulative probability distribution of for the *Nash* and $\%_{1.5}$ scores for the same sets of UM parameters as in the section 2.3 (see Table 5). The same comments remain valid, i.e.: with a fixed parameter, the greater is the other one the worst is the indicator, C_1 has a stronger influence than α on the computed uncertainty. Therefore, both parameters are needed and results cannot be interpreted only with the help of γ_s. It can be added that the worst is the indicator the widest is the probability distribution. Similar results are found for the other scores. It appears that the values of the UM parameters used for the simulations have a strong influence on resulting cumulative distributions, suggesting that for practical applications the parameters should be carefully estimated with a particular emphasis on C_1.

Besides redefining the optimum of standard scores and setting values to which score variations should be compared, this work also suggests changing common practice when temporal evolution of rain rate or cumulative rainfall depth observed by rain gauge or disdrometer and the corresponding radar pixel are plotted on the same graph. This is the last standard way of comparing the output of the two measurement devices to be addressed in this paper. The observation scale gap between the two devices should be visible directly on the plot. A way of achieving this is to explicitly display the range of “realistic” values at the rain gauge scale for a given radar pixel measurement, in order to give an immediate insight into this issue to the reader and suggest whether to look for other explanations than small scale rainfall variability. This is currently not done in usual comparison. We propose to proceed as in the section 3 and to plot the 5, 25, 50 and 95 % quantiles for both rain rate and cumulative depth along with the radar curves. This is done in Fig. 8 for the 9 Feb. 2009 and 14 Jul. 2010 rainfall event for one rain gauge. For the 9 Feb. 2009, the cumulative depth (Fig. 8.b) is clearly outside the uncertainty range of the radar measurement at rain gauge scale meaning the instrumental error are likely to have affected at least one of the devices. Concerning the 14 Jul. 2010, the rain gauge cumulative depth is in agreement with the radar measurement (Fig. 8.d). With regards to the rain rate (Fig. 8.c), the rain gauge measurements are in the lower portion of the realistic values for the first peak, outside of it for the second peak (suggesting the effect of instrumental errors), and in the upper one for the third peak. More generally these results suggest that to compare the measurements of two devices that observe the same physical phenomenon at two different scales, it should become a common practice to first simulate an ensemble of realistic outputs at the smallest available scale of observations among

two devices, and to compare the latter's output to the generated ensemble. The example of radar and rain gauge was discussed in this paper, but similar techniques could also be implemented on the comparison between satellite and radar data that also do not correspond to the same scale, while with a smaller scale gap compare to the case discussed in this paper.

These results also hint at some ways of revisiting standard interpolation and merging techniques that, in spite being beyond the main scope of this paper, can take advantage from obtained herein results. Indeed the validity of a UM model of rainfall down to very small scale suggests that developing a multifractal interpolation algorithm would be feasible. Some basic ideas on how to proceed can be found in Tchiguirinskaia et al. 2004, but there is still some work to be done in order to have an operational algorithm. Of course the output of such process would not be a single field but an ensemble of realistic fields, conditioned by the observed rainfall data. With regards to the merging between radar and rain gauge data the work here also suggest some new ideas. Indeed in this paper the rainfall at the rain gauge scale was simulated from the radar, but since we have validated a mathematical representation of rainfall between the two observation scales it is possible to do the inverse. More precisely, it would also be possible to compute an ensemble of possible radar values that could result in the observed data at the rain gauge scale. Such information could be used to modify in new ways the radar measurements according to the rain gauge data, which is a common step of merging techniques.

5) Conclusion

In this paper the issue of representativeness of point measurement with regards to larger scale measurements is revisited in the context of comparison between rain gauge and radar rainfall measurement. More precisely the influence the small scale rainfall variability occurring below the radar observation scale (1 km in space and 5 min in time here) on the standard comparison scores is investigated. It appears that this influence is twofold. First the target values of the scores are not the optimum ones because rainfall variability "naturally" worsens them. This worsening, which is neglected in numerous published comparisons, is significant. Second, because of the random position of the point measurements within a radar pixel there is an expected uncertainty on a computed score. The two effects are quantified in details on a case study with radar and rain gauge data from the 237 Km^2 Seine-Saint-Denis County (France) and appears to be significant. This result is assessed with the help of a robust methodology relying on an explicit theoretical representation of the small scale rainfall variability not grasped by the radar (C-band one here). Indeed the parsimonious Universal Multifractals, which rely on only two parameters furthermore not event based in this study, are used to perform a realistic downscaling of the radar data to the point-measurement scale. This downscaling process is validated with the help of two very dense (i.e. 16 within an area of 1 km^2) networks of disdrometers in Lausane (Switzerland) and rain gauges in Bradford (United-Kingdom). The disparities observed between the point measurements are in agreement with the theoretical expectations.

The two effects of small scale rainfall variability indentified on standard comparison tools are unfortunately usually not taken into account by meteorologists and hydrologists when they carry out standard comparison to either evaluate new radar quantitative estimation precipitation algorithms or compare two. Doing it could lead to qualifying some otherwise

straightforward conclusion. The results obtained on this case study show that the assessed values for standard scores are not fully explained by small scale rainfall variability. This means that a methodology to properly distinguish the instrumental error from the representativeness issue should be developed within that framework of multifractal modelling of rainfall. The validation of a downscaling process is also a first step in improving existing merging techniques between the two rainfall measurements devices which can help in providing the accurate fine scale rainfall needed for urban hydrology applications. Further investigation would be needed to achieve these two aims.

