



# Correcting bias in radar Z–R relationships due to uncertainty in point rain gauge networks



Mohammad Mahadi Hasan<sup>a</sup>, Ashish Sharma<sup>a,\*</sup>, Fiona Johnson<sup>a</sup>, Gregoire Mariethoz<sup>a,c,d</sup>, Alan Seed<sup>b</sup>

<sup>a</sup> School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia

<sup>b</sup> Weather and Environmental Prediction Group, Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, VIC 3001, Australia

<sup>c</sup> ETH Zürich, Department of Earth Sciences, Zurich, Switzerland

<sup>d</sup> University of Lausanne, Institute of Earth Surface Dynamics, Switzerland

## ARTICLE INFO

### Article history:

Received 12 April 2014

Received in revised form 17 July 2014

Accepted 20 September 2014

Available online 28 September 2014

This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Emmanouil N. Anagnostou, Associate Editor

### Keywords:

Radar Z–R relationship

Sub-grid variability

Point gauge uncertainty

Spatial variability

Error model

SIMEX

## SUMMARY

One of the key challenges in hydrology is to accurately measure and predict the spatial and temporal distribution of rainfall. Rain gauges measuring at point locations are often considered as the “ground truth” for grid based radar rainfall calibration. Usually, no consideration is given to the uncertainty in the measurement that varies depending on the number of rain gauges that fall within each grid cell. If this uncertainty in the rain gauge network measurements is ignored, the Z–R relationship used to convert reflectivity (Z) to rainfall (R) will be biased. We investigate the effects of point gauge rainfall uncertainty on parameter bias in the Z–R relationship. An error model is developed to compute point gauge rainfall uncertainty at the radar grid resolution. This error model has two components: (1) the error in the gauge measurement itself, and (2) the error introduced by the gauge not capturing the spatial variability within a radar pixel. The Simulation Extrapolation method (SIMEX) is used to determine the extent of parameter bias present in the rainfall–reflectivity relationship as a result of this uncertainty. When considering the point gauge rainfall uncertainty a 4% decrease in the average radar rainfall estimates is found.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Rainfall is one of the most important inputs to hydrological analysis and modelling. Any error in this input will propagate through the model and will introduce uncertainty in subsequent predictions (Hossain et al., 2004; Morin et al., 2005; Pessoa et al., 1993; Sharif et al., 2002; Vieux and Bedient, 1998; Vivoni et al., 2007; Zhu et al., 2013). One source of uncertainty and possible errors is the temporal and spatial variability of rainfall (AghaKouchak et al., 2010; Faurès et al., 1995; Goodrich et al., 1995; Shah et al., 1996). Therefore, accurately measuring and predicting the spatial and temporal distribution of rainfall is an important challenge in hydrology.

Remote sensing provides an increasingly important source for spatial estimation of precipitation. Using measurements of reflectivity, weather radars can produce estimates of precipitation over large geographic areas, and therefore provide information about rainfall patterns at high temporal and spatial resolution. As a

result, the potential applications of weather radars for hydrologic modelling have been the subject of extensive research (Bonnifant et al., 2009; Cole and Moore, 2008, 2009; Collier, 2009; Keblouti et al., 2013; Looper and Vieux, 2012; Viviroli et al., 2009).

Traditionally, radar reflectivities  $Z$  ( $\text{mm}^6/\text{m}^3$ ) are converted to ground rainfall rates  $R$  ( $\text{mm/h}$ ) using a power–law relationship ( $Z = AR^b$ ) known as the Z–R relationship. The relationship between radar reflectivity and rainfall rate depends on the nature of rainfall, and in particular the drop size distribution (DSD) (Chumchuan et al., 2006a, 2008; Islam et al., 2012; Uijlenhoet and Pomeroy, 2001). One way to calibrate the Z–R relationship is through the use of disdrometers which can explicitly measure the DSD and therefore the relationship with the radar reflectivity (Ochou et al., 2011; Verrier et al., 2013). However in many parts of the world, including Australia, gauge based calibration of the Z–R relationships is routinely used (Brangi et al., 2011; Rendon et al., 2013). Furthermore, in developing countries where disdrometer data is not available and even sub-daily rainfall measurements are infrequent, daily rainfall gauges are the sole source of information on ground rainfall estimates (Mapiam et al., 2009). In these cases gauge densities are also likely to be very low compared to developed countries such as the United States of America and Europe.

\* Corresponding author at: School of Civil and Environmental Engineering, University of New South Wales, Kensington, NSW 2052, Australia. Tel.: +61 2 9385 5768; fax: +61 2 9385 6139.

E-mail address: [a.sharma@unsw.edu.au](mailto:a.sharma@unsw.edu.au) (A. Sharma).

Thus errors in the recorded gauge rainfall can bias the Z–R relationship, resulting in errors in the radar rainfall estimates. A biased Z–R relationship may be due to an inadequate consideration of the spatial subgrid scale gauge rainfall variability and its representation through the handful of gauges that are available for use (Ciaich et al., 2007; Jordan et al., 2003) and therefore methods that take account of gauge uncertainty can be very useful.

The performance of the radar rainfall estimates is evaluated by comparing them with point gauge measurement at the ground (Anagnostou et al., 1999; Habib and Krajewski, 2002). These point gauge measurements are therefore considered as the “ground truth” (Lebel and Amani, 1999; Wolff et al., 2005). However, there is a spatial discrepancy between the two data sources, since radar rainfall estimates are provided as spatial averages with a resolution 1–4 km<sup>2</sup> (Krajewski and Smith, 2002). In contrast, whilst rain gauges measure precipitation at fixed point locations, these are often too sparse to properly represent the spatial variability and areal structure of rainfall (Morrissey et al., 1995; Rodríguez-Iturbe and Mejía, 1974). Furthermore, the uncertainty associated with spatial averages over a single radar grid cell is considerable, and is a function of the number of gauges that are within the cell being considered. In addition, rain gauge measurements can contain a variety of errors including wind effects, evaporation and mechanical tipping bucket errors (Groisman and Legates, 1994). An important question is whether the parameters estimated for the Z–R relationship would remain the same if consideration was given to the nature of the errors associated with spatial averaging of the gauge rainfall over the radar grid scale. We argue that if these errors are significant (and vary with space as they will if the rainfall and the gauge density are not uniform across the network), the estimated parameters will have a bias with respect to the true parameters that ought to be used.

This research aims to address whether the A parameter in the radar Z–R relationship needs to change to account for uncertainty in the point gauge rainfall network. It has been reported that the parameter A carries most of the variability in the Z–R relationship, whereas the uncertainty in *b* can be seen as second-order (Chumchean et al., 2003, 2006b; Steiner et al., 1999). Marshall and Palmer (1948) proposed the power law relation,  $Z = 200 R^{1.6}$  with specified values for A equal to 200 and *b* equal to 1.6. Since then, several studies have been conducted to find appropriate A and *b* parameter values in different settings. Given the need for consistent estimates of rainfall, for operational purposes the Australian Bureau of Meteorology (BOM) specifies a fixed value for *b* equal to 1.53 for most of the weather radars that form its rainfall measuring network. Fixing *b* for operational networks has also previously been used by other researchers (Steiner and Smith, 2004; Verrier et al., 2013). We use the Simulation Extrapolation method (SIMEX) (Cook and Stefanski, 1994) to investigate the significance of point gauge uncertainty on the Z–R relationship parameter A. The SIMEX method involves developing a relationship between the gauge uncertainty distribution and input gauge rainfall, which can then be used to assess the bias in the parameters of the Z–R relationship.

