Missouri University of Science and Technology
Scholars' Mine

Civil, Architectural and Environmental Engineering Faculty Research & Creative Works
Civil, Architectural and Environmental Engineering

01 Jan 2014

Spatial And Temporal Modeling Of Radar Rainfall Uncertainties

Gabriele Villarini

Bong Chul Seo
Missouri University of Science and Technology, bongchul.seo@mst.edu

Francesco Serinaldi

Witold F. Krajewski

Follow this and additional works at: https://scholarsmine.mst.edu/civarc_enveng_facwork

Part of the Civil and Environmental Engineering Commons

Recommended Citation

G. Villarini et al., "Spatial And Temporal Modeling Of Radar Rainfall Uncertainties," *Atmospheric Research*, vol. 135 thru 136, pp. 91 - 101, Elsevier, Jan 2014.
The definitive version is available at https://doi.org/10.1016/j.atmosres.2013.09.007

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Civil, Architectural and Environmental Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

Atmospheric Research 135–136 (2014) 91–101

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Spatial and temporal modeling of radar rainfall uncertainties

Gabriele Villarini [a,*], Bong-Chul Seo [a], Francesco Serinaldi [b,c], Witold F. Krajewski [a]

[a] IIHR—Hydroscience & Engineering, The University of Iowa, Iowa City, IA, United States
[b] School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, United Kingdom
[c] Willis Research Network, London, United Kingdom

ARTICLE INFO

Article history:
Received 6 May 2013
Received in revised form 30 August 2013
Accepted 4 September 2013

Keywords:
Weather radar
Rainfall
Uncertainty
Space–time modeling
NEXRAD

ABSTRACT

It is widely acknowledged that radar-based estimates of rainfall are affected by uncertainties (e.g., mis-calibration, beam blockage, anomalous propagation, and ground clutter) which are both systematic and random in nature. Improving the characterization of these errors would yield better understanding and interpretations of results from studies in which these estimates are used as inputs (e.g., hydrologic modeling) or initial conditions (e.g., rainfall forecasting).

Building on earlier efforts, the authors apply a data-driven multiplicative model in which the relationship between true rainfall and radar rainfall can be described in terms of the product of a systematic and random component. The systematic component accounts for conditional biases. The conditional bias is approximated by a power-law function. The random component, which represents the random fluctuations remaining after correcting for systematic uncertainties, is characterized in terms of its probability distribution as well as its spatial and temporal dependencies. The space–time dependencies are computed using the non-parametric Kendall's τ measure. For the first time, the authors present a methodology based on conditional copulas to generate ensembles of random error fields with the prescribed marginal probability distribution and spatio-temporal dependencies.

The methodology is illustrated using data from Clear Creek, which is a densely instrumented experimental watershed in eastern Iowa. Results are based on three years of radar data from the Davenport Weather Surveillance Radar 88 Doppler (WSR-88D) radar that were processed through the Hydro-NEXRAD system. The spatial and temporal resolutions are 0.5 km and hourly, respectively, and the radar data are complemented by rainfall measurements from 11 rain gages, located within the catchment, which are used to approximate true ground rainfall.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Studies examining uncertainties in radar rainfall estimates span over three decades, with the first studies being completed by Harrold et al. (1974), Wilson and Brandes (1979), and Austin (1987), just to cite a few. Over this time period, radar rainfall estimation has improved significantly: compared to Wilson and Brandes (1979), Krajewski et al. (2010) found a reduction on the order of 33% in the average differences between radar and rain gages. These comparisons are promising and reveal that we are on the right path towards reducing uncertainties in radar rainfall estimation. Further reductions will result from the increasing availability of polarimetric measurements of rainfall (e.g., Zrnić and Ryzhkov, 1999; Bringi and Chandrasekar, 2001; Ryzhkov et al., 2005). These measurements will provide more accurate rainfall estimates by improving hydrometeor classification, attenuation correction, melting layer identification, and rainfall estimation in the presence of beam blockage and returns from non-hydrometeorological targets.

* Corresponding author at: IIHR—Hydroscience & Engineering, The University of Iowa, Iowa City, IA 52242, United States.
E-mail address: gabriele-villarini@uiowa.edu (G. Villarini).

0169-8095/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.atmosres.2013.09.007

Despite this progress, we will still have to account for uncertainties in rainfall estimates by weather radars for the foreseeable future (consult Villarini and Krajewski (2010b) for a recent review of the topic). Over the past few years, different modeling approaches have been proposed and developed to accomplish this task, and Mandapaka and Germann (2010) provide an extensive overview on the topic. There are two main ways to describe the uncertainties in radar rainfall estimates: in the first approach, the discrepancies between radar and ground "truth" data (generally based on rain gage measurements) represent the total uncertainties and reflect the overall contribution of different sources of uncertainties; in the second approach, the main sources of uncertainties are characterized separately and then combined to provide a description of the overall radar rainfall uncertainties. Both of these approaches have advantages and disadvantages, as discussed in Berne and Krajewski (2013). Due to the complications of parsing out the relative contribution of each of the sources of uncertainties and their inter-relationship, the direct description of the total rainfall uncertainties, albeit simpler, has been the selected approach in most of the recent studies. Therefore, this is the approach we will focus on in this study.

The methodology of this study is to directly compare rain gage measurements with corresponding radar rainfall estimates and develop models that can describe their discrepancies. We focus on error models that were developed to characterize the full statistical properties of the radar rainfall errors as opposed to just their first and second moments (consult Mandapaka and Germann (2010)). Ciach et al. (2007) described the radar rainfall uncertainties using a multiplicative error model which accounted for both systematic and random errors. This model was also used in other studies, including Habib et al. (2008), Villarini and Krajewski (2009b), and Villarini and Krajewski (2010a). Villarini and Krajewski (2010a) extended the model proposed by Ciach et al. (2007) to the additive form (see also Habib and Qin (2013) and Kirstetter et al. (2010)). Germann et al. (2009) developed a model in which the radar rainfall errors were proportional to the logarithm of the ratio between rain gage measurements and radar rainfall estimates. AghaKouchak et al. (2010b) described the errors as the sum of two random components: a purely random one and one that is a function of the radar rainfall values (consult AghaKouchak et al. (2010a) for an overview of other models by the same authors).

Each of these models has strengths and weaknesses. The main weaknesses of these models are related to their use of a Gaussian distribution for a multiplicative error model (e.g., Ciach et al., 2007; Villarini et al., 2009a) and their difficulties in generating random error fields with the prescribed spatial and temporal dependencies. This latter problem is one of the most significant impediments to the development of generators that are able to reproduce all of the error characteristics. Villarini et al. (2009a) discussed why it is not possible to simulate temporal dependencies in the model proposed by Ciach et al. (2007). Habib et al. (2008) considered the temporal correlation of the radar rainfall random errors to be low, and assumed that the random errors were independent in time (see also AghaKouchak et al. (2010a,b) for a similar assumption). Germann et al. (2009) accounted for temporal dependencies not at the pixel level but rather for basin-averaged lag-1 and lag-2 temporal correlations.

In this study, we will build on the error model described in Ciach et al. (2007) and show that it is possible to move away from the Gaussian distribution in favor of the more flexible mixture of gamma distributions. Moreover, both spatial and temporal dependencies computed with respect to the reference data can be preserved at the pixel scale. Our approach is general enough to be applicable for a wide range of probability distributions and spatio-temporal structures. Our improved capacity to generate ensembles which can reproduce the error characteristics computed with respect to the observational records will lead to improved understanding and interpretation of the results from those studies in which radar rainfall estimates are used as input, initial conditions, or reference datasets (e.g., Habib et al., 2008; Germann et al., 2009; Villarini and Krajewski, 2009a; Villarini et al., 2009b, 2010; Schröter et al., 2011; Liguori and Rico-Ramirez, 2012; Cunha et al., 2012).