Acknowledgements

The authors acknowledge Météo-France and especially Pierre Tabary and Valérie Vogt, and "Direction Eau et Assainissement" of Seine-Saint-Denis and especially Natalija Stancic and François Chaumeau, for providing respectively the radar rainfall estimates and the rain gauges data in an easily exploitable format. The authors greatly acknowledge partial financial support form the Chair "Hydrology for Resilient Cities" (sponsored by Veolia) of Ecole des Ponts ParisTech, and EU INTERREG RainGain Project. The purchase of the Bradford rain gauges was made possible through a research grant (RG110026) from The Royal Society.

References

Berne, A., Delrieu, G., Creutin, J.D. and Obled, C., 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. Journal of Hydrology 299 (3-4), 166-179.

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

Ciach, G.J., Habib, E. and Krajewski, W.F., 2003. Zero-covariance hypothesis in the error variance separation method of radar rainfall verification. Advances in Water Resources 26 (5), 573-580.

Ciach, G.J. and Krajewski, W.F., 1999. On the estimation of radar rainfall error variance. Advances in Water Resources 22 (6), 585-595.

Ciach, G.J. and Krajewski, W.F., 2006. Analysis and modeling of spatial correlation structure in small-scale rainfall in Central Oklahoma. Advances in Water Resources. 29 (10), 1450-1463.

de Lima, M.I.P. and Grasman, J., 1999. Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. Journal of Hydrology 220 (1-2), 1-11.

de Montera, L., Barthes, L., Mallet, C. and Gole, P., 2009. The Effect of Rain-No Rain Intermittency on the Estimation of the Universal Multifractals Model Parameters. Journal of Hydrometeorology 10 (2), 493-506.

Diss, S. et al., 2009. Ability of a dual polarized X-band radar to estimate rainfall. Advances in Water Resources 32 (7), 975-985.

Douglas, E.M. and Barros, A.P., 2003. Probable maximum precipitation estimation using multifractals: Application in the eastern United States. Journal of Hydrometeorology, 4(6), 1012-1024.

Emmanuel, I., Andrieu, H. and Tabary, P., 2012. Evaluation of the new French operational weather radar product for the field of urban hydrology. Atmospheric Research 103, 20-32.

Figueras i Ventura, J., Boumahmoud, A.-A., Fradon, B., Dupuy, P. and Tabary, P., 2012. Long-term monitoring of French polarimetric radar data quality and evaluation of several polarimetric quantitative precipitation estimators in ideal conditions for operational implementation at C-band. Quarterly Journal of the Royal Meteorological Society.

Gires, A., Tchiguirinskaia, I., Schertzer, D. and Lovejoy, S., 2011. Analyses multifractales et spatio-temporelles des precipitations du modele Meso-NH et des donnees radar. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques 56 (3), 380-396.

Gires, A., Tchiguirinskaia, I., Schertzer, D. and Lovejoy, S., 2012a. Influence of the zero-rainfall on the assessment of the multifractal parameters. Advances in Water Resources, 45, 13-25.

Gires, A., Tchiguirinskaia, I., Schertzer, D. and Lovejoy, S., 2012b. Multifractal analysis of an urban hydrological model on a Seine-Saint-Denis study case. Urban Water Journal.

Gires, A., Tchiguirinskaia, I., Schertzer, D. and Lovejoy, S., 2013. Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework. Nonlinear Processes in Geophysics 20(3), p. 343-356.

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

Hubert, P., 2001. Multifractals as a tool to overcome scale problems in hydrology. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques 46 (6), 897-905.

Jaffrain, J. and A. Berne, 2011. Experimental Quantification of the Sampling Uncertainty Associated with Measurements from PARSIVEL Disdrometers. Journal of Hydrometeorology 12 (3), 352-370.

Jaffrain, J. and A. Berne, 2012. Quantification of the Small-Scale Spatial Structure of the Raindrop Size Distribution from a Network of Disdrometers. Journal of Applied Meteorology and Climatology 51 (5), 941-953.

Kolmogorov, A.N., 1962. A refinement of previous hypotheses concerning the local structure of turbulence in viscous incompressible fluid at high Reynolds number. J. Fluid. Mech. 83, 349.

Krajewski, W.F., Villarini, G. and Smith, J.A., 2010. Radar - Rainfall uncertainties : Where are We after Thirty Years of Effort? Bulletin of the American Meteorological Society 91 (1), 87-94.

Ladoy, P., Schmitt, F., Schertzer, D. and Lovejoy, S., 1993. The multifractal temporal variability of nimes rainfall data. Comptes Rendus de l'Academie Des Sciences Serie II 317(6), 775-782.

Mandapaka, P.V., Lewandowski, P., Eichinger, W.E. and Krajewski, W.F., 2009. Multiscaling analysis of high resolution space-time lidar-rainfall. Nonlinear Processes in Geophysics 16 (5) 579-586.

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

Massey, F.J., 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association 46(253), 68-78.

Moreau, E., Testud, J. and Le Bouar, E., 2009. Rainfall spatial variability observed by X-band weather radar and its implication for the accuracy of rainfall estimates. Advances in Water Resources 32 (7), 1011-1019.