The paper is organised in seven sections. Section 2 outlines the logic behind the SIMEX approach. Section 3 describes the radar and gauge data used in this study, whilst Section 4 presents the development of an error model for the lowest radar pixel resolution (1 km<sup>2</sup>). The application of SIMEX method on radar Z–R relationship is presented in Section 5 followed by results and discussion in Section 6. The main findings are summarised in Section 7.

## 2. Simulation Extrapolation (SIMEX)

SIMEX is a method for parameter estimation that attempts to ascertain model parameters taking into account the error

distribution associated with each predictor variable. It estimates parameter values that should have resulted if the covariates were error-free. The general idea behind the method is that if the error in the predictors causes bias in the parameter estimates, then adding more error should cause the parameter estimates to become even more biased (Benoit et al., 2009). A relationship between the bias and the added error can be developed and this can be extrapolated back to the case where there is no error. This process is shown in Fig. 1 where the naive estimate is the estimation of some parameter  $\beta$  using the recorded data, including its inherent errors. Estimates to the right of the naive estimate represent the cases where more error has been added to give even more biased estimates of  $\beta^*$ . The trend of the bias versus the error is then extrapolated back to the case of no error, which is known as the SIMEX estimate. The error variance is multiplied by a factor  $\lambda$  termed the Multiple of Error Variance. The difference between two consecutive  $\lambda$ 's is a variance inflation factor.

Consider a simple linear regression  $Y = \beta X$ . Suppose that, instead of observing the covariate *X*, we observe an erroneous covariate *W* with an error distribution *U* having a zero mean and variance  $\sigma_u^2$  (or,  $W = X + U$ ). The estimate for the coefficient  $\beta$  (referred to as  $\beta_{\text{naive}}$  and termed the naive estimate) is likely to be different from the estimate that would result if *W* were error free (or, equal to *X*). In SIMEX, more erroneous realisations of *W* with variance of  $\lambda \times \sigma_u^2$  are drawn (these being termed  $W^*$ ), and the resulting parameters (with an expected value of  $\beta^*$ , (shown with a circle in Fig. 1) are estimated. Finally, a relationship is developed with multiple of error variances ( $\lambda$ ) and the expected value  $\beta^*$  corresponding to each  $\lambda$ . The relationship is then extrapolated to the condition where no additive error is present, i.e.  $\lambda = -1$ , giving an unbiased SIMEX estimate of the target parameter which is denoted as  $\beta_{\text{SIMEX}}$ .

SIMEX has been popular in statistical research due to its ability to produce an efficient parameter estimate for linear, non-linear, logistic and non-parametric models (Carroll et al., 1996, 1999; Cook and Stefanski, 1994; Holcomb, 1999; Marcus and Elias, 1998; Staudenmayer and Ruppert, 2004). Chowdhury and Sharma (2008) applied the SIMEX method in hydrology to determine the bias in the Sacramento Rainfall Runoff Model's key storage parameters using synthetic data. They found that compared to naive estimates, SIMEX estimates were closer to the true values. SIMEX has also been used to reduce biases in parameter estimate of future droughts, which are due to uncertainties and errors in General Circulation Model (GCM) projections of rainfall (Woldemeskel et al., 2012). To date, SIMEX has not been applied in radar–rainfall calibration and this paper aims to test the utility of the method in this context.

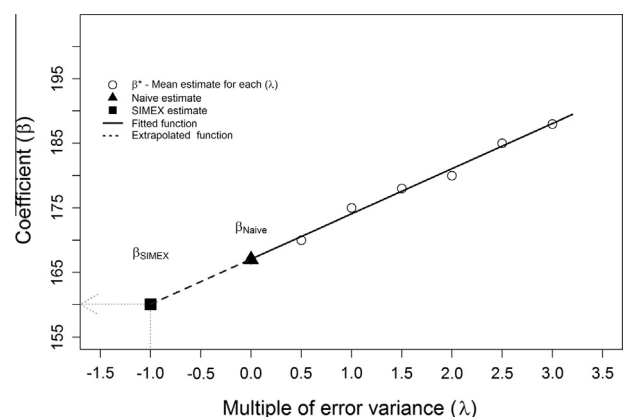


Fig. 1. Synthetic example of the SIMEX approach.

It is important to note that the SIMEX method requires the prior knowledge of the error distribution associated with the intended model parameter. Chowdhury and Sharma (2007) illustrates the SIMEX method using a synthetic example with additive error in a linear regression model. However, Tian et al. (2013) report that the multiplicative error models are more suitable than additive error models when dealing with precipitation measurements. Therefore in this paper, we consider a multiplicative point gauge rainfall measurement error. The following section discusses the data used in this study and this is followed by the development of the multiplicative error model.

### 3. Data

In this research, the reflectivity data were obtained from the Australia Bureau of Meteorology for the Terrey Hills radar (Sydney, Australia) during the period from November 2009 to December 2011. The Terrey Hills radar is an S-band Doppler radar with 6 min temporal resolution and 1 km spatial resolution. The radar covers a region of 256 km by 256 km extent, with bandwidth of 1° and wavelength of 10.7 cm. The climatological freezing levels in Sydney are about 2.5 km (Chumchean et al., 2003). Therefore the 1.5 km Constant Altitude Plan Position Indicator (CAPPI) reflectivity data were used to avoid bright band effects. The noise and hail effect was nullified by only considering reflectivity ranges from 15 dBz to 53 dBz (Chumchean et al., 2004, 2006a).

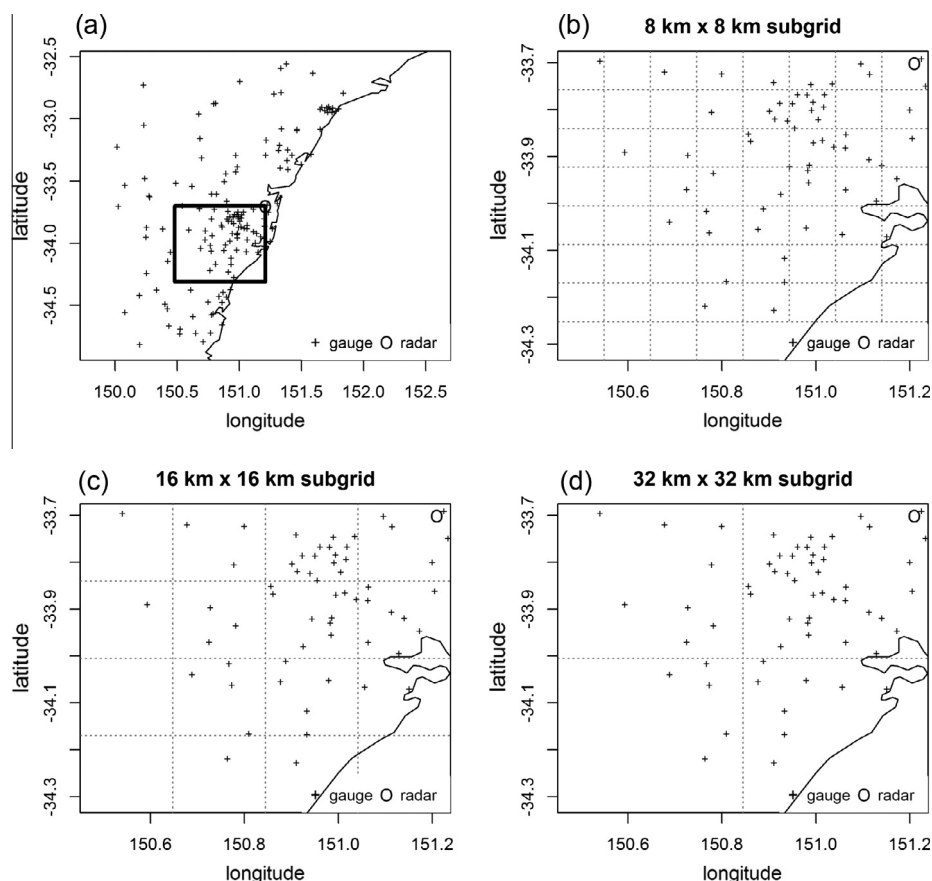
The rain gauge network used for the calibration of the radar includes 150 tipping bucket gauges located within 128 km of the radar as shown in Fig. 2a. The Australian Bureau of Meteorology operates and maintains these stations, with the tipping bucket size