The paper is organized as follows. In Section 2, we provide an overview of the error model and describe the data. Section 3 describes the results of our analysis. In Section 4, we summarize the main points and conclusions of this study.

2. Data and methodology

2.1. Radar rainfall error model

The radar rainfall error model we use takes its lead from Ciach et al. (2007). We will consider a multiplicative error model formulation, in which the relationship between true rainfall R_{true} and radar rainfall R_r can be described as the product of a systematic function $h(\cdot)$ and a random component $\varepsilon(\cdot)$:

$$R_{true} = h(R_r) \cdot \varepsilon(R_r). \tag{1}$$

The systematic component accounts for conditional biases, whereas the random component for the residual fluctuations remains after correcting for systematic uncertainties. The true rainfall R_{true} is approximated by rain gage measurements. Tian et al. (2013) showed that this formulation is a better choice with respect to an additive error model in terms of separating random and systematic errors, large variability in rainfall estimates, and predictive skill. Consult Villarini and Krajewski (2010a) for a comparison between multiplicative and additive model formulations in the context of radar rainfall error modeling.

Prior to estimating the model's components, we compute the unconditional bias B_0 and correct the radar rainfall estimates for it. It is computed as the ratio between rain gage measurements R_g and concurrent and co-located radar rainfall estimates over the domain of interest:

$$B_0 = \frac{\sum_i R_{g,i}}{\sum_i R^*_{r,i}}, \tag{2}$$

where $R_{g,i}$ is the rainfall accumulated by the ith rain gage over a given accumulation time and $R^*_{r,i}$ is the corresponding concurrent and co-located uncorrected radar rainfall estimate. We remove the unconditional bias with the following formula:

$$R_r = B_0 R_r^*, \tag{3}$$

where R_r represents the radar rainfall estimate corrected for the unconditional bias.

We then estimate the systematic component of the error model as a conditional expectation function:

$$h(r_r) = E[R_{true}|R_r = r_r], \tag{4}$$

where r_r represents the specific value of the random variable R_r. The benefit of conditioning on the R_r is that we can correct for conditional biases everywhere within the radar umbrella. Had we conditioned on the rain gage measurements, we would have been able to correct for the conditional biases only at those locations in which we had rain gages. Villarini et al. (2008b) compared both parametric and non-parametric methods to estimate the systematic function and found only minor differences between the two. In this study, we estimate the systematic function using a cubic spline (e.g., Hastie and Tibshirani, 1990) with five degrees of freedom.

After correcting the radar rainfall estimates for unconditional and conditional biases, we can estimate the random component:

$$\varepsilon(R_r) = \frac{R_{true}}{h(R_r)} \tag{5}$$

The random component is completely characterized once we have information about its probability distribution and spatial and temporal dependencies. Previous studies (e.g., Ciach et al., 2007; Villarini and Krajewski, 2009b, 2010a) described the random component with a Gaussian distribution. Moreover, $\varepsilon(\cdot)$ exhibited significant spatial and temporal correlations. The modeling of the error component in these previous studies was subject to a series of limitations. First of all, by using a Gaussian distribution in a multiplicative error model, it is likely that some of the generated variates would be negative, resulting in negative rainfall values which are not physically possible. As discussed in Villarini et al. (2009a), these values were automatically set to zero in generating ensembles of error fields, which impacts the error dependence structure. Moreover, because the error variance was a function of R_r, it was not possible to generate ensembles of random error fields that were correlated in both space and time; only the spatial component was preserved (Villarini et al., 2009a). As discussed in Section 1 and summarized by Mandapaka and Germann (2010), the error models developed so far cannot reproduce all the characteristics of the random errors.

The methodology we introduce in this study addresses all of these limitations building on the modular framework proposed by Serinaldi and Kilsby (2012) to simulate monthly rainfall series at multiple sites. Instead of using a Gaussian distribution to describe marginal distribution of the random component, we focus on a mixture of gamma distributions. The advantages of using the gamma distribution are its flexibility (it can be highly skewed or symmetric) and the fact that values are defined only over the positive axis; this addresses the problem of potential negative rainfall. The use of a mixture of gamma distributions allows a high degree of flexibility in describing the random component. We can formally write the probability density function of a mixture distribution as:

$$f(\varepsilon) = \sum_{j=1}^{n} w_j p_j(\varepsilon), \tag{6}$$

where w_j are the weights such that $w_j \geq 0$ (for $j = 1, 2, \ldots, n$) and $\sum_{j=1}^{n} w_j = 1$. Here, p_j refers to the jth probability density function and is represented by a gamma distribution:

$$p_j\left(\varepsilon|\mu_j, \sigma_j\right) = \frac{1}{\left(\sigma_j^2 \mu_j\right)^{1/\sigma_j^2}} \frac{\varepsilon^{\frac{1}{\sigma_j^2}-1} e^{-\varepsilon/(\sigma_j^2 \mu_j)}}{\Gamma\left(\sigma_j^2\right)}, \tag{7}$$

where $\mu_j > 0$ and $\sigma_j > 0$. Because of the multiplicative nature of this error model, we focus on the modeling of the random component for radar rainfall values larger than 0.3 mm.

The selection of the appropriate value of n will be based on the Akaike Information Criterion (AIC; Akaike, 1974). Values of the AIC will be computed for mixtures with different number of terms. We will select the value of n for which AIC is the smallest. This methodology is in agreement with the principle of parsimony, and results in a trade-off between model complexity and goodness of fit, with final models with few parameters, simple structures and reasonable accuracy. The goodness-of-fit of this mixture distribution is assessed through comparison of the empirical and fitted survival functions, and examination of the residuals, which should be white noise if the selected model is able to describe the systematic behavior in the data. We will accomplish this by computing the first four moments of the residuals, together with their Filliben correlation coefficient (Filliben, 1975), and by visual examination of their worm plot (van Buuren and Fredriks, 2001), which are detrended versions of quantile–quantile plots.

The spatial and temporal dependences in previous studies were generally estimated using Pearson's correlation coefficient ρ due to the assumption of a Gaussian distribution of the residuals. In our study, we use an estimator that is not affected by departures of the data from the Gaussian distribution (e.g., Habib et al., 2001; Serinaldi, 2008). We use the Kendall's τ measure of association (e.g., Nelsen, 2006) which is non-parametric (no distributional assumptions are required).

