European Water 57: 279-285, 2017.

Integration of rain gauge errors in radar-rain gauge merging techniques

F. Cecinati[1*], A. Moreno Rodenas[2] and M.A. Rico-Ramirez[1]
[1] *University of Bristol, Department of Civil Engineering*
[2] *Technische Universiteit Delft, Water Management Department*
[*] *e-mail: fc14509@bristol.ac.uk*

Abstract: Kriging interpolation methods are often applied without considering measurement uncertainty, but rarely rain gauge uncertainty is negligible. The most used technique to consider measurement uncertainty in Kriging is to include a nugget effect in the variogram model. The nugget is a measurement of the variance that one can expect between two measurements at short distance. Including a nugget in the variogram model is appropriate when the uncertainty is stationary in space; if a time-variant variogram is not used, the uncertainty is considered stationary in time as well. This is not the case for rain gauge interpolation, because rain gauge uncertainty is proportional to the rainfall rate, which is highly variable in space and time. Kriging for Uncertain Data (KUD) is a technique proposed in literature to solve this problem, modifying the nugget effect for each time step and for each measurement point. Nevertheless, the application to rain gauge interpolation showed that the methodology is not stable in certain conditions. Several factors can affect the performance: the position and the number of rain gauges, their accuracy, spatial characteristics of the rainfall field, rainfall intensity, or the relative position of peaks compared to the rain gauge position. This work investigates the performance of KUD with three synthetic experiments, simulating the variation of rainfall spatial variability, rain gauge density, and rain gauge accuracy. The results suggest that rain gauge density in relation to the rainfall variability plays an important role in the performance of KUD.

Key words: rain gauge interpolation, kriging, kriging for uncertain data

1. INTRODUCTION

Rain gauge measurements are probably the most used source of rainfall information for hydrological applications. Often rain gauge interpolations are accomplished neglecting measurement uncertainties (AghaKouchak et al., 2010; Ciach et al., 2007; Dai et al., 2014; Germann et al., 2009; Rico-Ramirez et al., 2015; Villarini and Krajewski, 2009). Rain gauge uncertainty is reduced when temporal accumulation is performed, when more expensive high-accuracy devices are used and when the data is correctly managed and calibrated (Habib et al., 2004, 2008; Molini et al., 2005; Nešpor and Sevruk, 1999; Sevruk, 1996). Ideally, accurate and dense rain gauge networks should be used for hydrological applications, especially at urban scale, when high spatial resolutions is required (Schilling, 1991). In reality, rain gauge networks, especially high-accuracy device networks, are rarely sufficiently dense. To improve the spatial resolution, rain gauge networks with different accuracy characteristics are often integrated (Peleg et al., 2013; Villarini et al., 2008). The integration of accurate and less accurate rain gauges networks is often unavoidable to reach a sufficient density and measurement uncertainty cannot be neglected. Additionally, rain gauge uncertainty is proportional to the rainfall rate, thus highly variable in space and time (Ciach, 2003; Habib et al., 2001).

Kriging interpolation methods provide rainfall estimations with associated variance, offering a starting platform for uncertainty estimation. The nugget effect in the variogram model is a way to represent the rain gauge measurement uncertainty (Clark, 2010), but it assumes homoscedasticity, therefore the uncertainty is modelled uniform in space. If a time invariant variogram is used, the modelled uncertainty may be modelled uniform in time as well (Cressie, 1993). De Marsily (1986), proposes an approach named Kriging for Uncertain Data (KUD) to use a different nugget for different measurement points. The formulation is refined by Mazzetti and Todini (2009). However, we are not aware of KUD applications to rain gauge interpolation and of its performance evaluation.

This work studies the application of KUD to the interpolation of rain gauge measurements. The formulation in this work is a simplification of Mazzetti and Todini (2009). Since the application to real case studies does not allow to vary the factors affecting the KUD performance, here synthetic tests are used. The application of KUD is compared to standard ordinary kriging (OK) in three synthetic tests to investigate the effect of different factors on the KUD performance.