being either 0.2 or 0.5 mm. In this study, a threshold on the minimum rainfall intensity of 1 mm/h was used to avoid quantification error at low rainfall intensity (Chumchean et al., 2004, 2006b).

For this research, storms are defined as discrete meteorological events through visual inspection of the time series of radar images. The start of each storm was chosen such that there was significant rainfall over a spatially coherent area (i.e. scattered small systems were not used in the analysis). The end of each storm was found when there was no longer any significant rainfall occurring in the region. Thus the duration of each storm is allowed to vary and the final dataset is compiled from all hours belonging to any storm. The selected 107 storm events lead to 1259 h of rainfall and contain a wide range of storm durations varying between 1 and 79 h, although 59% of the storm event duration ranges are less than 10 h. It is therefore concluded that the data set contains a good mix of convective and stratiform events. The radar reflectivity and gauge rainfall data are extracted with the same event start and end time, assuming no time lag between radar and gauge rainfall. The radar rainfall software MAPVIEW (Chumchean et al., 2006b; Seed and Jordan, 2002) was used to import the reflectivity files and to accumulate the rainfall.

### 4. Error model

The rain gauge data cannot be realistically assumed as error-free when calibrating the Z–R relationship. The SIMEX technique allows to account for the various sources of errors in the rainfall totals and to remove any corresponding bias from our estimate of the parameter A. As outlined above, there are two main sources of error in the rain gauge observations with respect to the



**Fig. 2.** (a) Location of point gauges around Sydney Terrey Hills Radar. The highlighted area shows enlarged view of different number of rain gauges for subgrid sizes of (b) 8 km × 8 km (c) 16 km × 16 km and (d) 32 km × 32 km respectively.

radar–rainfall relationship. The first one is due to recording errors in the tipping bucket gauge. The second one is the error introduced by not properly sampling the spatial variability of rainfall at the scale of the radar pixels. These two error components can be estimated separately and the total error calculated by combining them.

#### 4.1. Measurement error model

As discussed previously, rain gauge measurements contain a variety of errors including wind effects, evaporation and mechanical error (Groisman and Legates, 1994) as well as the spatial sampling error (Morrissey et al., 1995). Using a network of rain gauges, Habib et al. (2001) investigated tipping bucket sampling error with the emphasis on gauge ability to present temporal variability for the case of small scale rainfall. Their research showed sampling interval and bucket size are the dominant factors that controls gauge performance and its associated error. In contrast, Ciach (2003) found that gauge measurement error are highly reliant on rainfall intensity and timescale. However, both studies found that as the rainfall rate and/or accumulation time increases, the gauge measurement error generally decreases.

For the error in tipping bucket gauge measurements, the error model from Ciach (2003) is adopted. In this model, the gauge measurement uncertainty is a function of the rainfall intensity and the time period over which the rainfall is being aggregated. Shorter time scales were found by Ciach (2003) to have larger relative errors than longer time aggregations where the errors in the tipping buckets tended to cancel out. The tipping bucket error model is presented in Eq. (1) and based on Ciach (2003) where the error is defined as the standard error ( $\sigma_g$ ) of the gauge measurement.

$$\sigma_g(T, P_T) = e_0(T) + \frac{R_0(T)}{R_i} \quad (1)$$

where  $e_0$  and  $R_0$  are the model coefficients at time scale  $T$ , which for the 1 h rainfall totals take the value of 0 and 0.16 respectively (Ciach (2003)). The rainfall rate is denoted by  $R_i$  in units of mm/h.

#### 4.2. Spatial disaggregation error model

Zawadzki (1973) provided an analytical description of the errors and the fluctuations of areal rainfall considering uniformly distributed rain gauges. Rodríguez-Iturbe and Mejía (1974) considered the spatial distribution of rain gauge and developed a method to estimate the standard error of areal mean for stratified and random rain gauge network configurations. Morrissey et al. (1995) further developed the model for any spatial configuration of measuring sites by considering the correlation between a number of gauges as well as their spatial distribution. Krajewski et al. (1998) attempted to estimate the radar rainfall uncertainty at sub-pixel scales. Later, the radar rainfall estimation error was differentiated from rain-gauge sampling error by introducing an error separation technique. Other studies (Krajewski et al., 2000; Villarini et al., 2008) used a variance reduction factor (VRF) to compute the accuracy of the pixel-average rainfall. Jensen and Pedersen (2005) studied rainfall variation for single radar pixel (500 by 500 m) of nine gauges and their results showed that within a pixel, the gauge rainfall varies up to 100% among neighbouring stations. The studies of Villarini et al. (2008) used the VRF as an indicator of uncertainties associated with the average of point measurements with the approximation of the true areal value. They suggest that VRF depends on the network density, its configuration and the spatial correlation of the sampled process. Villarini and Krajewski (2008) pointed out that the spatial sampling uncertainty should be considered if a single gauge is used to approximate the areal estimates. The authors proposed that a

minimum number of gauges are required to represent areal rainfall, for example they found that 25 gauges are required to estimate the true areal rainfall with 20% accuracy for a basin size of 12 km by 16 km. The radar rainfall estimation method as currently used is strictly valid only when the rainfall measured by a single gauge represents that measured in a radar pixel, as one could then argue that the spatial sampling uncertainty is relatively uniform. However, there are clearly significant scale differences between these two measurements.

Despite the extensive previous research on rainfall spatial variability, we did not find a simple error model that could account for the error from incorrectly sampling the sub-grid variability at the radar pixel scale. We have therefore developed such a model using data for the study region we have focussed on. Our hypothesis is that the uncertainty in the rainfall spatial variability should be a function of rainfall rate, gauge density and the size of the area that the point rainfall is being assumed to represent. The desirable characteristics of the error model are:

- The uncertainty in the gauge rainfall should approach zero as the area over which spatial variability becomes smaller.
- As the number of gauges in the area increases, the standard deviation should represent the true variability across the studied area.
- As the rainfall intensity increases, the standard deviation should also increase.

