Hydrol. Earth Syst. Sci. Discuss., 9, 937–982, 2012
www.hydrol-earth-syst-sci-discuss.net/9/937/2012/
doi:10.5194/hessd-9-937-2012

Hydrology and Earth System Sciences Discussions

Copula-based assimilation of radar and gauge information to derive bias corrected precipitation fields

S. Vogl[1,2], P. Laux[2], W. Qiu[2], G. Mao[2], and H. Kunstmann[1,2]

[1]University of Augsburg, Institute for Geography, Regional Climate and Hydrology, 86135 Augsburg, Germany
[2]Karlsruhe Institute of Technology (KIT), Institute for Meteorology and Climate Research (IMK-IFU), Kreuzeckbahnstrasse 19, 82467 Garmisch-Partenkirchen, Germany

Received: 8 December 2011 – Accepted: 10 January 2012 – Published: 17 January 2012

Correspondence to: S. Vogl (stefanie.vogl@geo.uni-augsburg.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This study addresses the problem of combining radar information and gauge measurements. Gauge measurements are the best available source of absolute rainfall intensity albeit their spatial availability is limited. Precipitation information obtained by radar mimics well the spatial patterns but is biased for their absolute values. In this study Copula models are used to describe the dependence structure between gauge observations and rainfall derived from radar reflectivity at the corresponding grid cells. Only the positive pairs (radar > 0, gauge > 0) are considered. As not each grid cell can be assigned to one gauge, the integration of point information, i.e. gauge rainfall intensities, is achieved by considering the structure and the strength of dependence between the radar pixels and all the gauges within the radar image. Two different approaches namely *Maximum Theta* and *Multiple Theta* are presented. They finally allow for generating precipitation fields which mimic the spatial patterns of the radar fields and correct them for biases in their absolute rainfall intensities. The performance of the approach, which can be seen as a bias-correction for radar scenes, is demonstrated for the Bavarian Alps. The bias-corrected rainfall fields are compared to a field of interpolated gauge values (Ordinary Kriging) and are validated with the available gauge measurements. The simulated precipitation fields are compared to an operationally corrected radar precipitation field (RADOLAN). This comparison of the Copula-based approach and RADOLAN by different validation measures indicates that the Copula-based method successfully corrects for errors in the radar precipitation.

1 Introduction

For many hydrological analyses spatially distributed precipitation information is indispensible.

Whenever gauge data alone is the basis to derive a precipitation field, a wide range of interpolation methods is in use. Some examples include nearest neighbour (e.g. Isaaks

HESSD

9, 937–982, 2012

and Srivastava, 1989), inverse distance weighting, regression models (e.g. Bourrough and McDonell, 1998), trend surface analysis (e.g. Colins and Bolstadt, 1996), Splines (e.g. Hutchinson, 1998a,b; Bourrough and McDonell, 1998) and Kriging for which a large set of sub-methods has been developed (e.g. Isaaks and Srivastava, 1989; Bollerlslev, 1986; Goovaerts, 2000; Haberlandt, 2007). There has been also made use of mixed methods such as regression combined with Kriging (e.g. Erxleben et al., 2002) and others. Some methods attempt to assimilate additional information such as elevation of the terrain or additional measurements such as radar or from remote sensing (e.g. Haberlandt, 2007).

Even though the methods are different in nature they have one thing in common: the performance is highly dependent on the density of the observation network and on the complexity of the underlying terrain. This is problematic insofar as especially in regions with large height gradients usually the number of available meteorological stations, providing reliable rainfall measurements, is limited. Radar precipitation fields are supposed to be a good supplement as they are also covering areas with complex terrain in a high spatio-temporal resolution and the patterns of rainfall are assumed to be realistically reproduced.

There are multiple sources of errors when precipitation is derived from radar reflectivities such as the empirical (Z/R) reflectivity-rainfall relationship and errors induced by the radar measurement itself such as backscatter or shadowing effects (e.g. Joss and Lee, 1995). Whenever radar fields are used to drive meteorological or hydrological models these errors have to be taken into account (e.g. Cole and Moore, 2008; Singh, 1997) as they will directly propagate into the predicted variables. There have been many attempts to quantify the uncertainty of the radar measurement (e.g. Mandapaka et al., 2009; AghaKouchak et al., 2010c). To reduce the uncertainties in the radar derived precipitation fields many different approaches exist. It is very common to assimilate gauge information to the radar field (e.g. Brandes, 1975; Krajewski, 1987; Mazzetti and Todini, 2004; Ehret, 2003) to correct for errors in the absolute values. Usually, the existing approaches follow two steps: First the available gauge data is used to generate an interpolated precipitation field, and in the second step radar and interpolated field are combined.

The underlying assumption of Gaussian behaviour is restricting the performace of standard approaches, as many studies showed that the interdependence in hydrological or meteorological data sets is usually more complex (e.g. Gomez-Hernandez and Wen, 1998; Bárdossy, 2006; Bárdossy and Li, 2008).

Alternatively, a Copula approach can be used to describe the complex spatio-temporal dependence structure and assess for non-linear behaviour (e.g. Genest and Favre, 2007; Dupois, 2007). The Copula method is advantageous in several respects and has e.g. been used sucessfully in risk asessment (e.g. Embrechts et al., 2001; Frees and Valdez, 1998). Over the past years there is a remarkable increase in applications of Copulas in hydrometeorology. Copula-based models have been introduced for multivariate frequency analysis, geostatistical interpolation and multivariate extreme value analyses (e.g. De Michele and Salvadori, 2003; Dupois, 2007; Bárdossy, 2006; Genest and Favre, 2007; Renard and Lang, 2007; Schölzel and Friederichs, 2008; Bárdossy and Li, 2008; Zhang and Singh, 2008). For rainfall modelling, De Michele and Salvadori (2003) used Copulas to model intensity-duration of rainfall events. Favre et al. (2004) utilized Copulas for multivariate hydrological frequency analysis. Zhang and Singh (2008) carried out a bivariate rainfall frequency analysis using Archimedean Copulas. Renard and Lang (2007) investigated the usefulness of the Gaussian Copula in extreme value analysis. Kuhn et al. (2007) employed Copulas to describe spatial and temporal dependence of weekly precipitation extremes. Serinaldi (2008) studied the dependence of rain gauge data using the non parametric Kendall's rank correlation and the upper tail dependence coefficient (TDC). Based on the properties of the Kendall correlation and TDC, a Copula-based mixed model for modelling the dependence structure and marginals is suggested. Recently, Copula-based models for estimating error fields of radar information are developed (Villarini et al., 2008; AghaKouchak et al., 2010a,b). Most of these studies are carried out in the bivariate framework, describing dependency between two variates but there are also few examples of multivariate applications. Bárdossy (2006), Bárdossy and Li (2008) and Bárdossy and Pegram (2009) suggested a new method for geostatistical interpolation based on Copulas. They used multivariate Copulas to describe the spatial variability of groundwater quality parameters and developed a methodology to spatially interpolate these quantities.