2. METHODOLOGY

2.1 Kriging for Uncertain Data (KUD)

In ordinary kriging, the interpolation at each point x_0 is a weighted average:

$$\hat{R}(x_0) = \sum_{\alpha=1}^{n} w_\alpha \cdot R(x_\alpha) \tag{1}$$

where $\hat{R}(x_0)$ is the interpolated rainfall in x_0, $R(x_\alpha)$ are the rain gauge measurements at locations x_α, n is the number of rain gauges, and w_α are kriging weights, estimated through the kriging system:

$$\begin{cases} \sum_{\alpha=1}^{n} w_\alpha(x_0) = 1 \\ \sum_{\alpha=1}^{n} w_\alpha(x_0) \cdot C(x_\beta - x_\alpha) + \mu = C(x_\beta - x_0) \quad \beta = 1, \dots, n \end{cases} \tag{2}$$

where x_α and x_β are generic rain gauge locations and μ is the Lagrange parameter (Cressie, 1993). $C(d)$ is the covariance function that describe the covariance between measurements at two different locations, as function of their distance d.

Equation 2 can be written in matricial form:

$$\boldsymbol{W} = \boldsymbol{C}^{-1} \cdot \boldsymbol{D} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix} \tag{3}$$

where $\boldsymbol{W}$ is a vector containing the kriging weights and the Lagrange parameter, $\boldsymbol{C}$ is the covariance matrix, where elements C_{ij} are the short notation for $C(x_i - x_j)$, and $\boldsymbol{D}$ is a vector containing the elements C_{i0} , i.e. the covariance function applied to the distances between each measurement point x_i and the prediction point x_0.

The covariance function used in this work has an exponential form:

$$C(d) = \begin{cases} c + c_0 & for\ d = 0 \\ c - c\left(1 - exp\left(-\frac{3d}{r}\right)\right) & for\ d > 0 \end{cases} \tag{4}$$

where c is the sill, r is the effective range, and c_0 is the nugget. The nugget effect can be interpreted as measurement uncertainty (Clark, 2010) and using this formulation it appears only at distance *d=0*.

The conventional modelling of measurement uncertainty is done assuming a spatially and often temporally invariant nugget effect in the definition of the covariance function, which is then applied to all the elements of Equation 3. The KUD approach instead assumes $c_0 = 0$ in the covariance function definition; then it modifies the diagonal of the covariance matrix $\boldsymbol{C}$ at each time step, so that each element C_{ii} is substituted with:

$$C_{ii} = c + c_{0i} \tag{5}$$

where c_{0i} is the nugget corresponding to the uncertainty associated to the i^{th} rain gauge, i.e. the estimated measurement variance. Applying the operation at each time step, the measurement uncertainty is modelled variant in space and in time.

2.2 Synthetic tests

The performance of KUD is expected to be affected by many factors. In particular the performance is evaluated considering three factors: 1) the spatial variability of the precipitation field, represented by the range of the covariance function; 2) the density of measuring points; 3) the accuracy of measurement points. To investigate the impact of these factors on the KUD performance three synthetic experiments are carried out. The experiments are designed to reproduce a possible case study in which a certain number of accurate rain gauges are complemented with a certain number of less accurate rain gauges.

2.2.1 Setup for all the experiments

Synthetic Gaussian fields θ, are generated with given mean m = 1 and variogram sill s = 0.1 using unconditional Gaussian simulations on a 106 by 106 pixels grid. Although the experiment is unitless, pixels can be representative of a square kilometre, and the field values to rainfall intensity in [mm/h]. Each realisation of θ is sampled by more or less accurate rain gauge simulators that sample the field θ obtaining measurement values that contain errors. The measurement value is drawn from a Gaussian distribution with mean equal to the true θ value; accurate simulators use a standard deviation equal to 5% of the true θ value, while less accurate simulators use a standard deviation equal to 20% of the true θ value. For KUD, the nugget c_0, which is a measurement of variance, is the square of the error standard deviation. This model is realistic for different rain gauges types and reproduce the proportionality of the errors to the rainfall intensity, as observed in real rain gauges (Ciach, 2003; Habib et al., 2001). The rain gauge position changes randomly for every realisation of θ.