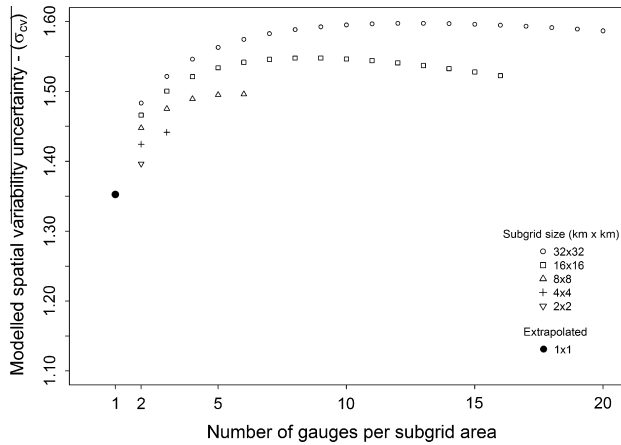
For our study area, a relationship between the number of gauges and the standard deviation of the rainfall across a finite analysis area was formulated. Because of the relatively low density of gauges across a large area, most radar pixels only contain at most a single gauge and therefore a relationship between the sub-grid spatial variability and the gauge rainfall could not be determined. To overcome this, increasing levels of spatial aggregation were considered so that multiple gauges were contained in each aggregated grid cell. For the consideration of spatial aggregation, the same area was divided into subgrid sizes of 2 km by 2 km, 4 km by 4 km, 8 km by 8 km, 16 km by 16 km and 32 km by 32 km. Fig. 2 shows the location of point gauges around Sydney Terrey Hills radar. The highlighted area is enlarged to illustrate the number of rain gauges per cell for different subgrid sizes.

We start by standardising the data results across all rainfall intensities. To this end, the coefficient of variation (CV) (Pedersen et al., 2010) was computed as

$$CV = \frac{\sigma}{\bar{R}} \quad (2)$$

where  $\sigma$  is the standard deviation of rainfall in each subgrid and  $\bar{R}$  is the mean rainfall for that subgrid. For each rainfall event and for all combinations of aggregation area, the numbers of gauges recording rainfall and the CV of the rainfall across those gauges were calculated for each subgrid. The final data set of area, number of gauges and CV values was then used to formulate an empirical relationship using the Eureka model exploration framework (Schmidt and Lipson, 2013). Eureka gives a number of alternate relationships depending on their associated complexity and a  $R$ -square goodness of fit. The adopted functional relationship is presented in Fig. 3 for different area aggregations from 4 km<sup>2</sup> to 1024 km<sup>2</sup>. This relationship was then extrapolated back to the case of 1 km<sup>2</sup> which is the radar pixel resolution the rainfall is derived at, and this final relationship is also shown in Fig. 3. Then, for any radar pixel, the number of gauges within the pixel is counted and the corresponding CV ascertained. When combined with the average rainfall for the gauges in the radar pixel for that time period, the CV is an estimate of the spatial uncertainty in the gauge rainfall ( $\sigma_{cv}$ ).





**Fig. 3.** Illustration of Eureka modelled spatial variability uncertainty ( $\sigma_{cv}$ ) for different subgrid sizes having different number of gauges in each grid. This model is used to extrapolate spatial variability uncertainty for a subgrid size of 1 km by 1 km.

#### 4.3. Total error model

Finally the gauge measurement error is estimated using Eq. (1), and the spatial variability uncertainty is estimated from Fig. 3 for each radar pixel of 1 km<sup>2</sup>. The total uncertainty ( $\sigma_u$ ) from the gauge rainfall for a 1 km<sup>2</sup> radar pixel is dependent on the rainfall rate as well as the gauge density and given by

$$\sigma_u = \sqrt{(\sigma_g^2 + \sigma_{cv}^2)} \quad (3)$$

By combining the error model with the recorded rainfall, new rainfall simulations can be created that are consistent with the original recorded data but account for the uncertainty due to spatial variability and gauge measurement error. A multiplicative model for the rainfall error is used:

$$R'_{t,k} = R_{t,k} \times \varepsilon_{t,k} \quad (4)$$

where  $R'$  is the simulated rainfall at time  $t$  and location  $k$  based on the observed rainfall  $R$ . The error  $\varepsilon$  is defined using a lognormal distribution with zero mean in the log space and a total variance expressed as:

$$\varepsilon_{t,k} \sim \text{LN}(0, \sigma_{u,t,k}^2) \quad (5)$$

Here,  $\sigma_u^2$  is the error variance of the point gauge rainfall uncertainty obtained from the error model and LN is the lognormal distribution, which is appropriate to use for the assumed multiplicative error structure (Tian et al., 2013).

The results from applying this error model to the measured gauge rainfalls are presented in Fig. 4, where 300 realisations are compared to the observed hourly rainfall intensities. Fig. 4a shows the random errors introduced by the gauge measurement error ( $\sigma_g$ ), whilst Fig. 4b shows the spatial variability ( $\sigma_{cv}$ ) and Fig. 4c presents the total error (i.e. based on Eq. (3)). The distribution of modelled percent uncertainties for gauge measurement uncertainty and spatial variability uncertainty are presented in Fig. 4d. It is clear that the spatial variability uncertainty dominates the total uncertainty estimate. The gauge measurement uncertainty is quite small, although proportionally is more important for the lower rainfall intensities. The average spatial variability uncertainty is around 15% as shown in Fig. 4d.

#### 5. Use of SIMEX to improve the radar Z–R relationship

This section presents the algorithm that was used to estimate the unbiased  $A$  parameter using the SIMEX method. Suppose, instead of observing the gauge rainfall  $R$ , that we actually observe

the erroneous rainfall  $W$  where,  $W = R \times \delta$  and  $\delta$  is the multiplicative error distribution given by Eq. (5) (assuming  $\varepsilon \sim \text{LN}(0, \lambda\sigma_u^2)$ ). Here,  $\sigma_u^2$  is the error variance of the point gauge rainfall uncertainty obtained from the error model and  $\lambda$  is the multiple of error variance. In the simulation step, additional independent measurement errors with variance  $\lambda\sigma_u^2$  are generated and multiplied with the original  $W$ , thereby creating data sets with successively larger measurement error variance realisations  $W^*$ . Next, sample estimates for the Z–R parameter  $A$  are obtained, and their expected value denoted  $A^*$ . It is assumed that the radar reflectivity  $Z$ , radar power law parameter  $b$  and realisation  $W^*$  are related through the relationship  $Z^b = A^b W^*$ . The value for  $A$  is estimated through a linear regression of  $Z^b$  onto  $W^*$ .