Royer, J.-F., Biaou, A., Chauvin, F., Schertzer, D. and Lovejoy, S., 2008. Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. C.R Geoscience, 340, 431-440.

Schertzer, D. and Lovejoy, S., 1987. Physical modelling and analysis of rain and clouds by anisotropic scaling and multiplicative processes. J. Geophys. Res. 92(D8), 9693-9714.

Schertzer, D. and Lovejoy, S., 1992. Hard and soft multifractal processes. Physica A 185(1-4): 187-194.

Schertzer D, Tchichuirinskaia, I., Lovejoy, S. and Hubert P., 2010. No monsters, no miracles: in nonlinear sciences hydrology is not an outlier! Hydrological Sciences Journal 55(6), 965 – 979

Schertzer, D. and Lovejoy, S., 2011. Multifractals, generalized scale invariance and complexity in geophysics. International Journal of Bifurcation and Chaos 21 (12), 3417-3456.

Tabary, P., 2007. The new French operational radar rainfall product. Part I: Methodology. Weather and Forecasting 22 (3), 393-408.

Tchiguirinskaia, I., Schertzer, D., Bendjoudi, H., Hubert, P., and Lovejoy, S., 2004. Multiscaling geophysics and sustainable development. Scales in Hydrology and Water Management, IAHS Publ. 287, 113-136.

Verrier, S., de Montera, L., Barthes, L. and Mallet, C., 2010. Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem. Journal of Hydrology 389(1-2), 111-120.

Vuerich, E., Monesi, C., Lanze, L.G., Stagi, L., Lanzinger, E., 2009. World meteorological organization, Instruments and observing methods, Report Nr 99 'WMO field intercomparison of rainfall intensity gauges'.

Wilson, J.W. and Brandes, E.A., 1979. Radar Measurement of Rainfallâ€"A Summary. Bulletin of the American Meteorological Society 60 (9), 1048-1058.

Zhang, Y., Adams, T. and Bonta, J.V., 2007. Subpixel-Scale Rainfall Variability and the Effects on Separation of Radar and Gauge Rainfall Errors. Journal of Hydrometeorology 8 (6), 1348-1363.

Figures

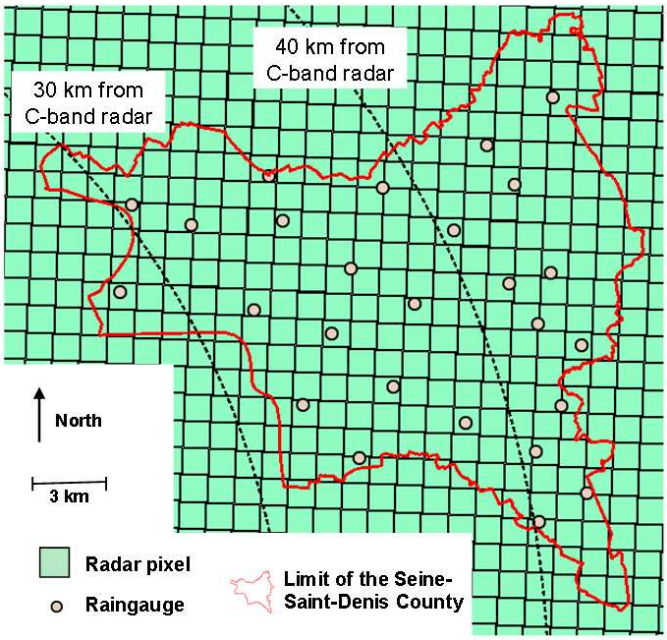


Figure 1: Map of the 26 rain gauges of Seine-Saint-Denis used in this study, with the radar pixels of the Météo-France mosaic

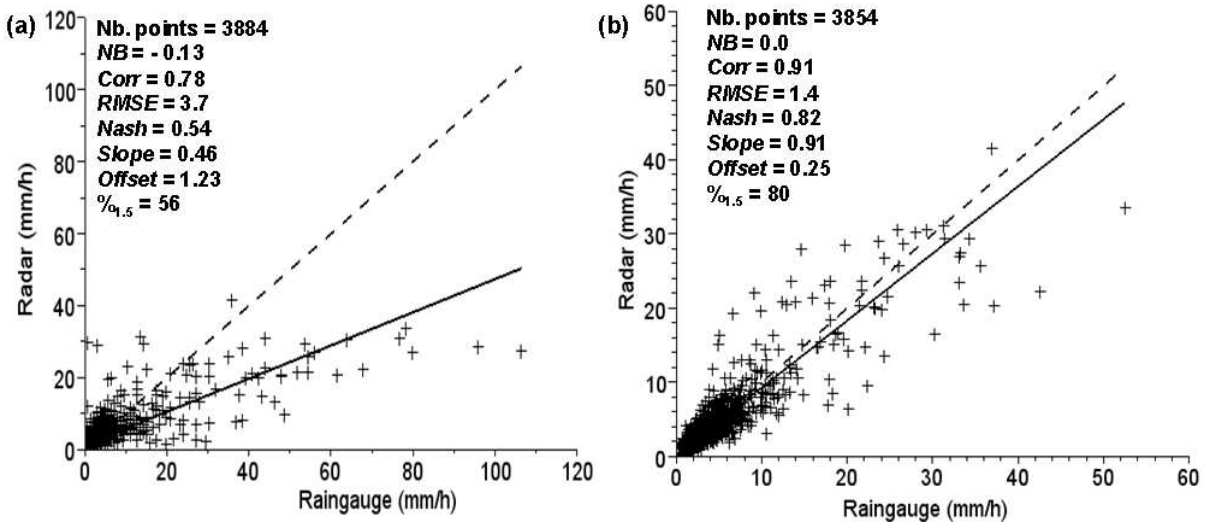


Figure 2: Scatter plot for all the events with a 15 min time steps: (a) Radar vs. rain gauges measurements, (b) Radar vs. a set of virtual rain gauges (one per radar pixels)