Modeling and generating random fields that are distributed according to a mixture of gamma distribution with a given spatio-temporal dependence structure is performed using the concept of conditional copulas (Patton, 2006). This approach is well known in econometrics (e.g., Fantazzini, 2008; Patton, 2009, and references therein) but is uncommon in hydrometeorology (e.g., Laux et al., 2011; Vogl et al., 2012). Copulas are multivariate distributions that allow for construction of joint distributions with arbitrary marginals using Sklar's theorem (Sklar, 1959). Focusing on a generic d-dimensional case, we can write $F_{\mathbf{X}}(x_1, \ldots, x_d) = C\left(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)\right)$, where $\mathbf{X} =$

$\{X_1, \ldots, X_d\}$, is a vector of d generic random variables with marginal distributions F_{X_i}, for $i = 1, \ldots, d$, and C is their copula. Referring to Nelsen (2006), Genest and Favre (2007), De Michele and Salvadori (2007) and Salvadori et al. (2007) for thorough introductions to copula theory and applications, we recall that the copula theory involves (time) independent and identically distributed (iid) multivariate random variables. However, in many real-world problems such as radar error field modeling, this hypothesis is not realistic. Fortunately, the copula approach can be applied in this case by resorting to the concept of conditional copulas introduced by Patton (2006). In particular, time dependence can be modeled by conditioning $\mathbf{X}_t$ on the previous observations $\mathbf{X}_t, \ldots, \mathbf{X}_{t-k}$. The conditioning variables can also be generic exogenous variables $\mathbf{Y}$. In the time series context, Sklar's theorem may be extended as follows (Patton, 2006; Fantazzini, 2008; de Melo Mendes and Costa Lopes, 2011):

$$F_{\mathbf{X},t}(x_1, \ldots, x_d|A_t) = C_t\left(F_{X_1,t}(x_1|A_t), \ldots, F_{X_d,t}(x_d|A_t)|A_t\right), \qquad (8)$$

where C_t is the copula at times t, and

$$A_t = \sigma\{\mathbf{x}_{t-1}, \ldots, \mathbf{x}_{t-k}, \mathbf{y}\}, \qquad (9)$$

is the σ-algebra generated by all past joint information up to time t provided by the sample $\mathbf{x}_{t-1} = \{x_{1,t-1}, \ldots, x_{d,t-1}\}, \ldots, \mathbf{x}_{t-k} = \{x_{1,t-k}, \ldots, x_{d,t-k}\}$ and eventual covariates $\mathbf{y} = \{y_1, \ldots, y_n\}$. Sklar's theorem for conditional distributions implies that the conditioning variable A_t must be the same for both marginal distributions and the copula. Failure to use the same information set for all components on the right-hand side of Eq. (8) generally implies that $F_{\mathbf{X},t}$ is not a valid conditional joint distribution function. Fermanian and Wegkamp (2012) consider the implications of a failure to use the same information set and define a so-called conditional pseudo copula. However, when each variable depends on its own previous lags but not on the lags of any other variable, Eq. (8) describes a valid conditional distribution. In this study, we assume that this hypothesis is a reasonable assumption. Since in the present case study the random variables $\mathbf{X}$ are the radar errors computed at d different sites and A_t denotes the at-site information related to the past error values, the previous hypothesis corresponds to assuming that the cross-correlation (spatial correlation) can be studied independently of the auto-correlation (temporal correlation), namely, that the spatio-temporal correlation function is separable. The conditional copula method is also denoted as the dynamic copula method when it implies the modeling of the time varying marginals and dependence structures.

From an operational point of view, the conditional copula method allows for splitting the analysis and modeling of marginals and dependence structure as follows (e.g., Grégoire et al., 2008; Reboredo, 2011). First, the d radar error time series are modeled by suitable time series models such as the auto regressive integrated moving average (ARIMA) models. Under the separability hypothesis, the residuals of the univariate time series models at different locations and times t are temporally independent copies of d spatially dependent random variables suitable to be modeled by a multivariate distribution. Since the proposed framework is devised for arbitrary dimensions d, for describing pairwise correlations, and simulating radar error maps over regular spaced grids, we adopt a meta-Gaussian copula, whose dependence structure, for high dimensions, corresponds to that of meta-Gaussian random fields with an appropriate covariance function. Of course, other dependence structures can be used; however, for high-dimensional problems the set of available models dramatically decreases and preliminary analyses showed that some competitors such as Student t-copulas do not improve the final output.

The modeling and simulation approach can be summarized as follows:

1. A mixture of gamma distributions is used to model the (at-site) radar error marginal distributions.
2. Therefore, the normal quantile transformation (NQT) is used to obtain marginals that are approximately Gaussian.
3. These transformed time series are then modeled by an autoregressive-moving average (ARMA) model, with both the AR and MA components being of order 1 [ARMA(1,1)]. ARMA is used to model the at-site information A_t. In this case, $A_t = \sigma\{\mathbf{x}_{t-1}, \boldsymbol{\xi}_{t-1}\}$, where $\mathbf{x}_{t-1}$ and $\boldsymbol{\xi}_{t-1}$ denote the radar error values and the ARMA residuals at times $t-1$, respectively. The term $\boldsymbol{\xi}_{t-1}$ corresponds with $\mathbf{y}$ in Eq. (9).
4. A power law covariance function is fitted to the empirical values of the cross-correlation of the ARMA(1,1) residuals computed at the rain gage sites (see Section 3). This step allows smoothing the empirical correlation matrix and exploiting the theory and tools developed for random fields modeling and simulation.
5. To simulate space–time dependent radar error fields (over a regular grid), this covariance function is used to simulate suitable Gaussian random fields, whose realizations are spatially correlated (according to the covariance function) and temporally independent.
6. The temporally independent sequences resulting from step (5) for each spatial location are used as (spatially correlated) innovations to feed at-site ARMA(1,1) models. This allows for the simulation of random fields that are both spatially and temporally correlated.
7. Finally, the simulated values are transformed back by using the inverse of the NQT which exploits the at-site fitted mixture of gamma distributions. NQT provides the link between the spatio-temporal meta-Gaussian dependence structure and the at-site (locally varying) marginal distribution.

Note that the described methodology is general enough to be applicable for a wide range of spatial and temporal dependence models. In this study, the spatial and temporal dependencies are estimated by using Kendall's τ and are then transformed into the corresponding Pearson correlation coefficient ρ as follows (e.g., Fang et al., 2002):

$$\tau = \frac{2}{\pi}\sin^{-1}\rho. \qquad (10)$$

This allows for a non-parametric estimation that is not affected by the failure of the hypothesis of Gaussian marginals and is effective if the spatial and/or temporal dependence structures are meta-elliptical.

We perform all the calculations in R (R Development Core Team, 2012) using the freely available gamlss (Stasinopoulos et al., 2007) and RandomFields (Schlather, 2012) packages.

2.2. Data

Our radar rainfall error model relies heavily on the availability of radar and rain gage data. For this study, we focus on Clear Creek, which is a densely instrumented experimental watershed in eastern Iowa (Fig. 1). We collected 19 (March 2010 to May 2012, excluding winter months) and five (June to October 2012) months of rain gage and radar data for model parameter estimation and validation, respectively.

Radar rainfall estimates used in this study were generated using the Hydro-NEXRAD algorithms documented in Krajewski et al. (2011). Clear Creek is between 100 km and 140 km from the Davenport, Iowa, Weather Surveillance Radar 88 Doppler (WSR-88D) radar (KDVN), with no significant orography affecting the radar beam propagation. Level II volume data for the KDVN radar were obtained from the National Climatic Data Center. Due to the location of the basin with respect to the radar site, we do not account for possible range effects. As a data quality control step, non-precipitation radar echoes such as ground clutters and anomalous propagation (AP) effects (e.g., Battan, 1973) were eliminated using the method proposed by Steiner and Smith (2002). The method identifies non-precipitation echoes based on the horizontal and vertical structure of measured radar reflectivity. We applied two threshold values (10 dBZ and 53 dBZ) to define effective minimum rainfall and hail contamination. We used multiple elevation angle data to construct reflectivity maps, and a non-parametric kernel function (e.g., Seo et al., 2011) was used to avoid discontinuity in the map due to the elevation angle transition. We then applied the NEXRAD default Z–R relationship ($Z = 300R^{1.4}$; Fulton et al., 1998) to transform radar reflectivity fields to rainfall intensities. To generate hourly rainfall accumulations, we integrated rainfall intensity maps over time. The spherical coordinates of radar observations were transformed into geographic coordinates with a spatial resolution of 1/4 arc-minute (approximately 0.5 km for the study area). Due to the fine spatial resolution and the hourly temporal scale, the potential effects of spatial sampling errors are negligible (e.g., Kitchen and Blackall, 1992; Villarini et al., 2008a; Villarini and Krajewski, 2008).