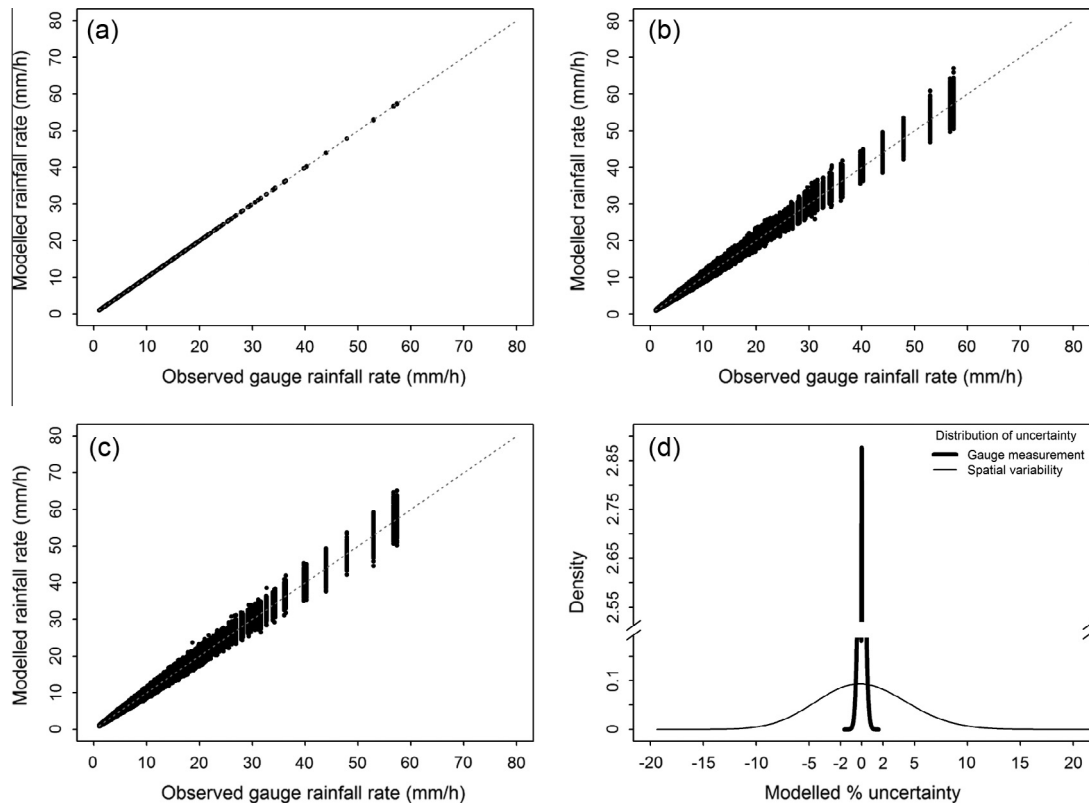
In the implementation of SIMEX adopted here, we used seven values for the error variance factors,  $\lambda \equiv [0, 0.5, 1.0, 1.5, 2.0, 2.5, \text{ and } 3.0]$ . For each  $\lambda$ , we run 300 simulations, resulting in 300 biased parameter estimates  $A$ . The expected estimate for each  $\lambda$ , denoted as  $A^*$ , is the mean of the 300 bias parameter estimates  $A$ . Finally, the SIMEX estimate is obtained from the relationship between the expected estimate of  $A^*$  with multiple of error variance  $\lambda$  and extrapolating back to a value of  $\lambda = -1$ . A quadratic function was chosen for the extrapolation.

The algorithm for unbiased  $A$  parameter estimation using SIMEX method is given below:

1. Specify the multiple of error variance factor  $\lambda$ . Here, we consider values ranging from 0 to 3, with steps every 0.5.
2. Specify the number of realisations ( $i$ ). In this case we select 300 realisations.
3. Generate random normal deviates  $\delta_{(i=1)}^*$  for  $\lambda = \lambda_1$ , where,  $\delta_{(i=1)} \sim \text{LN}(0, \lambda_{i=1}\sigma_u^2)$ .
4. Estimate the erroneous realisation,  $W_{(i=1)}^* = R \times \delta_{(i=1)}^*$ .
5. Estimate the slope  $A_{(i=1)}^*$  from  $Z_p = A_{(i=1)}^{\frac{1}{b}} W_{(i=1)}^*$  where,  $Z_p = Z^{\frac{1}{b}}$ ,  $b = 1.53$ .
6. The expected estimate  $A_{(i=1)}^*$  is the average of 300 realisations,  $A_{(i)}^* = \frac{1}{n} \sum_{i=1}^{n=300} A_{(i)}$ .
7. Repeat the steps 3–6 for  $W^* = [W_2^*, W_3^*, W_4^* \dots W_m^*]$  and  $\lambda = [\lambda_2, \lambda_3, \lambda_4 \dots \lambda_m]$ .
8. Develop an empirical relationship  $A^* = f(\lambda)$ . Here  $f(\cdot)$  is the relationship between the expected value of  $A^*$  and the multiple of error variance ( $\lambda$ ) which could be a linear or a nonlinear relationship.
9. Extrapolate to get  $A_{\text{SIMEX}} = f(\lambda = -1)$ .

To illustrate the SIMEX process, take the example of a recorded hourly rainfall of 30 mm. The assumption in this paper is that this recorded rainfall amount is uncertain. To sample this uncertainty, random noise is added to the measurement of 30 mm using a log normal distribution with a mean of zero and total error variance  $\lambda\sigma_u^2$  which is the combined gauge measurement error and spatial uncertainty multiplied by the SIMEX error variance factor. From this log-normal distribution a perturbation factor is sampled (e.g. 0.828) and used to create a new rainfall amount ( $30 \times 0.828 = 24.84$  mm). This process is repeated for all recorded rainfall amounts to create a new set of data ( $W$ ) that is consistent with the observed data and its error. From the new data set, the value for the parameter  $A$  is estimated using recorded reflectivity ( $Z$ ) and rearranging the Z–R relationship to give  $Z^{1/b} = AW$ . This process is repeated 300 times and the values of  $A$  averaged to give  $A^*$  for some value of  $\lambda$ .

The whole process described above (Steps 3–6 in the algorithm) is repeated for the other values of  $\lambda$  (from 0 to 3) with the assumption that as more error is added to the recorded rainfall amounts (i.e.  $\lambda$  increases) the value found for  $A^*$  will become more biased. The next step is to fit a relationship between the  $A^*$  values and  $\lambda$ .



**Fig. 4.** Sample estimates of modelled rainfall uncertainty for 300 realisations based on error model for (a) gauge measurement uncertainty, (b) spatial variability uncertainty and (c) total uncertainty. In panel (d) the probability density function averaged across all rainfall rates is shown for the gauge measurement uncertainty and spatial variability uncertainty.

values. Finally this can be extrapolated back to the case of  $\lambda = -1$  where there is no error in the measurements.

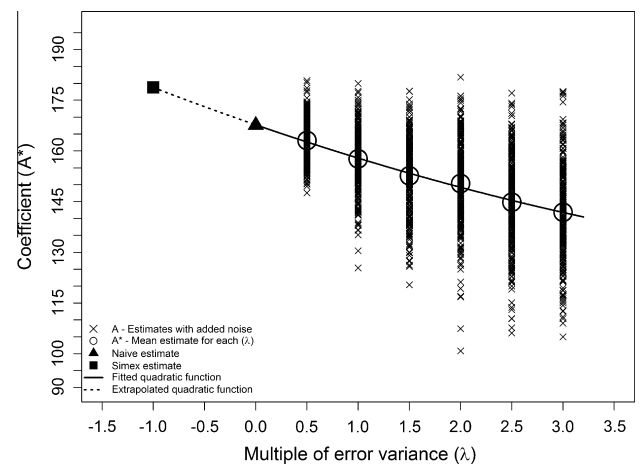
## 6. Results and discussion

### 6.1. SIMEX results

We have combined gauge measurement uncertainty ( $\sigma_g$ ) and spatial variability uncertainty ( $\sigma_{cv}$ ) to estimate point gauge uncertainty ( $\sigma_u$ ) (Fig. 4c). We have found that spatial variability uncertainty dominates (Fig. 4b) compared to the gauge measurement uncertainty (Fig. 4a). In addition, the same number of gauges in a larger area gives higher spatial variability uncertainty in the area of interest (Fig. 3).