In this study a new Copula-based method for correcting biases in radar precipitation fields is presented. Information from rain gauges is assimilated by investigating the dependence structure between rain gauges and radar fields.

The article is structured as follows: in Sect. 2 the study area and the data base is introduced. Section 3 reviews briefly the basic theory of Copulas and the procedure of simulating data using conditional CDFs. The new Copula-based bias-correction approach is introduced. Results of the application of the methodology in the Alpine space are presented in Sect. 4, followed by the discussion in Sect. 5 and the conclusions in Sect. 6.

2 Study area and data

2.1 Domain

In regions where precipitation reveals high spatio-temporal variability, such as alpine or prealpine terrain, it is a specific challenge to estimate realistic rainfall fields. Therefore the study area is chosen to be the Southern Bavarian Alps and alpine forelands in Germany. There are large gradients in elevation across the domain with lowest values in the north. Within the flat area of the domain, Munich is at 519 m a.s.l. The highest location is Mount Zugspitze at 2962 m a.s.l. Due to the complex orography and heterogeneity in topography, the domain is characterized by strong north southerly differentiations in soils, land-use, and climate. The mean annual temperature is around 7–8 °C in the alpine forelands and 4–5 °C in the southern part of the domain. The mean annual precipitation ranges from about 1100 mm in the northern part to more than 2000 mm in the south. Figure 1 shows an overview over the study area, highlighting the large north-southerly gradient in altitude.

In this study radar data from the Meteorological Observatory Hohenpeissenberg (MOHP) is used. The radar is a C-Band research weather radar operated by the German Weather Service (DWD). The observatory is located about 80 km southwest of Munich on Mount Hohenpeissenberg at an altitude of 1000 m a.s.l. The radar installation covers a circular area with a radius of 256 km, producing a scan every 5 min. For the study area, a square region of 100 × 100 grid cells with 1 km × 1 km resolution, centred at the radar station at mount Hohenpeissenberg is selected.

Preprocessing steps include a clutter correction of the radar reflectivities, which has been done by DWD. After that, rainfall amounts are derived by using the DWD standard Z (Reflectivity in dbz)/R (Rainfall in mm h^{-1}) relationship being

$$Z = 256 \cdot R^{1.42} \tag{1}$$

The data covers the summer months (June, July, and August) of the years 2005–2009. Precipitation data of 31 gauges within the chosen domain is retrieved from DWD, which covers the same period as the radar data. Figure 2 shows a snapshot of a radar based precipitation field measured at mount Hohenpeissenberg on 14 July 2008 (13:00). The coordinate system is centered at the position of the weather radar.

In addition to the radar precipitation derived by a standard Z/R relationship, operationally corrected radar-based rainfall time series are used. DWD developed a routine method for the online adjustment of radar precipitation by means of automatic surface precipitation stations (ombrometer) of DWD. The correction of DWD (hereinafter refered to as RADOLAN) includes a refined Z/R-relationship, contains orographic shading correction, statistical reduction of clutter, gradient smoothing and further preprocessing steps. The precipitation amounts observed at surrounding obrometer stations (more than 1000 for Germany) are interpolated and assimilated resulting in a bias-corrected radar precipitation field (Bartels et al., 2004).

Each of the 31 gauge stations used in this study is assigned to its corresponding grid cell in the radar domain. As the observation network is very sparse there is no grid cell

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

assigned to more than one gauge station. Figure 1 shows the positions of the available gauge stations and the position of the weather radar at Mount Hohenpeissenberg.

2.2 Data preprocessing and availability

The data pairs are checked for plausibility and erroneous or significantly anomalous values are removed. This procedure consists of the following steps:

1. High gauge values are checked for plausibility by comparing with nearby gauges.

2. Radar rainfall values smaller than 0.1 mm h^{-1} are set to zero as these measurements are considered to be erroneous.

3. The differences between neighbouring radar grid cells are calculated. Single values with absolute differences exceeding a threshold of 25 mm h^{-1}, revealing unrealistically large gradients in the radar field, are removed (Marx, 2007).

4. Only the remaining positive pairs (radar and gauge) of rainfall intensities are considered for further calculations in this study.

The remaining data pairs are divided into two subsets. The first set containing data of June, July, and August of 2006 and 2007 serves as calibration period. The second data set is containing data from the same months of 2008. Mean and standard deviation of gauge and radar (positive pairs) for the calibration period are listed in Table 1. Especially for station Hohenpeissenberg, the number of positive pairs is limited, only 70 pairs are available. As the gauge station is located very close to the radar observatory, the radar beam cannot capture the area above this gauge. This is the reason why mean and standard deviations of the positive pair differ significantly from those of the other stations.

3 Methodology

In the following section the theoretical background of Copula theory is briefly sketched including information about marginal distributions, a description of the new interpolation algorithm and applied validation methods.

3.1 Marginal distributions

Modelling the joint dependence structure with Copulas requires fitting marginal distributions to the data. The structure of the simulation technique allows to select a suitable distribution for each radar pixel and every gauge. In order to build the model as simple as possible, we search for one common theoretical distribution both for radar and gauge. In this study four different distribution functions are tested to identify the best fit.

The Normal distribution with mean μ and standard deviation σ

$$f(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)} \tag{2}$$

the Exponential distribution with parameter $\lambda \in \mathbb{R}_{>0}$

$$f_\lambda(x) := \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0\,, \\ 0 & \text{else} \end{cases} \tag{3}$$

the Weibull distribution with $\alpha > 0$, $\beta > 0$

$$f(x) := \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} \tag{4}$$

and the Gamma distribution with $b, p \in \mathbb{R}$ and $b > 0$, $p > 0$, where $\Gamma(p)$ denotes the value of the Gamma function at p

$$f_\lambda(x) := \begin{cases} \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} & \text{if} \quad x \geq 0\,, \\ 0 & \text{else} \end{cases} \tag{5}$$

HESSD
9, 937–982, 2012

To decide which univariate distribution is the best suitable for both radar and gauge data, additionally to the standard maximum likelihood approach the Akaike and the Bayesian information criterion is used. It is

$$\mathrm{AIC} = 2k - 2\ln(L) \tag{6}$$

and

$$\mathrm{BIC} = k\ln(n) - 2\ln(L)\,, \tag{7}$$

where k denotes the number of the free parameters of the model, n is the sample size and L is the maximized value of the likelihood function of the estimated model. The smallest value of AIC or BIC, respectively suggests the best fitting model/distribution.

For the application of a Copula model it is a prerequisite that the marginals are *iid* (independent and identically distributed). Daily precipitation time series exhibit some degree of autocorrelation and heteroscedasticity. Therefore, the data has to be checked before, and in case of violation of the *iid* prerequisite an ARMA-GARCH composite model is used to generate *iid* variables (Laux et al., 2011).