Once the fields are measured by the rain gauge simulators, both KUD and standard OK are performed and compared, as explained in section 2.3.

The three synthetic experiments are designed as follow:

2.2.2 First synthetic experiment: field spatial variability

The first experiment aims at assessing what is the impact of the rainfall spatial variability on the KUD performance. The range parameter is varied, using values of 10, 30, 50, 80 and 100 pixels. For each value, 500 realisations of θ are made and sampled. 10 accurate sampling points and 10 less accurate ones are used and their position is randomly sampled from a uniform distribution in the domain for each realisation.

2.2.3 Second synthetic experiment: rain gauge density

The second experiment assesses the impact of rain gauge density, for a given spatial variability of rainfall. The test compares the use of 4, 10, 20, 40, and 80 rain gauges in the domain, divided evenly between highly accurate measuring points and less accurate ones. 500 realisations are made for each of the values, changing the position of the measuring points randomly at each realisation.

2.2.4 Third synthetic experiment: sampling point accuracy

The third experiment looks at the impact of the relative number of accurate and less accurate rain gauges. 30 sampling points are used for all realisations, divided unevenly between accurate and less accurate ones. In particular the accurate/less accurate combinations are (5/25, 10/20, 15/15, 20/10, 25/5).

Table 1. Overview of the three experimental setups

	Experiment 1	Experiment 2	Experiment 3
Tested variable	Range	Total number of rain gauges (RG)	Ratio between accurate and less accurate rain gauges
Tested values	Range = (10, 30, 50, 80, 100)	Number of RG = (4, 10, 20, 40, 80)	Ratio = (5/25, 10/20, 15/15, 20/10, 25/5)
Number of realisation	500 for each range value	500 for each tested number of rain gauges	500 for each tested rain gauge ratio
Number of accurate rain gauges	10	2, 5, 10, 20, 40	5, 10, 15, 20, 25
Number of less accurate rain gauges	10	2, 5, 10, 20, 40	25, 20, 15, 10, 5
Range value [pixels]	10, 30, 50, 80, 100	50	50
Sill [-]	0.1	0.1	0.1
Mean [-]	1	1	1

2.3 KUD performance evaluation

In all the experiments, the performance of KUD is compared in the same way. Each time a realisation is done and sampled, the synthetic measurements are interpolated both with KUD and with standard OK, obtaining the fields θ_{KUD} and θ_{OK} respectively. An evaluation score β is then calculated as follow:

$$\beta = \sqrt{E\{(\theta - \theta_{OK})^2\}} - \sqrt{E\{(\theta - \theta_{OKUD})^2\}} \tag{5}$$

where the spatial mean $E\{\}$ is calculated throughout the pixels. β is designed to be positive when KUD performs better than OK and negative when OK performs better than KUD. A β value is calculated for each realisation, therefore 500 β values are calculated for each tested parameter.

Rainfall is not Gaussian in reality and the use of Gaussian fields may improve kriging performances. Nevertheless, both KUD and OK are affected in the same way and the use of a differential indicator assures that this assumption does not alter the results.

3. RESULTS

For each experiment, the average β, together with the 10% and 90% quantiles obtained out of the 500 realisations are presented. Additionally, the number of realisations with positive or negative β, indicating the number of times KUD outperforms or underperforms OK, are reported.

The results are summarised in Figure 1.

4. DISCUSSION

The results reported in Figure 1 show some interesting outcomes. The first thing to notice is that in none of the analysed configurations KUD performs consistently better or consistently worse than OK throughout the 500 realisations. This suggests that each combination of rain gauges configuration, error sampling, and rainfall field behaves in a different way. The reason for this is that KUD tends to smooth the field, giving less weight to the less accurate measurements. This has a positive effect if indeed the introduced error overestimates a peak or underestimate a low. Instead, the effect is counterproductive when it results in smoothing peaks and lows that were correctly estimated. When a sufficient number of measuring point is available, in relation to the field spatial variability, KUD is able to better estimate the direction of the errors and correct them, thanks to the comparison to adjacent measuring points; if instead the rain gauges are too sparse, in relation to the spatial variability of the field, KUD may result in a loss of spatial information.