The results of the SIMEX analysis are shown in Fig. 5. In this case the naive estimate is the traditional estimate of the parameter  $A$  using the point gauge data and not allowing for any errors in the gauge data. To the right of the naive  $A$  parameter estimate are the cases where more error has been included by applying the error model developed in Section 4 in the SIMEX framework. For each value of the multiple of error variance ( $\lambda$ ), the  $A$  values for each of the 300 simulations are shown, along with the mean estimate  $A^*$ . The fitted quadratic function is shown that provides the relationship between the mean estimates ( $A^*$ ) and the multiple of error variance ( $\lambda$ ). Finally the SIMEX estimate, which represents the case where there is no error in the rainfall, is shown with the dotted line which is the quadratic function extrapolated to  $\lambda = -1$ .

As can be seen in Fig. 5, the naive estimate of the  $A$  parameter is 167.6 and the SIMEX estimate is 178.8 – a difference of approximately 7% in the parameter estimates.



**Fig. 5.** Relationship between SIMEX error variance and estimated value of  $A$ . The naive estimate is found at  $\lambda = 0$  and SIMEX estimate at  $\lambda = -1$ .

Thus for a typical (average) reflectivity, the SIMEX rainfall intensity will be approximately 4% lower than if errors in gauge rainfall were not included (i.e. naive estimate). Although this is a small change compared to the other sources of error in hydrologic modelling, the important point is that this bias can be easily addressed using the method in this paper and therefore increases the accuracy of radar reflectivity measurements for characterising rainfall fields. The parameter  $A$  estimated by applying the SIMEX method is significantly different from the currently used biased naive estimate and shows that the consideration of point gauge uncertainty has an impact on the radar rainfall estimation. This

gauge uncertainty is classically ignored in radar Z–R relationship development by considering point gauge as a “ground truth”. Our result indicates that it is necessary to consider point gauge network uncertainty when developing radar Z–R relationships.

It is important to note that the correct Z–R relationship is also affected by random fluctuations introduced by variation in the rainfall drop size distribution both spatially and temporarily (Lee et al., 2007). Jaffrain and Berne (2012) reported that the spatial variability of the DSD within a radar pixel leads to variability of power law parameters, with deviations of between –2% and +15% in the pixel rainfall amount. The authors also reported relative variability in the Z–R relationship parameters  $A$  (–18% to 1%) and  $b$  (–4% to 1%).

## 6.2. Sensitivity of SIMEX method to assumptions

One of the main assumptions in the application of the SIMEX method is the form of the error distribution associated with the covariate. To test this, alternate error models for the gauge rainfall from Eureka were considered. All models led to a similar value for the extrapolated value of the spatial variability ( $\sigma_{cv}^2$ ) for the case of a single gauge in the  $1 \text{ km} \times 1 \text{ km}$  grid cell. When these models are used in the SIMEX framework, there are some differences in the final value for  $A^*$ , which are dependent on the magnitude of the errors in the adopted error model. In general, when errors are smaller, the difference between the two results is also smaller, something that would be expected from the rationale behind SIMEX. While this indicates that the SIMEX estimates are sensitive to the choice of the error model, one can also conclude that the presence of error is leading to a bias in a consistent direction for all models, pointing to the need for considering the uncertainty associated with the rain gauge network used in the formulation of any Z–R relationship.

The second major assumption in the SIMEX method is the choice of an extrapolation method for the bias. In this paper, a quadratic extrapolation function was used, but other extrapolation functions could also be used. Fig. 6 illustrates that the choice of the extrapolation function can have a small impact on the resulting parameter estimate. In this study we tested linear, quadratic and polynomial extrapolation functions. The quadratic extrapolation was adopted since all the methods lead to comparable results which provide confidence in the robustness of the extrapolation. It is clear that the uncertainty introduced by the extrapolation is much smaller than the sensitivity of the method to the adopted error model.

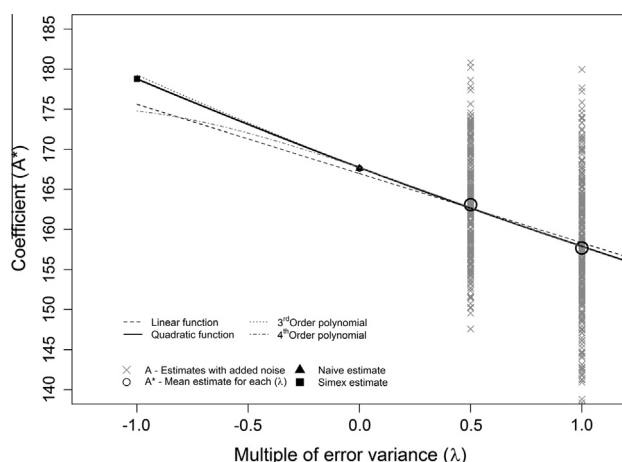


Fig. 6. SIMEX estimate of parameter  $A$  using a variety of functions to extrapolate to the no error case at  $\lambda = -1$ .

Other assumptions made in the SIMEX analysis include the number of realisations as well as the appropriate range and the increment for the values of the error variance factor. The choice of 300 realisations for the SIMEX method is arbitrary and could be varied. Previous hydrological applications of the SIMEX approach have used 300 realisations (Chowdhury and Sharma, 2008) and 500 realisations (Chowdhury and Sharma, 2007) and therefore 300 realisations were adopted for this current paper. We carried out sensitivity tests on the number of realisations and found that it only has a limited impact with variations of generally less than 1% in the value of the parameter  $A$ . However if SIMEX is used in other applications, then such parameters may become more important depending on the characteristics of the covariate-response sample. Similarly, a multiplicative error has been assumed for simplicity in the analysis reported here, and partly to maintain consistency with other studies on input uncertainty in the literature (McMillan et al., 2011; Renard et al., 2010; Tian et al., 2013). While choices such as these could be fine-tuned depending on the situation being analysed, the intention of this paper was to introduce the reader to the idea that point gauge uncertainty may affect radar rainfall relationships and demonstrate that SIMEX is an appropriate choice for resolving this issue.

## 7. Conclusion

Accurate spatial rainfall data is essential to get the best output from hydrological models. Generally, radar rainfall estimation accuracy is evaluated by comparing point gauge rainfall without considering gauge uncertainty. This paper makes a unique contribution as we have identified that point gauge uncertainty is a factor that may affect the radar rainfall estimation process. In this paper we have considered how this uncertainty impacts the Z–R relationship. We have identified that there are two main sources of error: the first one being due to recording errors in the tipping bucket gauge and the second one being the error introduced by not properly sampling the spatial variability of rainfall at the scale of the radar pixels. To address both sources of uncertainty, an error model is developed that quantifies the total error in the rain gauge measurements. The innovation in this error model is to consider large spatial aggregations of the data to overcome the low gauge density in the operational network, where at best there is only a single gauge for each radar pixel.