The research group at The University of Iowa operates high-quality density rain gage networks in the vicinity of Iowa City, Iowa. One of the networks in Clear Creek comprises 11 rain gage sites, and all sites are equipped with dual tipping-bucket gages, a data logger, and a cell phone. The temporal resolution of the processed rainfall data varies from 5-minute to daily. Detailed information on data quality control, processing, and transmission to data storage is documented in Ciach (2003) and Seo and Krajewski (2010). The period of rain gage data collection is identical to that of radar rainfall products.

3. Results

The first step in the development of the radar rainfall error model is to calculate the unconditional bias which is defined as the ratio between rain gage measurements and the corresponding radar rainfall estimates. Over the study period, the unconditional bias is equal to 0.7, which indicates an over-estimation by the radar with respect to the rain gages. This over-estimation is clear when we examine the scatterplots between radar and rain gages (Fig. 2, top panel). After correcting for the unconditional biases, we see that, overall, the points are scattered around the 1:1 line (Fig. 2, bottom panel) with a reduction of the root mean squared error (from 0.73 mm before bias correction to 0.52 mm after bias correction). After accounting for unconditional biases, we can compute the conditional biases as a conditional expectation function $h(\cdot)$. The radar tends to over-estimate large rainfall values, with $h(\cdot)$ lying below the 1:1 line (red line in Fig. 2, bottom panel). The conditional expectation function can be parameterized by the power-law function $y = x^{0.963}$. This parameterization is able to successfully reproduce the behavior of the conditional expectation function computed using cubic splines. The tendency for over-estimation of high rainfall values by the radar is consistent with what was observed in previous studies (e.g., Ciach et al., 2007; Villarini and Krajewski, 2009b, 2010a). The selected parameterization also preserves the zero-rainfall areas due to the multiplicative nature of our model.

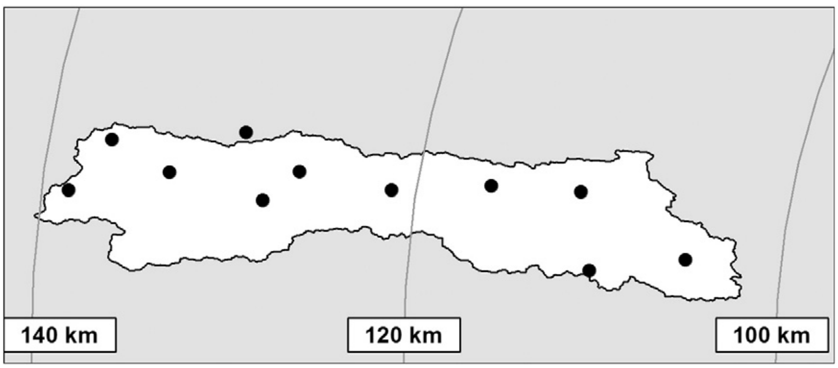

Fig. 1. Map showing the location of Clear Creek and of the 11 rain gages (black circles) used as ground reference. The distance of the watershed from the Davenport radar (KDVN) is also shown.

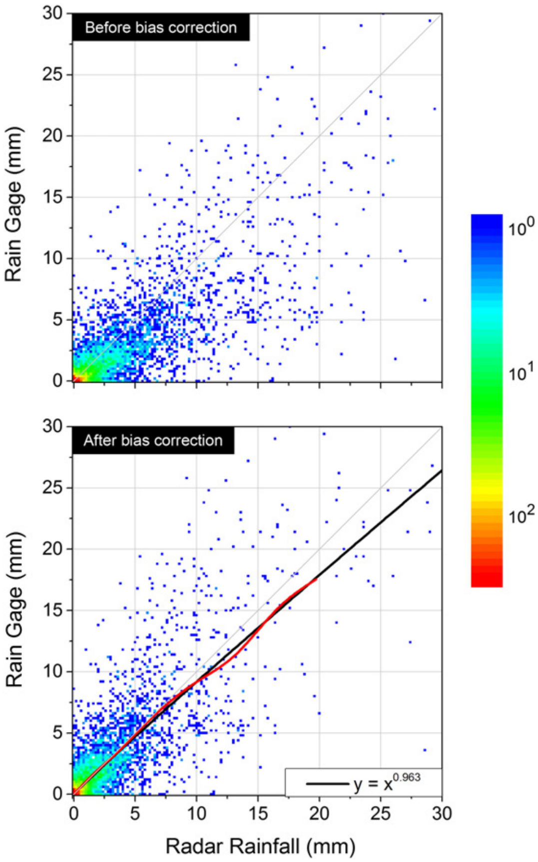

Fig. 2. Scatterplots of the rain gage measurements versus radar rainfall estimates before (top panel) and after (bottom panel) correction for the unconditional bias. In the bottom panel, the red line represents the systematic component (computed using a cubic spline with five degrees of freedom), while the solid black line represents its fit with a power law function.

Even though we estimated conditional and unconditional biases, there is still considerable variability in the scatterplot (Fig. 2, bottom panel). This variability is captured and described by the random component. A complete description of the random component requires characterizing their probability distribution and spatial and temporal dependencies. Once we have estimated all of these characteristics from the data, we can use the methodology based on conditional copulas (Section 2.1) to generate ensembles of radar rainfall fields conditioned on the observed radar rainfall maps and consistent with the observed error properties.

Let us start with the probability distribution. Instead of resorting to a Gaussian distribution, we use a mixture of gamma distributions because of its flexibility and because it is defined only for positive values. In previous studies (e.g., Ciach et al., 2007; Villarini and Krajewski, 2009b, 2010a; Habib and Qin, 2013), the standard deviation of the random component was considered a function of the radar rainfall values; here, we use a distribution with constant parameters. According to AIC, the random component can be described by a mixture of three gamma distributions, for a total of eight parameters to be fitted. The location μ_j (scale σ_j) parameters of the gamma distributions are 1.21, 0.82, and 3.28 (0.73, 0.44, and 0.95). The corresponding weights w_j are 0.623, 0.330, and 0.047, respectively. The visual examination of the results in Fig. 3 supports the choice of the model. There is a very good agreement between the empirical and fitted survival functions (Fig. 3, top panel), and the worm plots do not point to any problems with the fitted model (Fig. 3, bottom panel). Moreover, the first four moments of the residuals (mean equal to 0, variance equal to 1, coefficient of skewness equal to 0, coefficient of kurtosis equal to 2.99, and Filliben correlation coefficient equal to 0.9998) provide strong evidence that they are white noise. All of these diagnostics strongly suggest that the gamma mixture is able to describe very well the random component.