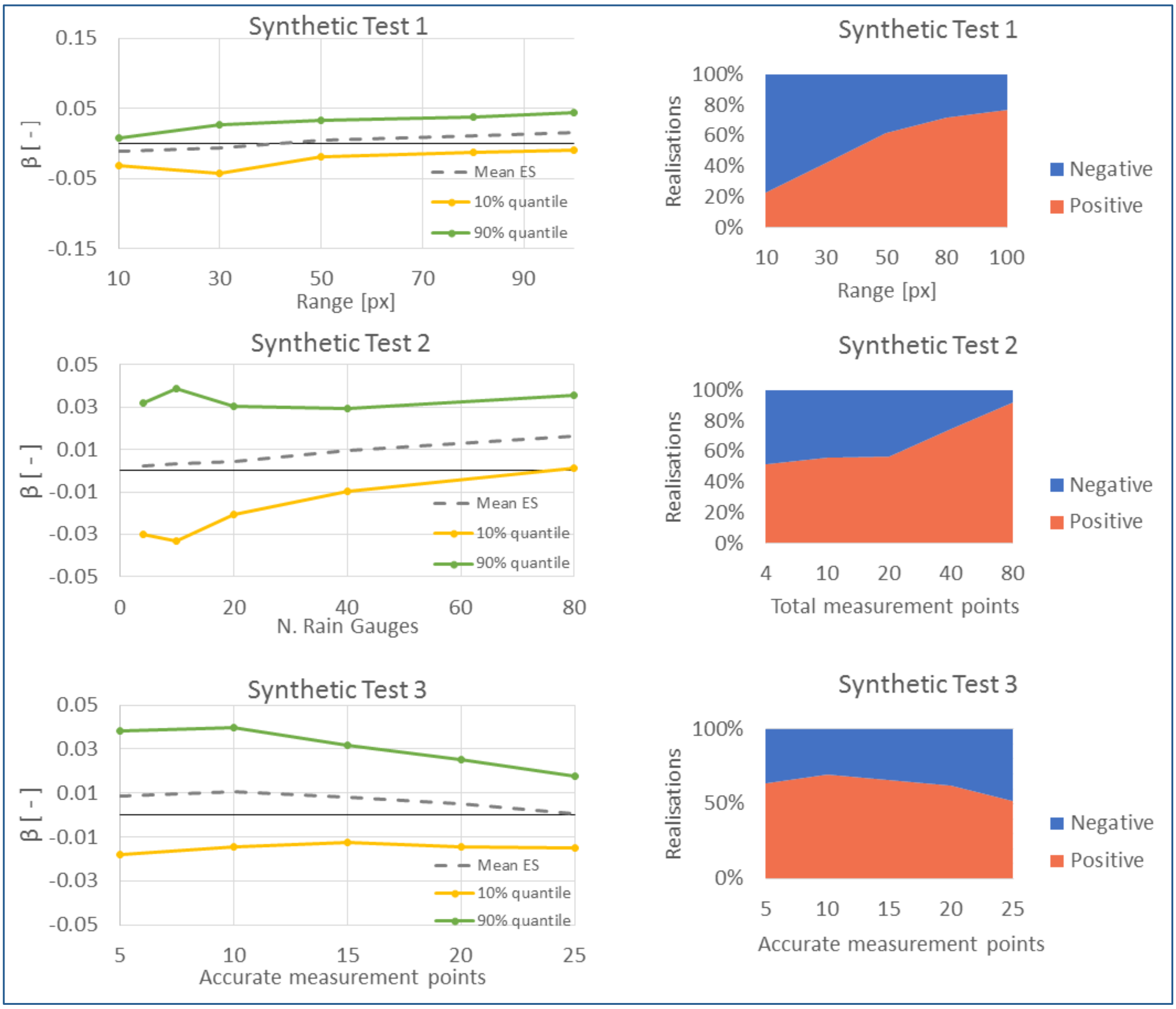

Figure 1. For each experiment, two plots are reported; the ones on the left show the mean, and the 10% and 90% quantiles of the β values. The plots on the right show how many times KUD outperforms OK (positive β) and how many times instead OK outperforms KUD (negative β). In the bottom panels, where the results of the third experiment are reported, the "x" axis reports only the number of accurate rain gauges, and it must be kept in mind that a corresponding number of less accurate rain gauges is present, such that they sum up to 30.

The results of the first experiment, reported in the top plots of Figure 1, show that KUD performs better when a high range value is used. In such a situation the fields θ are smoother and a lower number of rain gauges is sufficient. This experiment suggests that KUD could perform better in interpolating rain gauges when the studied rainfall event is stratiform, and worse when the field is convective.

The results of the second experiment, reported in the central panels of Figure 1, show that KUD performs better with a higher rain gauge density. The reason for this behaviour is again that a higher number of rain gauges allows us to have a correct spatial sampling of the field, identify highs and lows more precisely and balancing out the measurement errors. Another interesting feature that can be observed in the second experiment is that the variability of outcomes, represented by the difference between the 10% and 90% quantile, becomes lower for higher rain gauge densities. This means that the KUD and OK results tend to converge to a better estimation of the field.

Finally, the third experiment suggests that KUD performs better than OK when the rain gauges have a more variable accuracy, when the difference between the number of accurate and less accurate rain gauges is lower. The KUD method is able to give a higher weight to the good measurements, and a lower weight to the less accurate ones. When the sampling points have a more uniform accuracy, this advantage is less important, and the two methods perform more similarly, i.e. the β mean is closer to zero. Again the distribution is narrower for a higher number of accurate rain gauges indicates that more accurate measurements result in a convergence of KUD and OK results.

5. CONCLUSIONS