3.2 General introduction to Copula theory

In the following the basic definitions concerning Copulas are briefly reviewed. Let $(X_1,\ldots,X_n)$ denote a n-tupel of random variables and $(x_i,\ldots,x_n)$ a realization of it. Then Copulas are functions that link the multivariate distribution $F(x_1,\ldots x_n)$ to its univariate marginals $F_{X_i}(x_i)$, thus they are often also called dependence functions. Sklar (1959) proved that every multivariate distribution $F(x_1,\ldots x_n)$ can be expressed in terms of a Copula C and its marginals $F_{X_i}(x_i)$:

$$F(x_1,\ldots x_n) = C(F_{X_1}(x_1),\ldots,F_{X_n}(x_n)) \tag{8}$$

$$C : [0,1]^n \rightarrow [0,1]. \tag{9}$$

In the bivariate case, which is made use of in this study, a Copula is defined as a function C from $[0,1]^2$ to $[0,1]$ so that $\forall u,v \in [0,1]$:

$$C(u,0) = 0 = C(0,v) \tag{10}$$

$$C(u,1) = u \quad \text{and} \quad C(1,v) = v \tag{11}$$

and that $\forall u_1,u_2,v_1,v_2 \in [0,1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$ holds

$$C(u_2,v_2) - C(u_2,v_1) - (C(u_1,v_2) - C(u_1,v_1)) \geq 0\,. \tag{12}$$

This means that each Copula has uniform margins and that each subset $[u_1,u_2] \times [v_1,v_2] \subseteq [0,1]^2$ is mapped to a positive number in $[0,1]$.

In turn of linking multivariate distributions to their marginals, Copulas allow to merge the dependence structure from the marginal distributions to form their joint multivariate distribution. The Copula function is unique when the marginals are steady functions. As the Copula is only a reflection of the dependence structure itself, their construction is reduced to the study of the relationship between the correlated variables, giving freedom for the choice of the univariate marginal distributions. Further information about Copulas can be found e.g. in Joe (1997); Frees and Valdez (1998); Nelsen (1999) and Salvadori et al. (2007).

The Copula approach allows to account for the fact that the dependence structure between two variates (X,Y) is more complex than it can be modelled by the multivariate normal distribution or ordinary dependence measures such as e.g. the Pearson correlation coefficient. Another important property of Copula functions is the fact that they are invariant under increasing monotonic transformations. In practice this means that data may be transformed (e.g. by taking the logarithm or detrending) without changing its Copula. As a consequence the empirical Copula $C_n(u,v)$, which is defined on the ranks of the data, is an estimator for the unknown theoretical Copula distribution $C_\theta(u,v)$ associated with the pair (X,Y), having a set of parameters θ:

$$C_n(u,v) = 1/n \sum_{i=1}^{n} 1\left(\frac{r_i}{n+1} \leqslant u, \frac{s_i}{n+1} \leqslant v\right) \tag{13}$$

HESSD
9, 937–982, 2012
Copula-based assimilation of radar and gauge information
S. Vogl et al.

where $(r_1,s_1),...,(r_n,s_n)$ denote the pairs of ranks of the data $(x_1,y_1),...,(x_n,y_n)$, and $1(...)$ is the indicator function.

The density of a theoretical Copula $C(u,v)$ is calculated as

$$c_\theta(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}. \tag{14}$$

A special family of Copula functions are the so called Archimedian Copulas. Some of the Copula functions discussed in this study are members of this family.

Let $\varphi : [0,1] \to [0,\infty]$ a steady, strict monotonic function with $\varphi(1) = 0$ and let $\varphi^{[-1]}$: $[0,\infty] \to [0,1]$ the pseudo-inverse of φ (Nelsen, 1999):

$$\varphi^{[-1]} := \begin{cases} \varphi^{-1}(t) & \text{if} \quad 0 \le t \le \varphi(0), \\ 0 & \text{else} \end{cases} \tag{15}$$

then the function

$$\begin{aligned} C &: [0,1]^2 \to [0,1] \\ &(u,v) \mapsto \varphi^{[-1]}(\varphi(u) + \varphi(v)) \end{aligned} \tag{16}$$

defines a Copula only if φ is convex and φ is called the generator of the Archimedian Copula C.

In this paper the following Archimedian Copulas are used: the Frank Copula with $\varphi(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$ and $\theta > 0$:

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right) \tag{17}$$

the Clayton Copula with $\varphi(t) = \frac{1}{\theta}\left(t^{-\theta}-1\right)$ and $\theta > 0$

$$C_\theta(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}} \tag{18}$$

HESSD
9, 937–982, 2012

and the Gumbel-Hougaard Copula with generator $\varphi(t) = (-\ln(t))^{\theta}$ and $\theta > 1$

$$C_\theta(u,v) = e^{-\left((-\ln(u)^{\theta})+(-\ln(v)^{\theta})\right)^{\frac{1}{\theta}}}. \tag{19}$$

In addition to the Archimedean Copulas the bivariate Gaussian Copula

$$C_\theta(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left(-\frac{s^2-2\theta st+t^2}{2(1-\theta^2)}\right) \mathrm{d}s\mathrm{d}t \tag{20}$$

where $\theta \in [-1,1]$, and Φ denotes the inverse of the univariate Gaussian distribution, and the Student-T Copula

$$C_\theta(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-\theta^2)^{\frac{1}{2}}} \left(1+\frac{s^2-2\theta st+t^2}{v(1-\theta^2)}\right)^{-\frac{v+2}{2}} \mathrm{d}s\mathrm{d}t\,, \tag{21}$$

where t_v^{-1} is the inverse of the univariate Student-t distribution with v degrees of freedom, are considered.

In practice, modelling random samples of realisations from the dependence structure of a bivariate dataset consists of several steps:

1. The positive pairs of the data (x_i, y_i) are transformed to the rank space (r_i, s_i) with $i = 1, \ldots, n$ denoting the length of the dataset.
2. The empirical Copula $C_n(u,v)$ is calculated using the ranks (r_i, s_i).
3. The Copula parameters are estimated using different types of theoretical (Archimedean) Copula functions.
4. Goodness-of-fit tests are carried out to choose the appropriate Copula family and its parameter
5. Realizations are generated from the modelled random variates (called pseudo-observations in the sequel).

HESSD
9, 937–982, 2012

3.2.1 Copula goodness-of-fit test

Goodness-of-fit tests for Copulas are applied comparing the empirical Copula $C_n(u,v)$ (Eq. 13) with the parametric estimate of a theoretical Copula model C_θ derived under the null hypothesis. There are different goodness-of-fit tests as e.g. reviewed by (Genest and Remillard, 2008; Genest et al., 2009). One of the tests used in this study is based on the Cramér-von Mises statistic (Genest and Favre, 2007):

$$S_n = n \sum_{i=1}^{n} \{C_\theta(u_i,v_i) - C_n(u_i,v_i)\}^2. \tag{22}$$

As the definition of S_n involves the theoretical Copula function, the distribution of the test statistic depends on the unknown value of θ under the null hypothesis that C is from the class C_θ (Grégoire et al., 2008).

Additionally the AIC and BIC information criteria (Eqs. 6 and 7) are applied and K-plots (e.g. Genest and Favre, 2007) are used for visual inspection.