The spatial and temporal dependencies are described by Kendall's τ due to the non-Gaussian nature of random errors (Fig. 4, top panel). The values of the temporal and spatial correlations are comparable to other published studies for Oklahoma (Ciach et al., 2007; Villarini and Krajewski, 2010a), south–west England (Villarini and Krajewski, 2009b), and Louisiana (Habib and Qin, 2013). Using Eq. (10), we can transform Kendall's τ into Pearson's ρ, which represents the input correlation coefficient for our generator (Fig. 4, bottom panels). We have parameterized the Pearson-based spatial

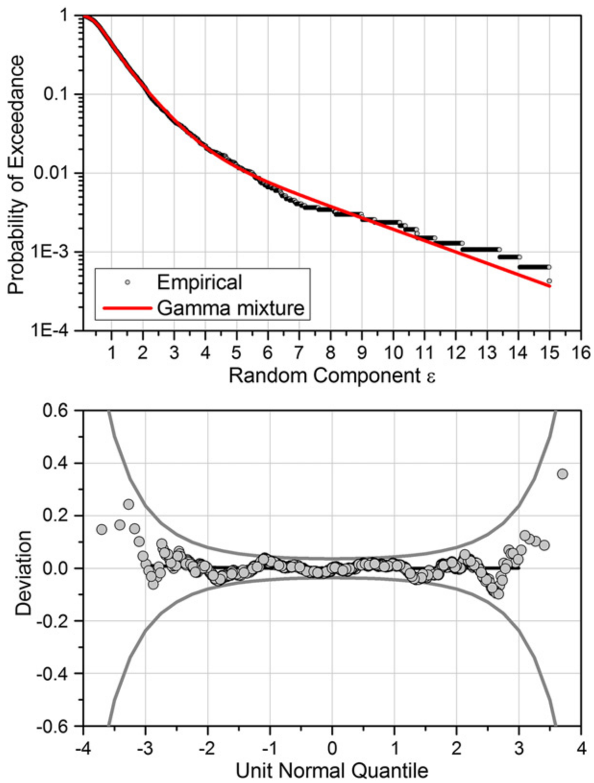

Fig. 3. Top panel: Survival function of the random component and of the fitted mixture of three gamma distributions. Bottom panel: Worm plot associated with the fit of the random component by means of the gamma mixture in the top panel.

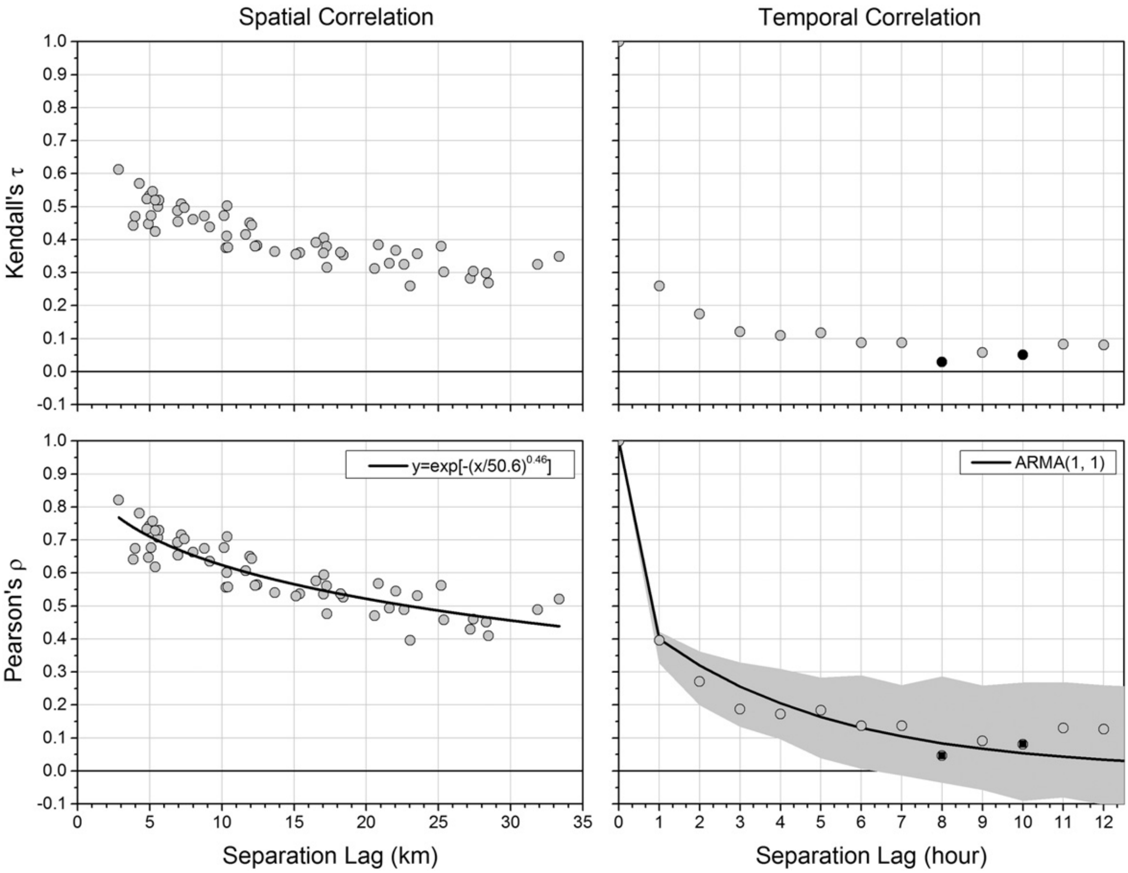

Fig. 4. Top panels: plots of the spatial and temporal correlations of the random component estimated using Kendall's τ. Bottom panel: Pearson's estimator ρ of the correlation derived using Eq. (10) from the results in the top panels. The solid black lines represent the fitting of the data with parametric models. The black circles indicate lags which are not significantly different from zero at the 5% level. The gray shaded area in the bottom-right panel shows the sampling uncertainties associated with the ARMA(1, 1) model (the value of the AR and MA parameters is 0.8 and −0.5, respectively).

correlation ($C(\cdot)$) using a power law model of the form (e.g., Chilés and Delfiner, 1999):

$$C(r) = \exp\left[-\left(\frac{r}{a}\right)^{\alpha}\right], \tag{11}$$

where r is the intergage distance, a is equal to 50.6 km, and α is equal to 0.46.

For the Pearson-based temporal correlation, we used an ARMA(1,1) model, with the value of the AR and MA parameters equal to 0.8 and −0.5, respectively. This model is comparable to a fractional Gaussian noise with a Hurst coefficient of 0.722 (O'Connell, 1971), indicating a power-law decay in the temporal correlation. This model describes well the temporal correlation for a number of lags (Fig. 4, bottom-right panel). Accounting for the sample size at each temporal lag, there is not enough statistical evidence (at the 1% confidence level) to reject the hypothesis that the estimated temporal correlation was generated by an ARMA (1,1) model with the parameters equal to 0.8 and −0.5 for the AR and MA components, respectively.

Now that we have identified the probability distribution and estimated the spatial and temporal dependencies, the random component is fully characterized. Using the approach described in Section 2.1, we generate Gaussian-distributed random fields correlated in space according to Eq. (11). These fields are used as innovation terms for the ARMA(1,1) models at each pixel. In this way, we have Gaussian fields correlated in space and time with the prescribed spatio-temporal structure. By using the inverse of the NQT, we can transform the Gaussian fields to gamma distributed fields. Fig. 5 shows that the spatial and temporal dependencies are generally preserved in the transformation and within the sampling uncertainties, even though it appears that the space-time fields generated according to our method tend to be less correlated than the theoretical correlation functions.