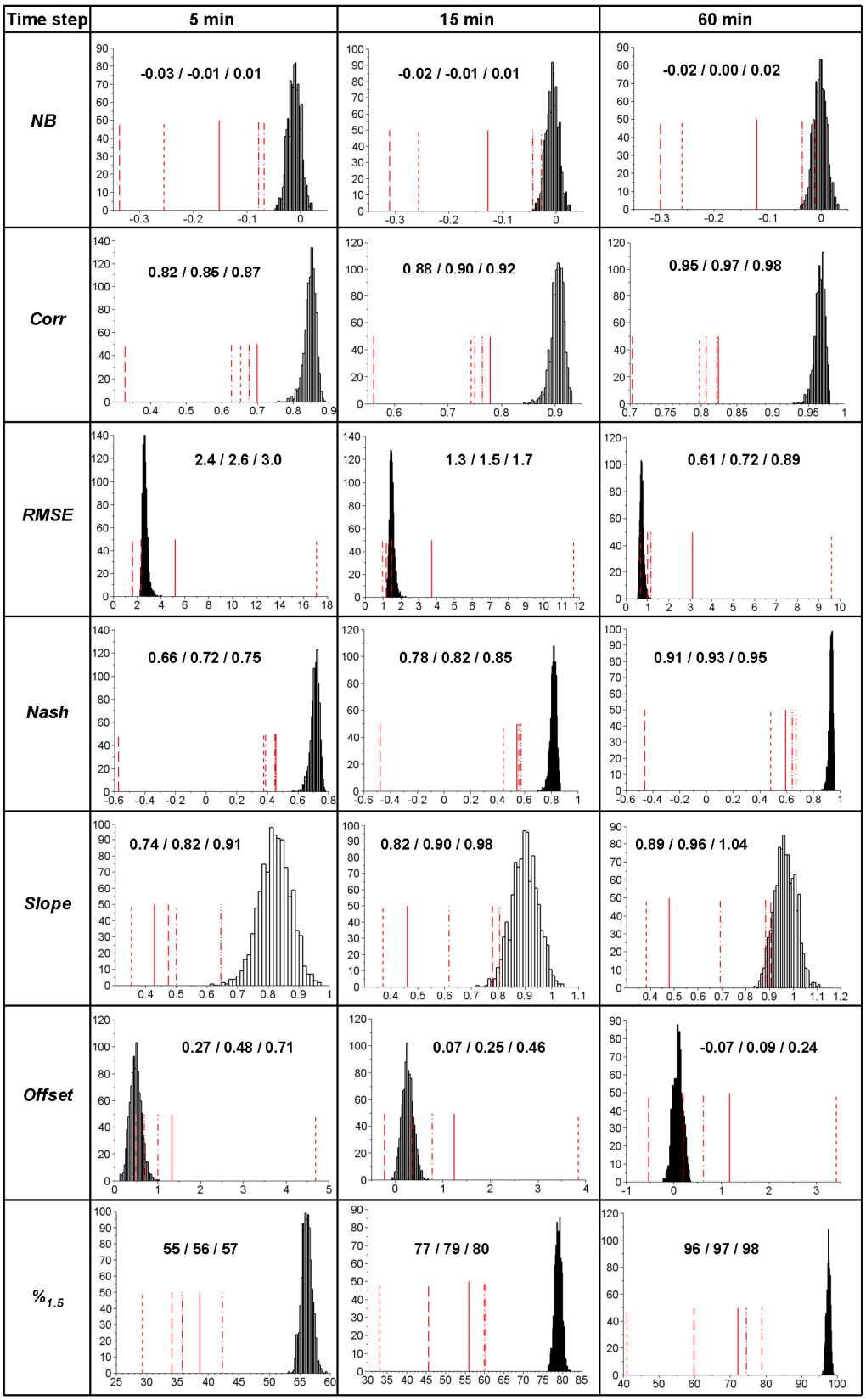


Figure 3: Histograms of the scores computed for the 1000 samples of possible combinations of virtual rain gauges. The values of the scores for all the events (solid) and the event of 9 Feb. 2009 (long dash), 14 Jul. 2010 (dash), 15 Aug. 2010 (dash dot) and 15 Dec. 2011 (dash bi-dot) are also displayed in red. The three figures associated with each distribution are the 5, 50 and 95% quantile.