3.2.2 Simulating from Copula distributions

With the calculated Copula dependence function and marginal distributions – that is $F_X(x)$, $F_Y(y)$ and $c_\theta(u,v)$ – conditional random samples are generated using the conditional probability density function.

The algorithm for simulating pseudo-observation from radar data is as follows:

1. Computation of $u = F_X(x)$, where x denotes one value of the observed radar rainfall and $F_X(x)$ is the marginal distribution of the variate X.

2. Generation of random samples for the variate v^* from the conditional PDF $c_{V|U}(v|u) = c_u(v)$ and calculation of $v = c_u^{-1}(v^*)$, where c_u^{-1} denotes the generalized inverse of c_u (Nelsen, 1999).

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

3. Calculation of the corresponding y-values using the probability integral transformation $F_Y^{-1}(v) = y$, with $F_Y(y)$ being the marginal distribution of the respective rain gauge.

3.3 Copula-based interpolation methods

Basis for the hereinafter presented Copula-based bias correction methods are the derived Copula maps. The Copula maps show the magnitude of dependence between radar and corresponding/surrounding gauge stations in terms of the Copula parameter θ for each grid in the domain. Thus the map is a reflection of the spatial dependence structure between radar and gauge data across the radar domain. These maps reflect anisotropies in the dependence structures between all gauges and the radar field. To assimilate the statistical chracteristics of all gauge stations at the same time, the Copula maps have to be combined in an appropriate way. Figure 3 schematically illustrates the different methods for one single time step. In any case the starting point is the ensemble of θ maps and of estimated marginal distributions. In the sequel the two different approaches are described in more detail.

3.3.1 Multiple Theta approach

The *Multiple Theta* approach consists of the following steps:

1. Based on the available gauge stations 31 Copula maps are derived, resulting in a set of 31 Copula parameters for a specific radar grid cell.
2. 31 Copula parameters, 31 marginal distributions of the gauge stations and one single marginal distribution of the specific radar grid cell are assigned to a specific radar grid cell.
3. For each set (marginal distribution gauge, marginal distribution radar, and Copula parameter) a sample of 100 members is simulated in the rank space.

HESSD
9, 937–982, 2012

4. The expectation values are calculated from the random samples.
5. The integral transformation is applied to the calculated expectation values to transform back to data space.
6. These steps are repeated for all radar grid cells.
7. Finally Inverse Distance Weighting (IDW) is applied to generate one single value for each radar grid cell.

3.3.2 Maximum Theta approach

The *Maximum Theta* approach comprises the following steps:

1. Based on the available gauge stations 31 Copula maps are derived, resulting in a set of 31 Copula parameters for a specific radar grid cell.
2. The set showing the maximum Copula parameter is assigned to a specific grid cell, retaining the information of the Copula parameter and the corresponding marginals (gauge and radar).
3. One sample of 100 members is simulated in the rank space for this set.
4. The expectation value is calculated from the random samples.
5. The integral transformation is applied to the expectation values to transform back to data space.
6. These steps are repeated for all radar grid cells.
7. A field containing the expectation values for each radar grid cell is obtained.

HESSD
9, 937–982, 2012

3.4 Validation measures

Once fields of pseudo-observations are generated the efficiency of the Copula based approach has to be tested. The sources that have to be compared are the original radar measurements, the informations from the different gauges in the domain and the simulated fields. In order to test the performance of the simulated field, Ordinary Kriging is applied to the gauge information exclusively to derive an interpolated field. Kriging uses the variogram of the regionalized variable (here precipitation), i.e. the variance between pairs of points that lie different distances apart. The best estimate of the values (BLUE – Best Linear Unbiased Estimator) is calculated considering the layout of the observation network relative to the interpolation grid. The major assumption of Ordinary Kriging is that the expected value of the regionalized variable is constant across the interpolated precipitation field. This is not the case for precipitation in the Alpine region. Nevertheless it is often used in a pragmatic way to obtain a first guess on the spatial distribution of rainfall.

Besides a purely visual inspection of the resulting fields a quantitative validation is done point-wise, using a cross-validation approach. Table 2 shows the different efficiency criteria used in this study, with o_i denoting the value of the observations and m_i the value of the model at time step $i = 1, \ldots, n$.

4 Results

To apply the algorithm described above, independent and identically distributed (*iid*) data is required. The autocorrelation function and the Box-Ljung test (Box et al., 1994) have been applied to test the positive pairs of the hourly time series for serial dependence and heteroskedasticity. Neither autocorrelation nor heteroskedasticity could be found (not shown).

HESSD
9, 937–982, 2012

4.1 Marginal distributions

As explained in Sect. 3.2.2 for the probability integral transformation the single marginal distributions are needed. The appropriate distribution function is estimated for all radar grids and all gauge stations separately. In general the algorithm presented in Sect. 3.3 allows to fit different theoretical marginal distribution functions. For both the radar and the gauge data four univariate distribution functions are considered, namely Normal, Exponential, Gamma and Weibull. The AIC/BIC values are calculated for each radar grid and all gauge stations. Table 3 lists exemplarily the results for the stations Garmisch-Partenkirchen, Oberammergau and Wielenbach, indicating that the Weibull distribution provides the best fit, followed by the Gamma and Exponential distribution. It is evident from this Table that neither radar nor gauge time series are normally distributed.

The empirical CDF (positive pairs only) as well as the fitted Weibull distribution for the station Garmisch-Partenkirchen are illustrated in Fig. 4. The gauge data corresponds perfectly to the fitted Weibull distribution, while for the radar data there is a slight deviation. For radar values $<5\,\mathrm{mm\,h^{-1}}$ ($>5\,\mathrm{mm\,h^{-1}}$), the theoretical distribution function underestimates (overestimates) the empirical distribution function. Even though a small number of grid cells/gauge stations show different characteristics, Table 3 and Fig. 4 corroborate an overall choice of the Weibull distribution function to ensure consistency throughout the study.

4.2 Fitting a theoretical Copula function

Once the univariate marginal distributions are fitted, the dependence structure between the time series has to be investigated. The first step is to calculate the empirical Copulas for radar-gauge pairs and then fit a theoretical bivariate Copula function. In this study five different theoretical Copula functions are tested, namely Gaussian, Student-T, Frank, Clayton and Gumbel-Hougaard. Only the positive pairs of radar and gauge are considered with 0.1 mm being the threshold for a rainy day.

HESSD

9, 937–982, 2012

Figure 5 shows the probability density functions of the empirical and the estimated theoretical Copula for radar-gauge pairs exemplarily for the station Garmisch-Partenkirchen. The empirical Copula density is asymmetric with respect to the opposite diagonal of the unit square. The density is highest for the upper right corner and is showing a second maximum in the lower left corner, indicating a strong upper and lower tail dependence.

For the five different theoretical Copula functions the goodness-of-fit test, which is based on the the Cramér-von Mises statistic, is applied (Genest and Remillard, 2008). 1000 values of the test statistic are sampled and the proportion of values larger than S_n are estimated by calculating the corresponding p-values. Additionally K-functions, Euler distances and p-values are used to identify appropriate Copula families. Based on the applied tests both Frank and Gumbel-Hougaard can be selected. However, the visual inspection of the empirical and both theoretical Copula density functions supports the decision for the Frank Copula.