Since the random component is independent of the radar rainfall values, we can generate large ensembles of gamma distributed random fields with the prescribed spatial and temporal dependencies off-line. We can then correct the radar rainfall fields for unconditional and conditional biases and super-impose the error fields in real time. This methodology allows for the generation of ensembles of rainfall fields conditioned on the radar rainfall maps and reflecting the uncertainties characterized by the comparison with respect to rain gages. Fig. 6 presents three realizations from the ensemble generator for a three-hour period (3 September

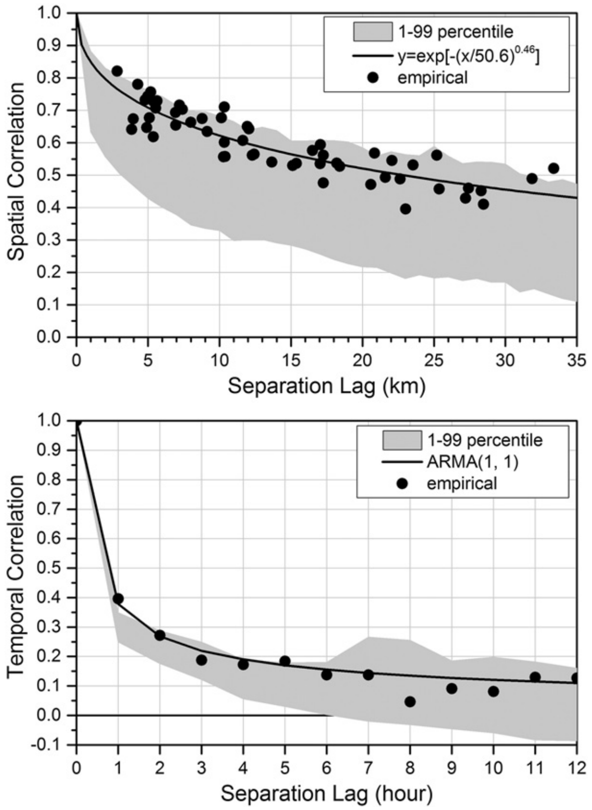

Fig. 5. Spatial (top panel) and temporal (bottom panel) correlations of the random component from the synthetic fields generated using conditional copulas. The solid black lines represent the theoretical (input) correlations. The gray areas represent the region between the 1st and 99th percentiles of the correlation sampling distribution. The black circles represent the correlations estimated from the data.

2011, 17–19 UTC). Fig. 7 compares rain gage measurements to radar rainfall estimates before and after accounting for radar rainfall uncertainties for a period not used for the calculations of the error model (June–October 2012). The radar tends to over-estimate rainfall with respect to the rain gage measurements. Moreover, it is clear how the largest uncertainties are associated with the largest rainfall values, highlighting the difficulties in estimating heavy rainfall.

4. Conclusions

This study focuses on the characterization of the radar rainfall uncertainties and on the development of a generator that is able to reproduce these error characteristics. The results are based on 19 months of radar rainfall data from the Davenport KDVN WSR-88D radar and 11 rain gages located within and near Clear Creek, Iowa. The temporal and spatial resolutions are hourly and about 0.5 km, respectively. The main conclusions of this work can be summarized as follows:

1. Scatterplots indicate a reasonably good visual agreement between the radar and rain gage data, with the former generally over-estimating the rain amounts with respect to the latter. The unconditional bias (defined as the ratio between rain gage measurements and the corresponding radar rainfall estimates) is equal to 0.7. This bias correction results in a reduction of the root mean squared errors and in points that are scattered around the 1:1 line.
2. The error model we used consists of a systematic and a random component. The systematic function accounts for conditional biases and is represented by a conditional expectation function. Accounting for the unconditional bias is enough to remove most of the biases, even though there is a tendency towards over-estimation by the radar for larger rainfall values.

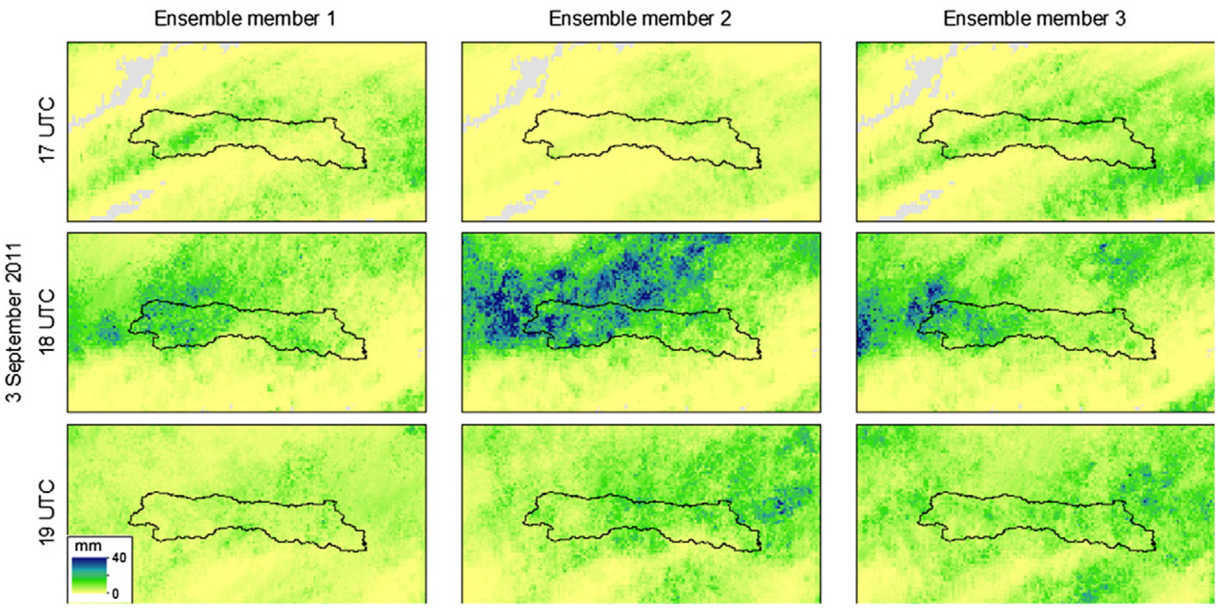

Fig. 6. Example of three realizations from the ensemble generator for the period 17–19 UTC on 3 September 2011.