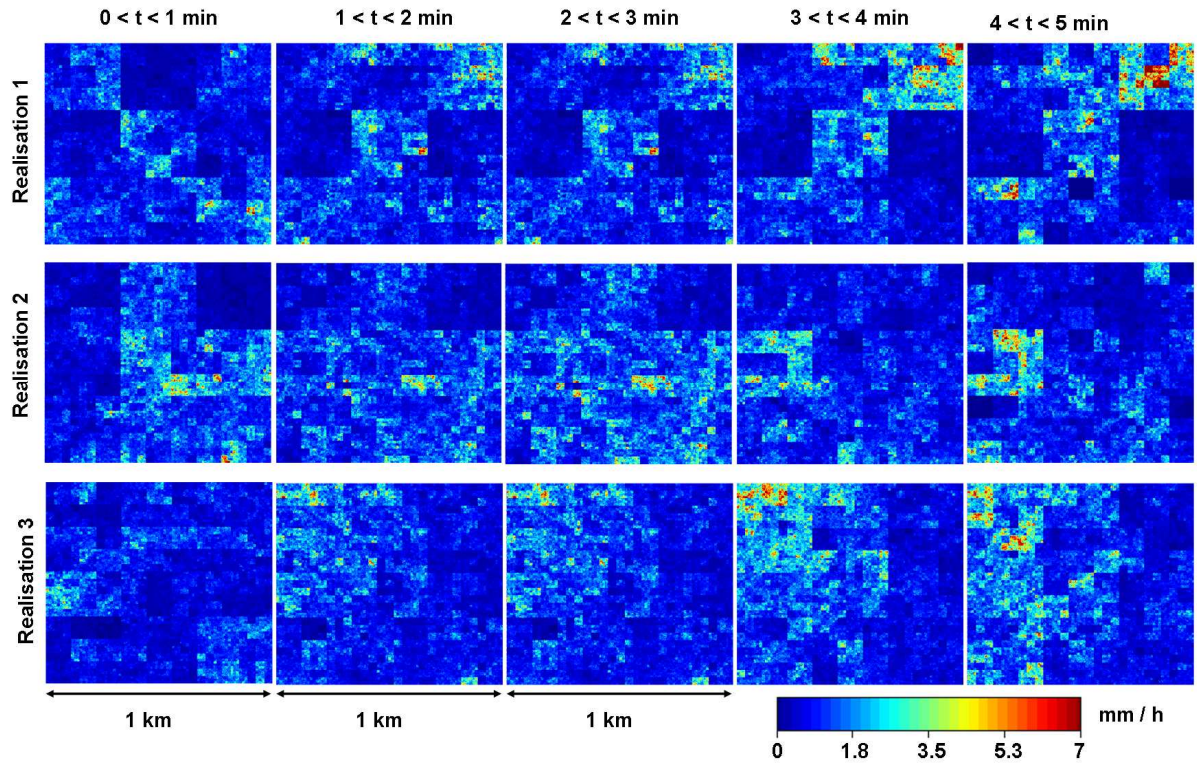


Figure 4: Illustration of the spatio-temporal downscaling process. Three realisations of the simulated rain rate with a resolution of 1 min in time and 46 cm in space starting with a uniform rain rate of 1 mm/h at the initial resolution of 5 min in time and 1 km in space.