The accuracy of the proposed algorithm is not reduced significantly by constraining the model to one specific theoretical Copula function. Nevertheless the corresponding Copula parameter is estimated individually.

4.3 Copula-based simulation of precipitation fields

In this section the results for the Copula-based simulation of precipitation fields are shown, starting with the estimation of the Copula parameter maps, followed by the description of the two different simulation algorithms (*Maximum Theta* and *Multiple Theta*). Derived precipitation fields are shown for one arbitrarily chosen time step and the obtained results are validated by visual inspection and different performance measures.

HESSD
9, 937–982, 2012

4.3.1 Copula parameter maps

As described in Sect. 3.3 the prerequisite of the Copula-based precipitation simulation is a precipitation field derived from radar reflectivities.

The radar field, derived by the standard Z/R relationship, is disturbed showing spokes caused by obstacles shading the radar beam (see Fig. 2). Besides these artifacts, effects of backscatter may also be present. Nevertheless, the measured radar field gives a realistic impression of the spatial distribution of precipitation at that certain time step. Even if the precipitation patterns are realistically displayed, the radar does not accurately reproduce the absolute precipitation amounts on the ground. In order to correct the radar field in terms of the absolute precipitation amounts, ground based measurements are assimilated. In this study 31 gauges are located in the domain of the radar field (see Fig. 1). Following the new algorithm proposed in Sect. 3.3, the parameter of the Frank Copula is estimated between one selected gauge and all grid cells of the radar field, showing the strength of the dependence between the two time series. All Copula parameters θ are visualized together building the so called *Theta maps* corresponding to the respective station. Figure 6 shows the Theta maps for station Garmisch-Partenkirchen (top) and Wielenbach (bottom). In general, the correlation between radar and gauge time series decreases with increasing distance. For station Wielenbach e.g. the Copula parameter map is almost radially symmetric. However, considering all available gauge stations in the domain, there are also cases where the dependence structure is highly anisotropic such as Garmisch-Partenkirchen (see Fig. 6).

Following the algorithm proposed in Sect. 3.3, all 31 available Copula parameter maps are combined in the first step. This is done by assigning the maximum value out of the set of all available Copula parameters to each grid cell.

Figure 7 shows the Maximum Theta map for the whole domain. The asymmetries from the single station maps are reflected by the field, making sure that all possible variabilities in the dependence structures are considered.

HESSD
9, 937–982, 2012

The anisotropic nature of the dependence structure has to be considered when Copula-based precipitation fields are modelled. The results of two different approaches namely *Multiple Theta* and *Maximum Theta* are presented in the sequel.

4.3.2 Multiple Theta

As described in Sect. 3.3 the single Copula maps based on the radar-gauge pairs are combined to simulate a bias-corrected precipitation field.

A precipitation field generated using the *Multiple Theta* approach is shown in Fig. 8. This field is based on 31×100 realisations for each grid cell. The pattern information from the radar precipitation field is preserved, and the absolute values are corrected towards the ground measurements. For the chosen time step the rainfall intensities and variances are reduced. Spokes and backscattering effects could be diminished compared to the original radar image (see Fig. 2).

4.3.3 Maximum Theta

Based on the *Maximum Theta* map 100 random realisations (rank space) are generated for each grid cell conditioned on the respective radar value. The resulting field is shown in Fig. 9. Compared to the results shown for the *Multiple Theta* approach the patterns from the radar are reproduced similarly well. Even if the simulated field is reduced in terms of variability compared to the radar field, more details are retained than for the *Multiple Theta* method. The absolute values are slightly higher.

4.4 Validation of the simulated precipitation fields

4.4.1 Visual inspection

Figure 10 shows an interpolated rainfall field derived from the observed precipitation values on 14 July 2008, (13:00). The observations from the 31 gauge stations in the radar domain are interpolated by application of an Ordinary Kriging approach (e.g. de Marsily, 1986; Kitanidis, 1997).

Albeit the major assumption of Ordinary Kriging is not valid throughout the whole research domain, i.e. the expectation value is not constant over space, the interpolated precipitation field (Fig. 10) roughly reproduces the rainfall patterns revealed by the radar measurement (compare to Fig. 2). The fine structures of the rainband are remarkably smoothed and the rainfall is concentrated in three local maxima. As gauge stations are only available for the center of the domain (see Fig. 1) there is no information included for the outer regions. Consequently, the interpolated field can not reproduce the variability of precipitation in that area. As the Copula-based simulated precipitation fields (see Figs. 8 and 9) not only include information from the rain gauges but also incorporate information from the radar measurement, they resolve better the fine structures in the overall pattern of precipitation.

Figure A1 shows the accumulated summer precipitation obtained by radar at Hohenpeissenberg derived by a standard Z/R-relationship, the RADOLAN correction of DWD, and simulated by following the *Maximum Theta*, and *Multiple Theta* approach for the years 2005–2008. Compared to the rain gauges, the radar observes the liquid and solid water in a certain volume of the atmosphere, which only partly reaches the ground as precipitation. Therefore, the total amount of precipitation is significantly different from that seen by the radar. In general, this leads to overestimation of the cumulated amount of precipitation by radar. This effect is supposed to be stronger in wet summers such as 2006 and 2007. For these seasons, the Copula-based methodologies *Maximum Theta* and *Multiple Theta* reduce significantly the total amount of rainfall, while for dry summers the spatial representation of annual amounts is similar for all approaches. While errors due to shading effects are accumulated over time in the radar (standard Z/R relationship) and the RADOLAN corrected fields, these artifacts are reduced by the Copula-based bias corrections.

HESSD
9, 937–982, 2012

4.4.2 Quantitative validation

The results of the *Maximum Theta* and *Multiple Theta* approaches are checked by cross-validation for all 31 stations, using the time series of gauge and pseudo-observations in the respective grid cell. Different validation measures are used to obtain a quantitative appraisal for the performance of the simulation algorithms (see Table A1). The mean correlation between the simulated pseudo-observations and the observed precipitation is 0.7 for both proposed methods indicating equally reasonable results. However, the results for RMSE, MAE and MSE show that the *Multiple Theta* approach is slightly superior. This finding is also supported by the inspection of the respective NSE values, being 0.31 for the *Maximum Theta* and 0.21 for the *Multiple Theta* method. Note that the validation results for station Hohenpeissenberg are not included to the mean values as this station is located close to the radar observatory. For this grid cell the radar precipitation is significantly biased, resulting in a very low correlation, negative NSE and the highest values for RMSE, MAE and MSE among all gauge stations. It is also important to appreciate how qualified the Copula-based bias-correction is compared to operational correction methods. Therefore precipitation time series corrected with RADOLAN are used for comparison (Bartels et al., 2004).