The work presented in this paper investigates the applicability of Kriging for Uncertain Data (KUD) to rain gauge interpolation, in order to take into account the measurement errors and their variability in space and time. The investigation is conducted by means of three synthetic experiments, in which random Gaussian fields correlated in spaces are generated to represent rainfall, and are sampled with rain gauge simulators, able to reproduce a realistic error structure in the measurements. The interpolation performed with KUD is compared to the interpolation done with the standard Ordinary Kriging (OK) algorithm. The comparison is done for multiple realisations of the Gaussian field, and for multiple rain gauge positions, varying factors that influence the KUD performance. In particular, the spatial variability of the field is studied in the first experiment, the rain gauge network density in the second, and the relative accuracy of the rain gauges in the third.

The presented work is not a complete and comprehensive overview of all the possible factors that can influence the KUD performance in rain gauge interpolation, the use of Gaussian fields is an approximation, and temporal correlation is not considered. However, the experiment results give some insight in the KUD properties, its strengths and its weaknesses. The following conclusions are drawn:

1. The use of KUD is advisable only if a sufficient rain gauge density is available, in relation to the rainfall field spatial variability. If the rain gauge network is too sparse, or the rainfall field decorrelation distance too short, the KUD performance drops.
2. KUD is particularly recommended when the used rain gauges have a variable accuracy, because it is able to give a higher weight to more accurate rain gauges.
3. The more accurate and dense the measurements are, the more KUD and OK estimates tend to converge, and KUD seems to outperform OK.
4. The strength and the limitation of KUD is the fact that it tends to smooth the field, in particular in the surroundings of less accurate measurements. If a sufficient sampling is performed, the smoothing primarily targets the deviation due to errors; if instead the spatial information is too sparse, the smoothing may affect real rainfall field features, like peaks and lows, and the quality of the estimation drops.

Given the above mentioned conclusions, the use of KUD is in general recommended, because it tends to outperform the standard OK algorithm in most of the situations, but there are conditions in which it is counterproductive. In particular, it is not advisable for sparse rain gauge networks, coupled with convective rainfall events.