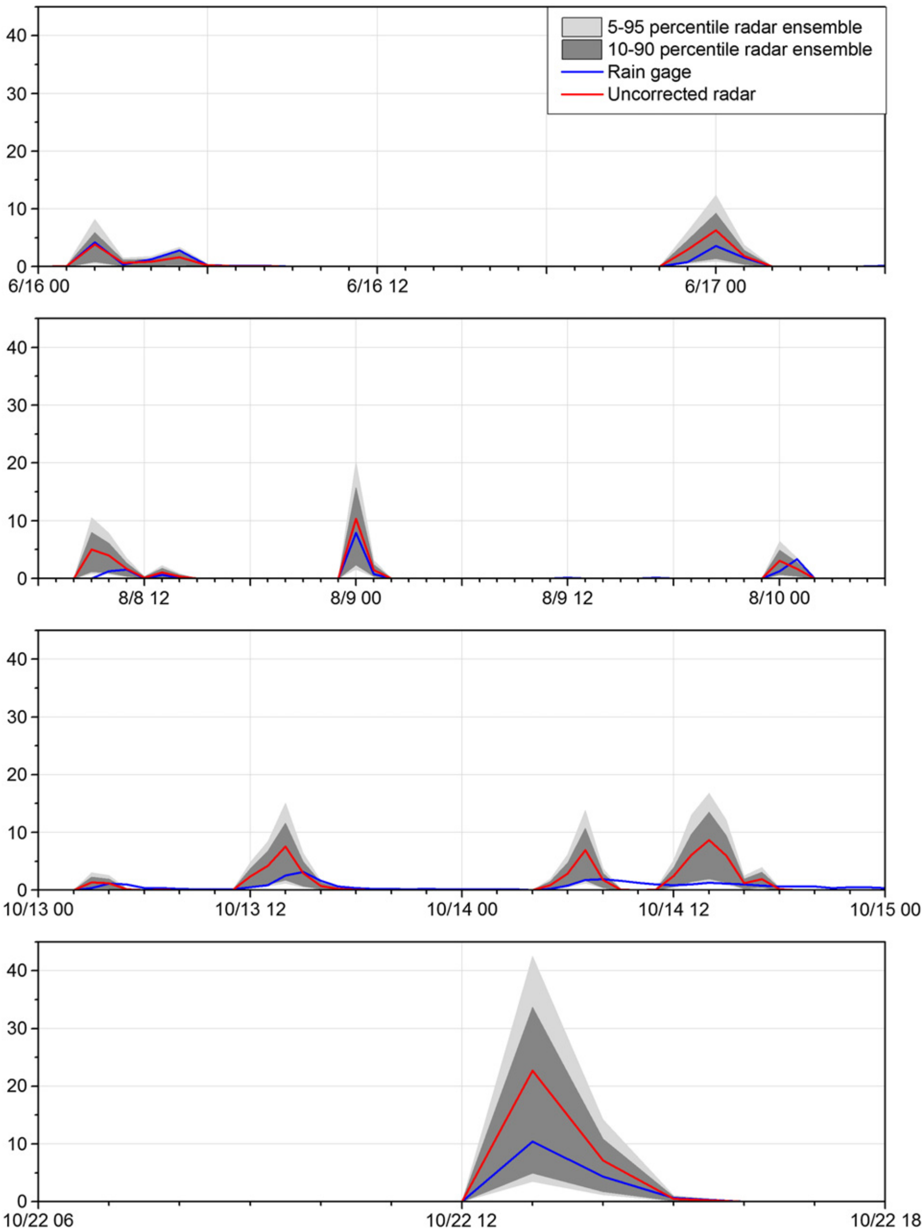

Fig. 7. Time series plots comparing the rainfall measured at one rain gage (blue lines) and the corresponding radar rainfall estimates before (red lines) and after (gray areas) accounting for radar rainfall uncertainties. The data refer to the year 2012.

3. The random component accounts for the variability remaining after correcting for conditional and unconditional biases. The comparison between radar and rain gage data indicates that there is a large degree of scatter that needs to be accounted for in order to fully describe the radar rainfall uncertainties. The probability distribution of the random component can be described by a mixture of three gamma distributions. We used Kendall's τ to describe the spatial and temporal dependencies of the random errors. The spatial correlation can be approximated by a power law model, and the temporal dependencies can be approximated by an ARMA(1,1) model.
4. Once we have characterized the structure of the random component in terms of its probability distribution and spatio-temporal dependencies, we have a generator that is able to reproduce them. This generator is based on conditional copulas and allows the successful generation of perturbation fields with the prescribed distribution and spatio-temporal characteristics. These random fields are then superimposed onto the bias-corrected radar rainfall

fields, allowing the generation of ensembles of rainfall fields with the prescribed spatio-temporal dependencies conditioned on the radar rainfall maps.

5. The formulation of this model is general enough that it can be used for any other radar rainfall products, including the new polarimetric-based estimates (e.g., Zrnić and Ryzhkov, 1999; Ryzhkov et al., 2005; Berne and Krajewski, 2013). At the same time, because the results are model-specific, it is not possible to make these results general for any other radars and/or products. While the unconditional bias depends on the radar calibration and is site specific, it is possible that the stratification of the data into different types of precipitation systems (e.g., convective, stratiform) could lead to a generalization of these results. Polarimetric measurements of rainfall could be extremely useful in performing this task.

Acknowledgments

Gabriele Villarini, Bong-Chul Seo, and Witold Krajewski acknowledge the financial support from the Iowa Flood Center, IIHR-Hydroscience & Engineering. Francesco Serinaldi acknowledges the financial support from the Willis Research Network. The authors would like to thank Dr. Stasinopoulos, Dr. Rigby, Dr. Akantziliotou, and Dr. Schlather for making the `gamlss` (Stasinopoulos et al., 2007) and `RandomFields` (Schlather, 2012) freely available in `R` (R Development Core Team, 2012).

References

AghaKouchak, A., Habib, E., Bárdossy, A., 2010a. A comparison of three remotely sensed rainfall ensemble generators. Atmos. Res. 98, 387–399.

AghaKouchak, A., Habib, E., Bárdossy, A., 2010b. Modeling radar rainfall estimation uncertainties: random error model. J. Hydrol. Eng. 15 (4).

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control. 19 (6), 716–723.

Austin, P.M., 1987. Relation between measured radar reflectivity and surface rainfall. Mon. Weather Rev. 115, 1053–1071.

Battan, L.J., 1973. Radar Observation of the Atmosphere. The University of Chicago Press.

Berne, A., Krajewski, W.F., 2013. Radar for hydrology: unfulfilled promise or unrecognized potential? Adv. Water Resour. 51, 357–366.

Bringi, V.N., Chandrasekar, V., 2001. Polarimetric Doppler Weather Radar. Princ. Applic. Cambridge University Press, Cambridge.

Chilés, J.P., Delfiner, P., 1999. Geostatistics. Wiley.

Ciach, G.J., 2003. Local random errors in tipping-bucket rain gauge measurements. J. Atmos. Oceanic Tech. 20 (5), 752–759.

Ciach, G.J., Krajewski, W.F., Villarini, G., 2007. Product-error-driven uncertainty model for probabilistic precipitation estimation with NEXRAD data. J. Hydrometeorol. 8 (6), 1325–1347.

Cunha, L.K., Mandapaka, P.V., Krajewski, W.F., Mantilla, R., Bradley, A.A., 2012. Impact of radar-rainfall error structure on estimated flood magnitude across scales: an investigation based on a parsimonious distributed hydrological model. Water Resour. Res. 48 (W10515). http://dx.doi.org/10.1029/2012WR012138.

de Melo Mendes, B.V., Costa Lopes, S.R., 2011. Dynamic copulas and long range dependence. Front. Finance Econ. 8 (2), 89–111.

De Michele, C., Salvadori, G., 2007. On the use of copulas in hydrology: theory and practice. J. Hydrol. Eng. 12 (4), 369–380.

Fang, H.B., Fang, K.T., Kotz, S., 2002. The meta-elliptical distributions with given marginals. J. Multivar. Anal. 82 (1), 1–16.

Fantazzini, D., 2008. Dynamic copula modelling for value at risk. Front. Finance Econ. 5 (2), 72–108.

Fermanian, J.-D., Wegkamp, M.H., 2012. Time-dependent copulas. J. Multivar. Anal. 110, 19–29.

Filliben, J.J., 1975. The probability plot correlation coefficient test for normality. Technometrics 17, 111–117.

Fulton, R., Breidenbach, J.P., Seo, D.J., Miller, D.A., O'Bannon, T., 1998. The WSR-88D rainfall algorithm. Weather Forecast. 13 (2), 377–398.

Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 347–368.

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

Grégoire, V., Genest, C., Gendron, M., 2008. Using copulas to model price dependence in energy markets. Energy Risk 5 (5), 58–64.

Habib, E., Qin, L., 2013. Application of a radar-rainfall uncertainty model to the NWS multi-sensor precipitation estimator products. Meteorological Applications 20, 276–286. http://dx.doi.org/10.1002/met.301.

Habib, E., Krajewski, W.F., Ciach, G.J., 2001. Estimation of rainfall interstation correlation. J. Hydrometeorol. 2 (6), 621–629.

Habib, E., Aduvala, V., Meselhe, E.A., 2008. Analysis of radar-rainfall error characteristics and implications for stream flow simulation uncertainty. Hydrol. Sci. J. 53, 568–587.

Harrold, T.W., English, E.J., Nicholas, C.A., 1974. The accuracy of radar-derived rainfall measurements in hilly terrain. Q. J. Roy. Meteorol. Soc. 100, 58–67.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall, London.