Table 4 shows the results of the comparison of radar precipitation derived with the standard Z/R-relationship and the DWD RADOLAN-correction, pseudo-observations simulated with the *Maximum Theta* approach and gauge observations for the three arbitrarily chosen stations Garmisch-Partenkirchen, Oberammergau and Wielenbach (positive pairs, summer 2005–2009). The correlations do not differ significantly between the three methods. However, the RMSE and the NSE indicate that the Copula-based approach performs best. Visual inspection of the respective time series shows, that high precipitation values adhere to the highest biases. These are most effectively corrected by the Copula approach which is also confirmed by the calculated NSE. Overall, the performance of the *Maximum Theta* method is found to be slightly better than the RADOLAN correction.

HESSD
9, 937–982, 2012

5 Discussion

Radar precipitation fields are a good supplement to gauge measurements as they provide a reasonable representation of spatial rainfall distribution. Nevertheless, precipitation fields from radar reflectivities incorporate significant errors due to the measurement itself (backscatter, shading etc.), the used Z/R-relationship or the fact that they do not provide values at the ground level. There are sophisticated operational bias-correction algorithms for radar products such as the RADOLAN, which contains orographic shading correction, statistical reduction of clutter, gradient smoothing and further preprocessing steps and assimilates data from ombrometer measurements. Here a purely statistical Copula-based approach to assimilate gauge information for bias-correction of radar precipitation fields is proposed.

Recently Laux et al. (2011) showed, that daily precipitation data incorporates autocorrelation and heteroskedasticity and therefore is not *iid*. They showed that an ARMA-GARCH algorithm can be used to transform the time series so that Copula-based methods can be applied.

A thorough investigation of the hourly radar and gauge precipitation used in this study showed, that the time series are intrinsically *iid* and no such transformation is necessary.

As shown by Villarini et al. (2008) it is possible to model radar-rainfall uncertainties using the joint Copula distribution of radar and gauge pairs. However, this study only considered data from certain grids where both radar and gauge measurements were available. Recently AghaKouchak et al. (2010a,b) presented a Copula-based method to simulate radar error fields using a network of gauge stations as reference. They used different multivariate Copula distributions such as Gaussian, t-Copula and their V-transform Copula, and showed that the multivariate dependence structure is not represented well by the Gaussian distribution. They generated rainfall ensembles with similar characteristics as the ground reference even for a very sparse observation network. Compared to the approach investigated in this study, it is computationally

HESSD
9, 937–982, 2012

extensive to estimate all parameters for a multivariate model and statistical tests have to be carried out to ensure reliability of the estimated parameter set and stability of the estimation process.

The method presented in this study tends to apply a pairwise Copula analysis and incorporates anisotropies in the dependence structure by combination of the derived Copula maps.

In general the correlation between gauge and radar decreases with increasing distance between gauge and radar grid cell. Besides, the Copula maps shows more details of the dependence structure. For certain gauge stations (see Fig. 6) the θ-field is strongly asymmetric, revealing the influence of orography on the dependence structure. This asymmetry is disregarded by standard interpolation methods such as inverse distance weighting although it is getting more important in complex terrain.

Asymmetries in the dependence structure can also be an indicator of air flow directions dominant for a certain location. This theory could be supported by investigation of different time scales.

For small time scales (<1 h), localized rainfall events are not resolved by the statistical analysis. In that case, the region of strongly correlated grid cells, visualized through high Copula parameters in the Copula map, is expected to be reduced compared to larger time scales. For different seasons different types of rainfall regimes are predominant. Therefore, the investigation of e.g. summer and winter season separately is expected to reveal preferential precipitation types, differentiating large scale winter precipitation from convective summer events.

For the results presented in this study one single type of marginal distribution and one single theoretical Copula function is used for all stations and all radar grid cells. However, the method allows for estimating different marginal distributions for each time series as well as different theoretical Copula functions for each radar/gauge pair. In the course of this study simulations for both the Frank and the Gumbel-Hougaard Copula are calculated as suggested by the goodness-of-fit tests. A comparison of the resulting fields shows only small differences, legitimating the restriction to only one theoretical

HESSD
9, 937–982, 2012
Copula-based assimilation of radar and gauge information
S. Vogl et al.

Copula function. Before the integral transformation is applied to the simulated fields, the marginals do not contribute to the simulation results. Therefore the performance of the Copula model itself can be assessed in the rank space. It is found that the correlation between the pseudo-observations and gauges is significantly higher in that case. Hence, the choice of the marginals is of crucial importance for the overall model results. An improvement of the results by allowing for different distribution functions for each gauge and radar grid cell can be expected.

Finally it is shown that even for a simplified setup the proposed bias-correction methods are able to combine the advantages of the two data sources: the rainfall pattern observed by the radar measurement are retained in the simulated field while the absolute values are successfully corrected towards the gauge observations.

The *Maximum Theta* apprach is found to be slightly better than the *Multiple Theta* method. This is due to the fact that for each radar grid cell the simulation is based on the highest θ and the respective marginal distribution while for the *Multiple Theta* case stations with low correaltion also slightly contribute to the simulated results.

As all ingredients for the proposed algorithms, namely the marginal distributions and the theoretical Copula functions, only have to be estimated once the proposed bias-correction is computationally not very demanding which facilitates operational application of the proposed methods in quasi real time. The effectiveness of the *Maximum Theta* approach is also emphasized by the results of the comparison with the RADOLAN correction. The correlations between corrected rainfall and gauge measurements are nearly the same for the two approaches but RMSE and NSE indicate the performance of the Copula based approach to be slightly better.

Recently, additionally to radar and gauge precipitation data there are also some remote sensing techniques such as dual microwave links or wireless communication networks (e.g. Messer et al., 2006) from which rainfall fields can be derived. Even though large uncertainties are associated with their application to derive precipitation information, these approaches are very promising due to the very dense network of observations.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

6 Conclusions

Two new Copula-based methods are proposed to bias-correct radar precipitation fields by assimilation of gauge information. It is found that the data is intrinsically *iid* and no transformation is necessary. For each gauge station and all radar grid cells a Copula map is derived revealing anisotropies in the dependence structure. Asymmetries in the Copula maps reflect the complexity of the terrain as well as predominant flow directions and rainfall types.

It is possible to constrain the Copula model to one theoretical Copula function while the choice of the marginals is of crucial importance for the quality of the resulting fields. Thus the simulation results could be further improved by allowing for different types of marginal distributions. The Copula model is suitable to merge the advantages of the different data sources: the spatial distribution of the radar rainfall field is preserved while absolute values are corrected towards gauge observations. As a consequence, they can be used as an alternative method to bias-correct radar precipiation fields. Considering the performance measures applied in this study, the *Maximum Theta* approach is modelling the gauge values slightly better than the bias-correction method RADOLAN, used as standard procedure by DWD.