Although the experiment is tailored to rain gauge rainfall interpolation, the conclusions can be interesting for the application of KUD to other fields as well.

ACKNOWLEDGEMENTS

This work was carried out in the framework of the Marie Skłodowska Curie Initial Training Network QUICS. The QUICS project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000. We would like to thank Prof. Gerard B.M. Heuvelink for the geostatistical feedback.

REFERENCES

AghaKouchak, A., Habib, E., Bárdossy, A., 2010. Modeling Radar Rainfall Estimation Uncertainties: Random Error Model. J. Hydrol. Eng. 15, 265–274. doi:10.1061/(ASCE)HE.1943-5584.0000185

Ciach, G.J., 2003. Local Random Errors in Tipping-Bucket Rain Gauge Measurements. J. Atmos. Ocean. Technol. 20, 752–759.

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

Clark, I., 2010. Statistics or geostatistics? Sampling error or nugget effect? J. South. African Inst. Min. Metall. 110, 307–312.

Cressie, N.A.C., 1993. Statistics for Spatial Data.

Dai, Q., Han, D., Rico-Ramirez, M., Srivastava, P.K., 2014. Multivariate distributed ensemble generator: A new scheme for ensemble radar precipitation estimation over temperate maritime climate. J. Hydrol. 511, 17–27. doi:10.1016/j.jhydrol.2014.01.016

De Marsily, G., 1986. Quantitative Hydrogeology.

Germann, U., Berenguer, M., Sempere-Torres, D., Zappa, M., 2009. REAL – Ensemble radar precipitation estimation for hydrology in mountainous region. Q. J. R. Meteorol. Soc 135, 445–456. doi:10.1002/qj.375

Habib, E., Ciach, G.J., Krajewski, W.F., 2004. A method for filtering out raingauge representativeness errors from the verification distributions of radar and raingauge rainfall. Adv. Water Resour. 27, 967–980. doi:10.1016/j.advwatres.2004.08.003

Habib, E., Krajewski, W.F., Kruger, A., 2001. Sampling Errors of Tipping-Bucket Rain Gauge Measurements. J. Hydrol. Eng. 6, 159–166. doi:10.1061/(ASCE)1084-0699(2001)6:2(159)

Habib, E.H., Meselhe, E. a., Aduvala, A. V., 2008. 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)

Mazzetti, C., Todini, E., 2009. Combining Weather Radar and Raingauge Data for Hydrologic Applications, in: Flood Risk Management: Research and Practice. Taylor & Francis Group, London.

Molini, A., Lanza, L.G., La Barbera, P., 2005. 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

Nešpor, V., Sevruk, B., 1999. Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. J. Atmos. Ocean. Technol. 16, 450–464. doi:10.1175/1520-0426(1999)016<0450:EOWIEO>2.0.CO;2

Peleg, N., Ben-Asher, M., Morin, E., 2013. Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network. Hydrol. Earth Syst. Sci. 17, 2195–2208. doi:10.5194/hess-17-2195-2013

Rico-Ramirez, M. A., Liguori, S., Schellart, A. N. A., 2015. Quantifying radar-rainfall uncertainties in urban drainage flow modelling. J. Hydrol. 528, 17–28. doi:10.1016/j.jhydrol.2015.05.057

Schilling, W., 1991. Rainfall data for urban hydrology : what do we need? Atmos. Res. 27, 5–21.

Sevruk, B., 1996. Adjustment of tipping-bucket precipitation gauge measurements. Atmos. Res. 42, 237–246. doi:10.1016/0169-8095(95)00066-6

Villarini, G., Krajewski, W.F., 2009. Empirically based modelling of radar-rainfall uncertainties for a C-band radar at different time-scales. Q. J. R. Meteorol. Soc. 1438, 1424–1438. doi:10.1002/qj

Villarini, G., Mandapaka, P. V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. J. Geophys. Res. 113, D11102. doi:10.1029/2007JD009214