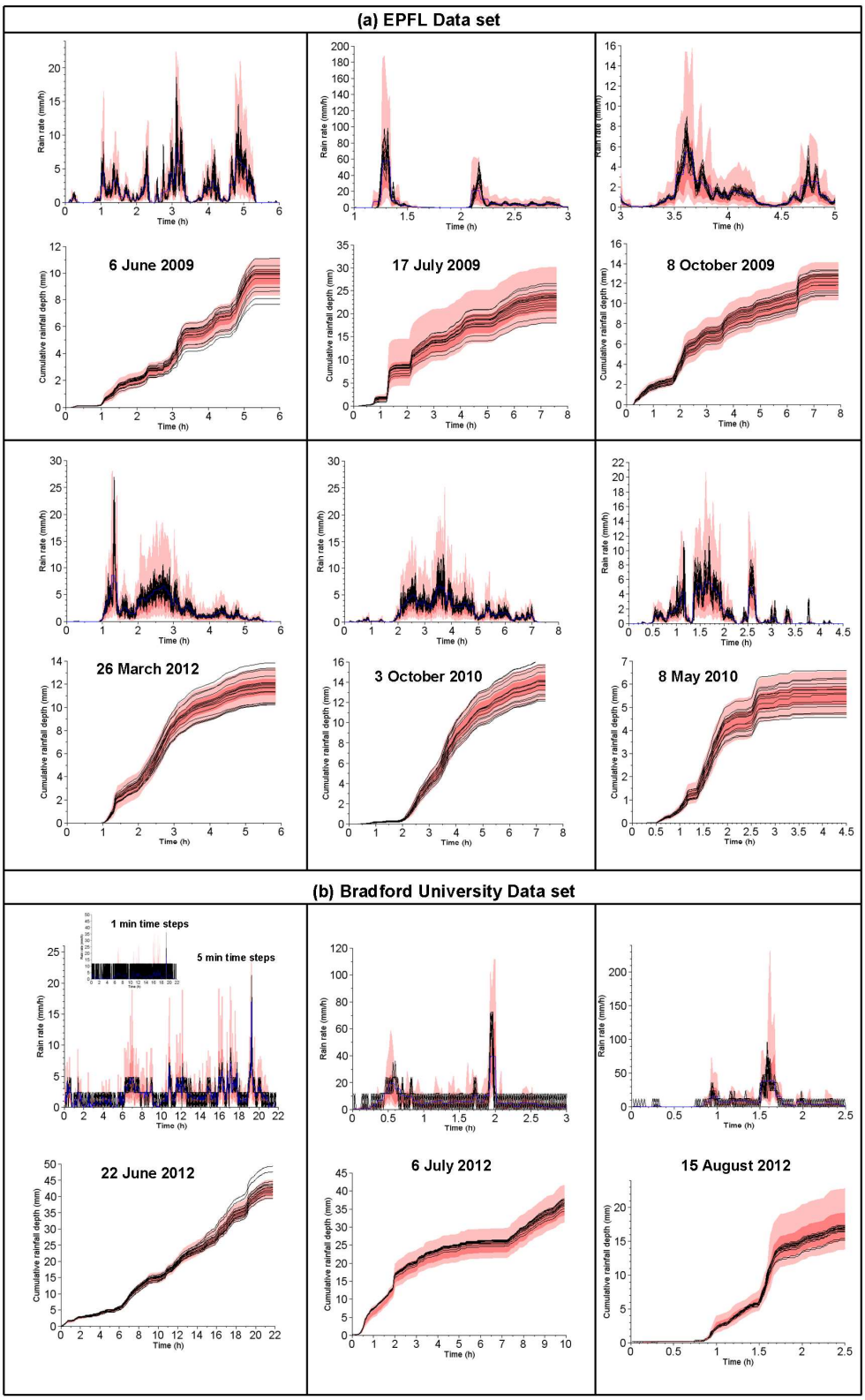


Figure 5: Temporal evolution of the rain rate and cumulative rainfall depth for point measurements for the EPFL data set (a) and the Bradford University data set (b). For each event the uncertainty range of the average measurement at the disdrometers or rain gauge observation scale is displayed (R_{25}(t) - R_{75}(t) or R_{25}(t) - R_{75}(t) and C_5(t) - C_{95}(t) or C_5(t) - C_{95}(t) are the limit of respectively the dark and the light area). Average measurement with 5 min resolution in blue.

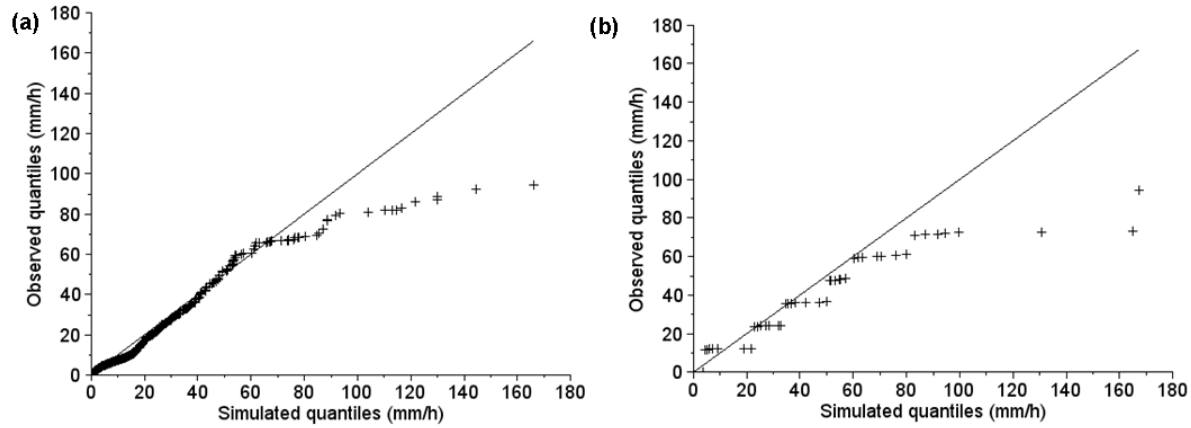


Figure 6: Quantile plot (including all the stations and all the available time steps) of the measured data versus a realisation of downscaled rainfall fields for the EPFL (a) and Bradford (b) data set.