Acknowledgements. This research was funded by the Bavarian State Ministry of the Environment and Public Health (reference number VH-ID:32722/TUF01UF-32722), by the Federal Ministry of Education and Research as part of the research project LUCCi (reference number 01LL0908C), HGF (research project: Terrestrial Environmental Observatories TERENO – http://teodoor.icg.kfa-juelich.de/overview-de and by the HGF Impuls und Vernetzungsfond (research project: Regional Precipitation Observation by Cellular Network Microwave Attenuation and Application to Water Resources Management, PROCEMA, Virtual Institute, reference number VH-VI-314). The help of the German Weather Service (DWD) which provided the rain gauge data and the radar data is highly appreciated. The authors also thank Jörg Seltmann from DWD, Meteorological Observatory Hohenpeissenberg (MOHP), for technical support concerning the radar data.

HESSD
9, 937–982, 2012
Copula-based assimilation of radar and gauge information
S. Vogl et al.

References

AghaKouchak, A., Bárdossy, A., and Habib, H.: Copula-based uncertainty modeling: Application to multi-sensor precipitation estimates, Hydrol. Process., 24, 2111–2124, doi:10.1002/hyp.7632, 2010a. 940, 959

AghaKouchak, A., Bárdossy, A., and Habib, H.: Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula, Adv. Water Resour., 33, 624–634, doi:10.1016/j.advwatres.2010.02.010, 2010b. 940, 959

AghaKouchak, A., Habib, E., and Bárdossy, A.: Modeling radar rainfall estimation uncertainties: random error model, J. Hydrol. Eng., 15, 265–275, 2010c. 939

Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res., 42, W11416, doi:10.1029/2005WR004754, 2006. 940, 941

Bárdossy, A. and Li, J.: Geostatistical interpolation using copulas, Water Resour. Res., 44, W07412, doi:10.1029/2007WR006115, 2008. 940, 941

Bárdossy, A. and Pegram, G. G. S.: Copula based multisite model for daily precipitation simulation, Hydrol. Earth Syst. Sci., 13, 2299–2314, doi:10.5194/hess-13-2299-2009, 2009. 941

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N.: Projekt RADOLAN, Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Abschlussbericht, Deutscher Wetterdienst, 2004. 942, 958

Bollerslev, T.: Generalized autoregressive conditional heteroscedasticity, J. Econometr., 31, 307–327, doi:10.1016/0304-4076(86)90063-1, 1986. 939

Box, G. E., Jenkins, G. M., and Reinsel, G. C.: Time Series Analysis: Forecasting and Control, 3rd Edn., Prentice Hall, Englewood Cliffs, New Jersey, 1994. 952

Burrough, P. A. and McDonnell, R. A.: Principles of Geographical Information Systems, Oxford University Press, Oxford, 1998. 939

Brandes, E. A.: Optimizing rainfall estimates with the aid of radar, J. Appl. Meteorol., 14, 1339–1345, 1975. 939

Cole, S. J. and Moore, J.: Hydrological modelling using raingauge and radar based estimators of areal rainfall, J. Hydrol., 358, 159–181, 2008. 939

Collins, F. B. and Bolstad, P. V.: A comparison of spatial interpolation techniques in temperature estimation, Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, NM, Santa Barbara, CA: National Center for Geographic

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

Information and Analysis, Santa Barbara, 1996. 939
de Marsily, G.: Quantitative Hydrogeology: Groundwater Hydrology for Engineers, Academic Press, Inc., New York, 1986. 957
De Michele, C. and Salvadori, J.: A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas, J. Geophy. Res.-Atmos., 108, 1–11, doi:10.1029/2002JD002534, 2003. 940
Dupuis, D.: Using copulas in hydrology: Benefits, cautions, and issues, J. Hydrol. Eng., 12, 381–393, 2007. 940
Ehret, U.: Rainfall and Flood Nowcasting in Small Catchments using Weather Radar, PhD dissertation, Stuttgart University, ISBN 3-933761-24-7, 2003. 939
Embrechts, P., McNeil, A. J., and Straumann, D.: Correlation and Dependency in Risk Management: Properties and Pitfalls, Cambridge University Press, New York, 2001. 940
Erxleben, J., Elder, K., and Davis, R.: Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains, Hydrol. Process., 16, 3627–3649, 2002. 939
Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., and Bobee, B.: Multivariate hydrological frequency analysis using copulas, Water Resour. Res., 40, 112, 2004. 940
Frees, E. W. and Valdez, E. A.: Understanding relationships using copulas, N. Am. Actuarial J., 2, 1–25, 1998. 940, 946
Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng., 12, 347–368, 2007. 940, 949
Genest, C. and Remillard, B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, Ann. I. H. Poincare B, 44, 1096–1127, 2008. 949, 954
Genest, C., Remillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: a review and a power study, Insur. Math. Econ., 44, 199–214, 2009. 949
Gomez-Hernandez, J. and Wen, X.: To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology, Adv. Water Resour., 21, 47–61, 1998. 940
Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113–129, 2000. 939
Grégoire, V., Genest, C., and Gendron, M.: Using copulas to model price dependence in energy markets, Energy Risk, 5, 58–64, 2008. 949
Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, J. Hydrol., 332, 144–157, 2007. 939

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

Hutchinson, M. F.: Interpolation of rainfall data with thin plate smoothing splines: I two dimensional smoothing of data with short range correlation, J. Geogr. Inf. Dec. Anal., 2, 139–151, 1998a. 939

Hutchinson, M. F.: Interpolation of rainfall data with thin plate smoothing splines: II analysis of topographic dependence, J. Geogr. Inf. Dec. Anal., 2,152–167, 1998b. 939

Isaaks, E. H. and Srivastava, R. M.: Applied Geostatistics, Oxford University Press, New York, 1989. 938, 939

Joe, H.: Multivariate Models and Dependence Concepts, Chapman and Hall, New York, 1997. 946

Joss, J. and Lee, R.: The application of radar-gauge comparisons to operational precipitation profile corrections, J. Appl. Meteorol., 34, 2612–2630, 1995. 939

Kitanidis, P. K.: Introduction to Geostatistics, Cambridge University Press, Cambridge, 1997. 957

Krajewski, W. F.: Cokriging radar-rainfall and rain-gauge data, J. Geophys. Res.-Atmos., 92, 9571–9580, 1987. 939

Krajewski, W. F. and Smith, J.: Radar hydrology: rainfall estimation, J. Hydrol., 25, 1387–1394, 2002.

Krause, P., Boyle, D. P., and Bäse, F.: Comparison of different efficiency criteria for hydrological model assessment, Adv. Geosci., 5, 89–97, doi:10.5194/adgeo-5-89-2005, 2005.