The SIMEX method is shown to be able to account for this input uncertainty and provide an unbiased estimate of the parameter  $A$  of the Z–R relationship. It is found that the SIMEX method leads to changes in the  $A$  parameter which will translate to biases in the resulting rainfall rates estimated from the Z–R relationship. Although the changes are small, they reflect a consistent bias in radar rainfall estimates that can be corrected using the ability of SIMEX to incorporate gauge uncertainty during the calibration of radar rainfall measurements. One main advantage of the SIMEX framework is that if different error models are found to be appropriate for other radar networks, then these can easily be used in place of the error model adopted in this paper. In addition, if a number of Z–R relationships are found to be appropriate for a particular location to reflect the physics of rainfall and meteorological of storm events, then a different error model could be used for each Z–R relationship. Finally, to know whether it is reasonable to ignore a source of error in modelling, first the impact of that error needs to be quantified. The SIMEX method provides a framework for this decision to be made.

In the current study, we have demonstrated that uncertainty in rain gauges measurements can be easily considered to produce unbiased Z–R relationships. Future work could couple this with

the uncertainty in the radar reflectivity (Jordan et al., 2003), particularly in the extrapolation between the levels of the CAPPI measurements and the ground level.

## Acknowledgements

The authors gratefully acknowledge the Australian Bureau of Meteorology for providing radar and rain gauge data for this study. The comments of two anonymous reviewers have greatly improved the presentation of the work. Authors acknowledge the Australian Research Council for partial funding for this work.

## References

- AghaKouchak, A., Habib, E., Bardossy, A., 2010. Modeling radar rainfall estimation uncertainties: random error model. *J. Hydrol. Eng.* 15 (4), 265–274.
- Anagnostou, E.N., Krajewski, W.F., Smith, J., 1999. Uncertainty quantification of mean-areal radar-rainfall estimates. *J. Atmosph. Oceanic Technol.* 16 (2), 206–215.
- Benoit, K., Laver, M., Mikheylov, S., 2009. Treating words as data with error: uncertainty in text statements of policy positions. *Am. J. Political Sci.* 53 (2), 495–513.
- Bonnifant, L. et al., 2009. Distributed hydrologic and hydraulic modelling with radar rainfall input: reconstruction of the 8–9 September 2002 catastrophic flood event in the Gard region, France. *Adv. Water Resour.* 32 (7), 1077–1089.
- Bringi, V.N., Rico-Ramirez, M.A., Thurai, M., 2011. Rainfall estimation with an operational polarimetric C-band radar in the United Kingdom: comparison with a gauge network and error analysis. *J. Hydrometeorol.* 12 (5), 935–954.
- Carroll, R.J., Küchenhoff, H., Lombard, F., Stefanski, L.A., 1996. Asymptotics for the SIMEX estimator in nonlinear measurement error models. *J. Am. Statist. Assoc.* 91 (433), 242–250.
- Carroll, R.J., Maca, J.D., Ruppert, D., 1999. Nonparametric regression in the presence of measurement error. *Biometrika* 86 (3), 541–554.
- Chowdhury, S., Sharma, A., 2007. Mitigating parameter bias in hydrological modelling due to uncertainty in covariates. *J. Hydrol.* 340 (3–4), 197–204.
- Chowdhury, S., Sharma, A., 2008. A simulation based approach for representation of rainfall uncertainty in conceptual rainfall runoff models. *Hydrol. Res. Lett.* 2, 5–8.
- Chumchean, S., Sharma, A., Seed, A., 2003. Radar rainfall error variance and its impact on radar rainfall calibration. *Phys. Chem. Earth* 28 (1–3), 27–39.
- Chumchean, S., Seed, A., Sharma, A., 2004. Application of scaling in radar reflectivity for correcting range-dependent bias in climatological radar rainfall estimates. *J. Atmosph. Oceanic Technol.* 21 (10), 1545–1556.
- Chumchean, S., Seed, A., Sharma, A., 2006a. Correcting of real-time radar rainfall bias using a Kalman filtering approach. *J. Hydrol.* 317 (1–2), 123–137.
- Chumchean, S., Sharma, A., Seed, A., 2006b. An integrated approach to error correction for real-time radar-rainfall estimation. *J. Atmosph. Oceanic Technol.* 23 (1), 67–79.
- Chumchean, S., Seed, A., Sharma, A., 2008. An operational approach for classifying storms in real-time radar rainfall estimation. *J. Hydrol.* 363 (1–4), 1–17.
- Ciach, G.J., 2003. Local random errors in tipping-bucket rain gauge measurements. *J. Atmosph. Oceanic Technol.* 20 (5), 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 (6), 1325–1347.
- Cole, S.J., Moore, R.J., 2008. Hydrological modelling using rain gauge- and radar-based estimators of areal rainfall. *J. Hydrol.* 358 (3), 159–181.
- Cole, S.J., Moore, R.J., 2009. Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Adv. Water Resour.* 32 (7), 1107–1120.
- Collier, C., 2009. On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models. *Meteorol. Appl.* 16 (1), 35–40.
- Cook, J.R., Stefanski, L.A., 1994. Simulation-extrapolation estimation in parametric measurement error models. *J. Am. Statist. Assoc.* 89 (428), 1314–1328.
- Faurès, J.-M., Goodrich, D.C., Woolhiser, D.A., Sorooshian, S., 1995. Impact of small-scale spatial rainfall variability on runoff modeling. *J. Hydrol.* 173 (1–4), 309–326.
- Goodrich, D.C., Faurès, J.-M., Woolhiser, D.A., Lane, L.J., Sorooshian, S., 1995. Measurement and analysis of small-scale convective storm rainfall variability. *J. Hydrol.* 173 (1–4), 283–308.
- Groisman, P.Y., Legates, D.R., 1994. The accuracy of United States precipitation data. *Bull. Am. Meteorol. Soc.* 75 (2), 215–227.
- Habib, E., Krajewski, W.F., 2002. Uncertainty analysis of the TRMM ground-validation radar-rainfall products: application to the TEFLUN-B field campaign. *J. Appl. Meteorol.* 41 (5), 558–572.
- Habib, E., Krajewski, W., Kruger, A., 2001. Sampling errors of tipping-bucket rain gauge measurements. *J. Hydrol. Eng.* 6 (2), 159–166.
- Holcomb, J.P., 1999. Regression with covariates and outcome calculated from a common set of variables measured with error: estimation using the SIMEX method. *Statist. Med.* 18 (21), 2847–2862.
- Hossain, F., Anagnostou, E.N., Dinku, T., Borga, M., 2004. Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty. *Hydrol. Process.* 18 (17), 3277–3291.
- Islam, T., Rico-Ramirez, M.A., Thurai, M., Han, D., 2012. Characteristics of raindrop spectra as normalized gamma distribution from a Joss–Waldvogel disdrometer. *Atmosph. Res.* 108, 57–73.
- Jaffrain, J., Berne, A., 2012. Influence of the subgrid variability of the raindrop size distribution on radar rainfall estimators. *J. Appl. Meteorol. Climatol.* 51 (4), 780–785.
- Jensen, N.E., Pedersen, L., 2005. Spatial variability of rainfall: variations within a single radar pixel. *Atmosph. Res.* 77 (1–4), 269–277.
- Jordan, P.W., Seed, A.W., Weinmann, P.E., 2003. A stochastic model of radar measurement errors in rainfall accumulations at catchment scale. *J. Hydrometeorol.* 4 (5), 841–855.
- Keblouti, M., Ouerdachi, L., Berhail, S., 2013. The use of weather radar for rainfall-runoff modeling, case of Seybouse watershed (Algeria). *Arab. J. Geosci.*, 1–11.
- Krajewski, W.F., Smith, J.A., 2002. Radar hydrology: rainfall estimation. *Adv. Water Resour.* 25 (8–12), 1387–1394.
- Krajewski, W.F., Kruger, A., Nespor, V., 1998. Experimental and numerical studies of small-scale rainfall measurements and variability. *Water Sci. Technol.* 37 (11), 131–138.
- Krajewski, W.F., Ciach, G.J., McCollum, J.R., Bacotiu, C., 2000. Initial validation of the global precipitation climatology project monthly rainfall over the United States. *J. Appl. Meteorol.* 39 (7), 1071–1086.
- Lebel, T., Amani, A., 1999. Rainfall estimation in the Sahel: what is the ground truth? *J. Appl. Meteorol.* 38 (5), 555–568.
- Lee, G.W., Seed, A.W., Zawadzki, I., 2007. Modeling the variability of drop size distributions in space and time. *J. Appl. Meteorol. Climatol.* 46 (6), 742–756.
- Looper, J.P., Vieux, B.E., 2012. An assessment of distributed flash flood forecasting accuracy using radar and rain gauge input for a physics-based distributed hydrologic model. *J. Hydrol.* 412, 114–132.
- Mapiam, P.P., Sriwongsitanon, N., Chumchean, S., Sharma, A., 2009. Effects of rain gauge temporal resolution on the specification of a Z–R relationship. *J. Atmosph. Oceanic Technol.* 26 (7), 1302–1314.
- Marcus, A.H., Elias, R.W., 1998. Some useful statistical methods for model validation. *Environ. Health Perspect.* 106, 1541–1550.
- Marshall, J.S., Palmer, W.M., 1948. The distribution of raindrops with size. *J. Appl. Meteorol.* 5, 165–166.
- McMillan, H., Jackson, B., Clark, M., Kavetski, D., Woods, R., 2011. Rainfall uncertainty in hydrological modelling: an evaluation of multiplicative error models. *J. Hydrol.* 400 (1), 83–94.
- Morin, E., Maddox, R.A., Goodrich, D.C., Sorooshian, S., 2005. Radar Z–R relationship for summer monsoon storms in Arizona. *Weather Forecast.* 20 (4), 672–679.
- Morrissey, M.L., Maliekal, J.A., Greene, J.S., Wang, J., 1995. The uncertainty of simple spatial averages using rain gauge networks. *Water Resour. Res.* 31 (8), 2011–2017.
- Ochou, A., Zahiri, E., Bamba, B., Koffi, M., 2011. Understanding the variability of Z–R relationships caused by natural variations in raindrop size distributions (DSD): implication of drop size and number. *Atmosph. Climate Sci.* 1 (3), 147–164.
- Pedersen, L., Jensen, N.E., Christensen, L.E., Madsen, H., 2010. Quantification of the spatial variability of rainfall based on a dense network of rain gauges. *Atmosph. Res.* 95 (4), 441–454.
- Pessoa, M.L., Bras, R.L., Williams, E.R., 1993. Use of weather radar for flood forecasting in the sieve river basin: a sensitivity analysis. *J. Appl. Meteorol.* 32 (3), 462–475.
- Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. *Water Resour. Res.* 46 (5), W05521.
- Rendon, S., Vieux, B., Pathak, C., 2013. Continuous forecasting and evaluation of derived Z–R relationships in a sparse rain gauge network using NEXRAD. *J. Hydrol. Eng.* 18 (2), 175–182.
- Rodríguez-Iturbe, I., Mejía, J.M., 1974. The design of rainfall networks in time and space. *Water Resour. Res.* 10 (4), 713–728.
- Schmidt, M., Lipson, H., 2013. Eureka Software. <<http://www.eureka.com/>>.
- Seed, A.W., Jordan, P.W., 2002. Mapview radar rainfall visualisation and processing software, Version 2.6. Centre for Australian Weather and Climate Research. The Australian Bureau of Meteorology.
- Shah, S.M.S., O'Connell, P.E., Hosking, J.R.M., 1996. Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. *J. Hydrol.* 175 (1–4), 89–111.
- Sharif, H.O., Ogden, F.L., Krajewski, W.F., Xue, M., 2002. Numerical simulations of radar rainfall error propagation. *Water Resour. Res.* 38 (8), 15–15–14.
- Staudenmayer, J., Ruppert, D., 2004. Local polynomial regression and simulation-extrapolation. *J. Roy. Statist. Soc.: Series B (Statist. Methodol.)* 66 (1), 17–30.
- Steiner, M., Smith, J.A., 2004. Scale dependence of radar-rainfall rates—an assessment based on raindrop spectra. *J. Hydrometeorol.* 5 (6), 1171–1180.
- Steiner, M., Smith, J.A., Burges, S.J., Alonso, C.V., Darden, R.W., 1999. Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* 35 (8), 2487–2503.
- Tian, Y. et al., 2013. Modeling errors in daily precipitation measurements: additive or multiplicative? *Geophys. Res. Lett.* 40 (10), 2060–2065.
- Uijlenhoet, R., Pomeroy, J., 2001. Raindrop size distributions and radar reflectivity? rain rate relationships for radar hydrology. *Hydrol. Earth Syst. Sci. Discuss.* 5 (4), 615–628.
- Verrier, S., Barthès, L., Mallet, C., 2013. Theoretical and empirical scale dependency of Z–R relationships: evidence, impacts, and correction. *J. Geophys. Res.: Atmosph.* 118 (14), 7435–7449.
- Vieux, B.E., Bedient, P.B., 1998. Estimation of rainfall for flood prediction from WSR-88D reflectivity: a case study, 17–18 October 1994. *Weather Forecast.* 13 (2), 407–415.