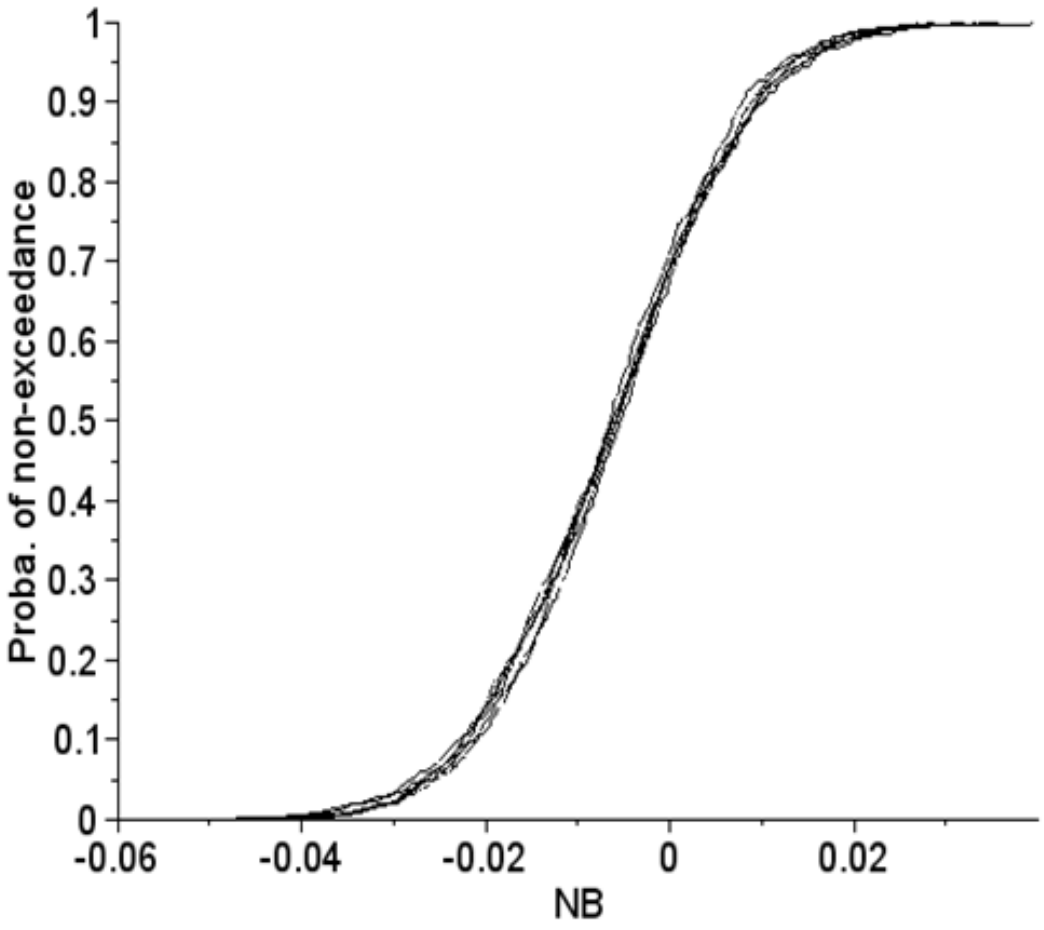


Figure 7: Cumulative probability functions for the *NB* with a 15 min time steps of the 6 samples generated to test the sensitivity of the results to the downscaling process and the selection process of the virtual rain gauges.

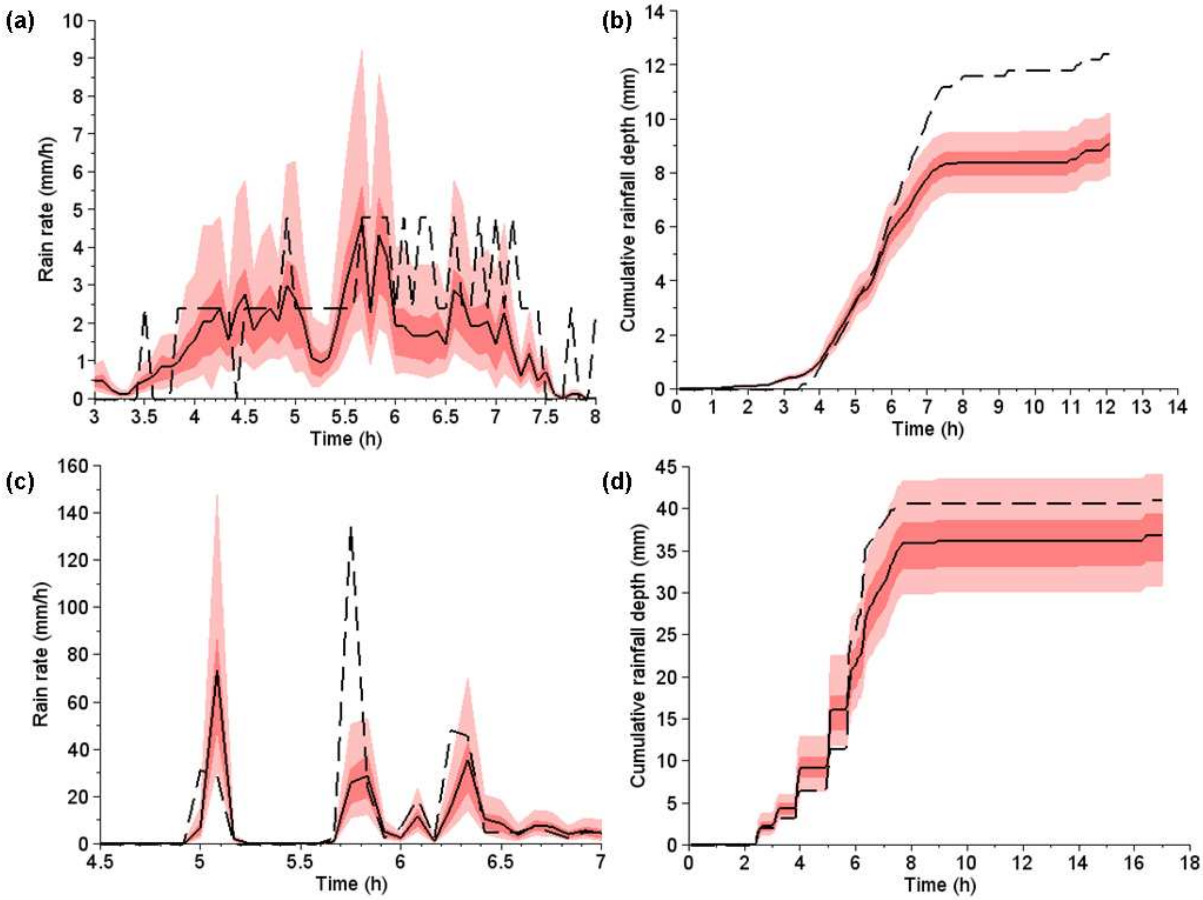


Figure 8: Rain gauge (dash), radar (solid), and uncertainty range of the radar measurement at the rain gauge scale (same as in Fig. 6) for 9 Feb. 2009 (top) and 14 Jul. 2010 (bottom) with the Seine-Saint-Denis data set.