Kirstetter, P.E., Delrieu, G., Boudevillain, B., Obled, C., 2010. Toward an error model for quantitative precipitation estimation in the Cévennes–Vivarais region, France. J. Hydrol. 394 (1–2), 28–41.

Kitchen, M., Blackall, R.M., 1992. Representativeness errors in comparisons between radar and gauge measurements of rainfall. J. Hydrol. 134 (1–4), 13–33.

Krajewski, W.F., Villarini, G., Smith, J.A., 2010. Radar-rainfall uncertainties: where are we after thirty years of effort? Bull. Am. Meteorol. Soc. 91 (7), 87–94.

Krajewski, W.F., Kruger, A., Smith, J.A., Lawrence, R., Gunyon, C., Goska, R., Seo, B.-C., Domaszczynski, P., Baeck, M.L., Ramamurthy, M.K., Weber, J., Bradley, A.A., DelGreco, S.A., Steiner, M., 2011. Towards better utilization of NEXRAD data in hydrology: an overview of Hydro-NEXRAD. J. Hydroinformatics 13 (2), 255–266.

Laux, P., Vogl, S., Qiu, W., Knoche, H.R., Kunstmann, H., 2011. Copula-based statistical refinement of precipitation in RCM simulations over complex terrain. Hydrol. Earth Syst. Sci. 15 (7), 2401–2419.

Liguori, S., Rico-Ramirez, M.A., 2012. Quantitative assessment of short-term rainfall forecasts from radar nowcasts and MM5 forecasts. Hydrological Processes 26, 3842–3857.

Mandapaka, P.V., Germann, U., 2010. Radar-rainfall error models and ensemble generators. In: Gebremichael, M., Testik, F. (Eds.), Rainfall: State of Science. Geophys. Monogr. Ser., Am. Geophys. Union, pp. 247–264.

Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer-Verlag, New York.

O'Connell, P.E., 1971. A simple stochastic modeling of Hurst's law. IAHS Mathematical Models in Hydrology Symposium 169–187.

Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. Int. Econ. Rev. 47 (2), 527–556.

Patton, A.J., 2009. Copula-based models for financial time series. In: Mikosch, T., Kreiß, J.-P., Davis, R.A., Andersen, T.G. (Eds.), Handbook of Financial Time Series. Springer Berlin Heidelberg, pp. 767–785.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (ISBN 3-900051-07-0).

Reboredo, J.C., 2011. How do crude oil prices co-move?: A copula approach. Energy Econ. 33 (5), 948–955.

Ryzhkov, A.V., Schuur, T.J., Burgess, D.W., Heinselman, P.L., Giangrande, S.E., Zrnić, D.S., 2005. The joint polarization experiment, polarimetric rainfall measurements and hydrometeor classification. Bull. Am. Meteorol. Soc. 86 (6), 809–824.

Salvadori, G., De Michele, C., Kottegoda, N.T., Rosso, R., 2007. Extremes in Nature. Springer (292 pp.).

Schlather, M., 2012. RandomFields: Simulation and Analysis of Random Fields.

Schröter, K., Llort, X., Velasco-Forero, C., Ostrowski, M., Sempere-Torres, D., 2011. Implications of radar rainfall estimates uncertainty on distributed hydrological model predictions. Atmos. Res. 100, 237–245.

Seo, B.-C., Krajewski, W.F., 2010. Scale dependence of radar-rainfall uncertainty: initial evaluation of NEXRAD's new super-resolution data for hydrologic applications. J. Hydrometeorol. 11 (5), 1191–1198.

Seo, B.-C., Krajewski, W.F., Kruger, A., Domaszczynski, P., Smith, J.A., Steiner, M., 2011. Radar-rainfall estimation algorithms of Hydro-NEXRAD. J. Hydroinformatics 13 (2), 277–291.

Serinaldi, F., 2008. Analysis of inter-gauge dependence by Kendall's τ_k, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. Stoch. Env. Res. Risk A. 22, 671–688.

Serinaldi, F., Kilsby, C.G., 2012. A modular class of multisite monthly rainfall generators for water resource management and impact studies. J. Hydrol. 464–465, 528–540.

Sklar, A., 1959. Fonction de répartition á n dimensions et leurs marges, 8. Publications de l'Institut de Statistique de L'Université de Paris, pp. 229–231.
Stasinopoulos, D.M., Rigby, R.A., Akantziliotou, C., 2007. GAMLSS: Generalized Additive Models for Location Scale and Shape, R Package Version 1.6-0.
Steiner, M., Smith, J.A., 2002. Use of three-dimensional structure for automated detection and removal of non-precipitating echoes in radar data. J. Atmos. Oceanic Tech. 19, 673–686.
Tian, Y., Huffman, G.J., Adler, R.F., Tang, L., Sapiano, M., Maggioni, V., Wu, H., 2013. Modeling errors in daily precipitation measurements: additive or multiplicative? Geophys. Res. Lett. 40, 2060–2065.
van Buuren, S., Fredriks, M., 2001. Worm plot: a simple diagnostic device for modeling growth reference curves. Stat. Med. 20, 1259–1277.
Villarini, G., Krajewski, W.F., 2008. Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. Adv. Water Resour. 31 (7), 1015–1023.
Villarini, G., Krajewski, W.F., 2009a. Inference of spatial scaling properties of rainfall: impact of radar-rainfall estimation uncertainties. IEEE Geosci. Remote Sens. Lett. 6 (4), 812–815.
Villarini, G., Krajewski, W.F., 2009b. Empirically based modelling of radar-rainfall uncertainties for a C-band radar at different time-scales. Q. J. Roy. Meteorol. Soc. 135 (643), 1424–1438.
Villarini, G., Krajewski, W.F., 2010a. Sensitivity studies of the models of radar-rainfall uncertainties. J. Appl. Meteorol. Climatol. 49 (2), 288–309.
Villarini, G., Krajewski, W.F., 2010b. Review of the different sources of uncertainty in single-polarization radar-based estimates of rainfall. Surv. Geophys. 31, 107–129.
Villarini, G., Serinaldi, F., Krajewski, W.F., 2008a. Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. Adv. Water Resour. 31 (12), 1674–1686.
Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008b. Rainfall and sampling errors: a rain gauge perspective. J. Geophys. Res. 113 (D11102). http://dx.doi.org/10.1029/2007JD009214.
Villarini, G., Krajewski, W., Ciach, G.J., Zimmerman, D.L., 2009a. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. Water Resour. Res. 45 (W01404). http://dx.doi.org/10.1029/2008WR006946.
Villarini, G., Krajewski, W., Smith, J.A., 2009b. New paradigm for statistical validation of satellite precipitation estimates: application to a large sample of the TMPA 0.25 degree 3-hourly estimates over Oklahoma. J. Geophys. Res. 114 (D12106). http://dx.doi.org/10.1029/2008JD011475.
Villarini, G., Krajewski, W.F., Ntelekos, A.A., Georgakakos, K.P., Smith, J.A., 2010. Towards probabilistic forecasting of flash floods: the combined effects of uncertainty in radar-rainfall and flash flood guidance. J. Hydrol. 394 (1–2), 275–284.
Vogl, S., Laux, P., Qiu, W., Mao, G., Kunstmann, H., 2012. Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. Hydrol. Earth Syst. Sci. 16 (7), 2311–2328.
Wilson, J.W., Brandes, E.A., 1979. Radar measurement of rainfall — a summary. Bull. Am. Meteorol. Soc. 60, 1048–1058.
Zrnić, D.S., Ryzhkov, A.V., 1999. Polarimetry for weather surveillance radar. Bull. Am. Meteorol. Soc. 80, 389–406.