- Villarini, G., Krajewski, W.F., 2008. Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. *Adv. Water Resour.* 31, 1015–1023.
- Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: a rain gauge perspective. *J. Geophys. Res. – Atmosph.* 113 (D11).
- Viviroli, D., Zappa, M., Gurtz, J., Weingartner, R., 2009. An introduction to the hydrological modelling system PREVAH and its pre-and post-processing-tools. *Environ. Model. Software* 24 (10), 1209–1222.
- Vivoni, E.R., Entekhabi, D., Hoffman, R.N., 2007. Error propagation of radar rainfall nowcasting fields through a fully distributed flood forecasting model. *J. Appl. Meteorol. Climatol.* 46 (6), 932–940.
- Woldemeskel, F.M., Sharma, A., Sivakumar, B., Mehrotra, R., 2012. An error estimation method for precipitation and temperature projections for future climates. *J. Geophys. Res.: Atmosph.* 117 (D22), D22104.
- Wolff, D.B. et al., 2005. Ground validation for the tropical rainfall measuring mission (TRMM). *J. Atmosph. Oceanic Technol.* 22 (4), 365–380.
- Zawadzki, I.L., 1973. Errors and fluctuations of rain gauge estimates of areal rainfall. *J. Hydrol.* 18 (3–4), 243–255.
- Zhu, D., Peng, D.Z., Cluckie, I.D., 2013. Statistical analysis of error propagation from radar rainfall to hydrological models. *Hydrol. Earth Syst. Sci.* 17 (4), 1445–1453.