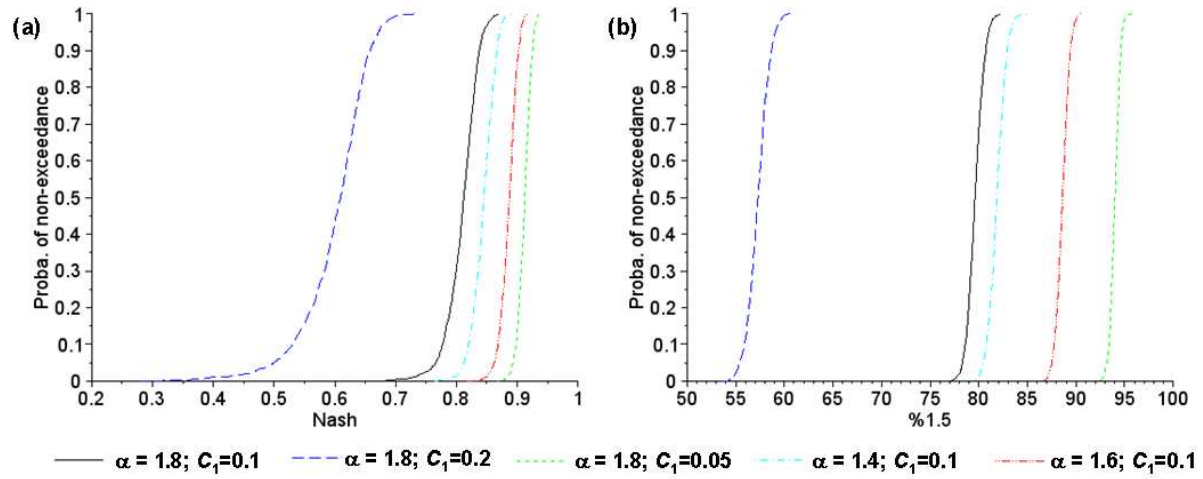


Figure 9: Cumulative probability functions for the *Nash* (a) and $\%_{1.5}$ (b) scores with a 15 min time steps for 5 different sets of UM parameters inputted to the downscaling process.

Tables

Table 1: General features of the studied rainfall events in Seine-Saint-Denis. For the cumulative depth the three figures correspond to the average, the maximum, and the minimum over the rain gauges or the corresponding radar pixels

	9 Feb. 2009	14 Jul. 2010	15 Aug. 2010	15 Dec. 2011
Approx. Event duration (h)	9	6	30	30
Available gauges	24	24	24	26
Gauge cumul. depth (mm)	11.4 (10 - 12.8)	37.9 (47.8 – 23.4)	50.1 (62.8 – 27.4)	22.4 (28.2 – 18.2)
Radar cumul. Depth (mm)	8.5 (9.3 – 7.5)	28.7 (35.8 – 21.2)	50.6 (59.2 – 36.0)	22.4 (28.2 – 19.8)

Table 2: Standard scores for the comparison between radar and rain gauges data for the 4 studied events. Only the time steps with one of data type exhibiting a rain rate greater than 1 mm/h are considered.

Score	5 min	15 min	60 min
Nb. of points	11412.	3884.	991.
NB	- 0.15	- 0.13	- 0.12
Corr	0.70	0.78	0.82
RMSE	5.19	3.71	3.09
Nash	0.46	0.54	0.59
Slope	0.43	0.46	0.48
Offset	1.33	1.24	1.17
$\%_{1.5}$	38.7	55.9	72.1

Table 3: Same as in Table 1 for the studied rainfall event in Lausane (EPFL data set)

	6 June 2009	17 July 2009	8 October 2009	26 March 2010	3 April 2010	5 August 2010
Approx. Event duration (h)	6	7.6	7.9	5.8	7.3	4.5
Nb of selected disdrometers	15	16	15	16	16	15
Disdrometer cumul. depth (mm)	9.7 (11.1 – 7.6)	22.9 (26.5 – 18.0)	12.2 (13.4 – 10.8)	11.8 (13.8 – 10.2)	14.0 (16.2 – 12.1)	5.5 (6.6– 4.6)
Maximum % difference between all selected disdrometers	46	47	24	35	34	43

Table 4: as in Table 1 for the studied rainfall events and selected rain gauge data in Bradford (Bradford University data set)

	22 June 2012	6 July 2012	15 August 2012
Approx. Event duration (h)	24	10	3
Nb of selected gauges	14	14	16
Maximum % difference within pairs of selected co-located rain gauges	4.1%	2.0%	3.7%
Gauge cumul. depth (mm)	43.2 (49.4 – 39.4)	36.7 (38.0 – 34.5)	16.8 (17.4 – 15.2)
Maximum % difference between all selected rain gauges	25%	10%	14%

Table 5: Sensitivity test to the values of the UM parameters for the 6 June 2009 of the EPFL data set

α	C_1	γ_s	CV'_{95} (%)
1.8	0.1	0.50	14
1.8	0.05	0.36	9.2
1.8	0.2	0.67	26
1.4	0.1	0.42	12
0.6	0.1	0.22	11