Kuhn, G., Khan, S., and Ganguly, A.: Geospatial-temporal dependence among weekly precipitation extremes with applications to observations and climate model simulations in South America, J. Adv. Water Resour., 30, 2401–2423, 2007. 940

Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H.: Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, Hydrol. Earth Syst. Sci., 15, 2401–2419, doi:10.5194/hess-15-2401-2011, 2011. 945, 959

Mandapaka, P. V., Krajewski, W. J., Ciach, G. J., Villarini, G., and Smith, J. A.: Estimation of radar-rainfall error spatial correlation, Adv. Water Resour., 32, 1020–1030, 2009. 939

Marx, A.: Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, PhD dissertation, Stuttgart University, ISBN 3-933761-64-6, 2007. 943

Mazzetti, C. and Todini, E.: Combining rain-gauges and radar precipitation measurements using a bayesian approach, Quant. Geol. G., 13, 401–412, doi:10.1007/1-4020-2115-1_34, 2004. 939

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

Messer, H., Zinevich, A., and Alpert, P.: Environmental monitoring by wireless communication networks, Science, 312, 713, 2006. 961

Nelsen, R. B.: An Introduction to Copulas, Springer-Verlag, New York, 1999. 946

Renard, B. and Lang, M.: Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, Adv. Water Resour., 30, 897–912, 2007. 940

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: Extremes in Nature: an Approach Using Copulas, Springer-Verlag, New York, 2007. 946

Schölzel, C. and Friederichs, P.: Multivariate non-normally distributed random variables in climate research – introduction to the copula approach, Nonlin. Processes Geophys., 15, 761–772, doi:10.5194/npg-15-761-2008, 2008. 940

Serinaldi, F.: Analysis of inter-gauge dependence by kendall's τ, upper tail dependence coefficient, and 2-copulas with application to rainfall fields, Stoch. Env. Res. Risk A., 22, 671–688, 2008. 940

Singh, V. P.: Effects of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph, Hydrol. Process., 11, 1649–1669, 1997. 939

Sklar, K.: Fonctions de repartition à n dimensions et leura marges, Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231, 1959.

Villarini, G., Serinaldi, F., and Krajewski, W. F.: Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches, Adv. Water Resour., 31, 1674–1686, 2008. 940, 959

Wang, W., Van Gelder, P. H. A. J. M., Vrijling, J. K., and Ma, J.: Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes, Nonlin. Processes Geophys., 12, 55–66, doi:10.5194/npg-12-55-2005, 2005.

Zhang, L. S. R. and Singh, V.: Bivariate rainfall frequency distributions using archimedean copulas, J. Hydrol., 32, 93–109, 2008. 940

HESSD

9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

Table 1. Geographical location of gauges and the rainfall radar as well as statistical measures for gauge and radar (positive pairs only) for period June, July, and August of 2006 and 2007.

ID	Station name	Altitude (m)	Lat (°)	Lon (°)	$\text{Gauge}_{\text{mean}}$ (mm h^{-1})	$\text{Gauge}_{\text{std}}$ (mm h^{-1})	$\text{Radar}_{\text{mean}}$ (mm h^{-1})	$\text{Radar}_{\text{std}}$ (mm h^{-1})
1	Bernbeuren-Prachtsried	936	47.74	10.75	1.76	2.27	1.33	2.23
2	Diessen	658	47.96	11.01	1.51	1.96	1.15	1.74
3	Deisenhofen	585	48.04	11.58	1.62	2.12	1.25	1.60
4	Ettal	940	47.57	10.96	1.59	1.85	1.36	1.88
5	Garmisch-Partenkirchen	719	47.48	11.06	1.61	2.02	1.48	2.13
6	Gilching	550	48.11	11.28	1.56	2.25	1.25	1.75
7	Griesen	801	47.48	10.95	1.48	1.75	1.49	2.13
8	Halblech	780	47.65	10.81	1.68	2.27	1.15	1.42
9	Hindelang	1015	47.46	10.43	1.75	2.28	1.59	2.67
10	Hohenpeissenberg	977	47.80	11.01	5.03	4.10	0.16	0.01
11	Kaufbeuren	716	47.87	10.60	1.66	2.16	1.26	1.93
12	Kochel	805	47.57	11.30	1.51	1.87	1.22	1.60
13	Kohlgrub, Bad	740	47.67	11.08	1.82	2.27	1.25	1.91
14	Kraftisried	831	47.77	10.46	1.70	2.25	1.30	2.12
15	Kreuth	895	47.61	11.65	1.72	2.13	1.43	2.11
16	Krün	873	47.50	11.28	1.71	2.18	1.46	1.97
17	Lenggries	737	47.59	11.55	1.73	2.20	1.55	2.40
18	Maisach	530	48.21	11.20	1.55	2.04	1.19	1.69
19	Marktoberdorf	790	47.72	10.64	1.77	2.30	1.32	2.08
20	München	515	48.16	11.54	1.53	2.21	1.43	1.86
21	Oberammergau	835	47.60	11.06	1.58	2.11	1.10	1.88
22	Oberschleissheim	484	48.24	11.55	1.55	2.13	1.24	1.81
23	Oy	885	47.64	10.39	1.84	2.14	1.52	2.09
24	Schwangau	796	47.58	10.72	1.63	2.13	1.47	2.04
25	Seeg	802	47.67	10.63	1.67	2.19	1.36	1.95
26	Schäftlarn	557	47.98	11.47	1.58	2.01	1.24	1.59
27	Steingaden	761	47.76	10.86	1.72	2.18	1.16	1.83
28	Schwaben	538	48.20	10.73	1.54	1.91	1.18	1.67
29	Schlehdorf	609	47.66	11.32	1.76	2.12	1.41	1.82
30	Vilgertshofen	685	47.97	10.92	1.66	2.16	1.21	1.71
31	Wielenbach	550	47.88	11.16	1.45	1.89	1.10	1.44

HESSD
9, 937–982, 2012

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

Table 2. Validation measures used in this study.

Validation measure	Abbreviation	Formula	Range	Perfect Fit
Pearson Correlation	r	$r = \frac{\sum_{i=1}^{n}(o_i-\bar{o}_i)(m_i-\bar{m}_i)}{\sqrt{\sum_{i=1}^{n}(o_i-\bar{o}_i)^2}\sqrt{\sum_{i=1}^{n}(m_i-\bar{m}_i)^2}}$	$[-1,1]$	$\lvert r\rvert = 1$
Root Mean Square Error	RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i-m_i)^2}$	$[0,\infty[$	RMSE = 0
Nash Sutcliffe Efficiency	NSE	$1-\frac{\sum_{i=1}^{n}(o_i-m_i)^2}{\sum_{i=1}^{n}(o_i-\bar{o}_i)^2}$	$]-\infty,1]$	NSE = 1

Table 3. AIC and BIC of radar and gauge (positive pairs only) for selected rain gauge stations and different univariate distribution functions (June, July, and August of 2006 and 2007). Smallest values of AIC/BIC indicate the best fit.

Station		Normal	Exponential	Gamma	Weibull
Garmisch-Partenkirchen	AIC (Radar)	2620	1678	1676	1662
	BIC (Radar)	2631	1684	1687	1673
	AIC (Gauge)	2556	1779	1774	1766
	BIC (Gauge)	2567	1785	1784	1777
Oberammergau	AIC (Radar)	2493	1516	1514	1498
	BIC (Radar)	2504	1525	1522	1508
	AIC (Gauge)	2634	1776	1767	1754
	BIC (Gauge)	2644	1781	1778	1765
Wielenbach	AIC (Radar)	1806	1111	1112	1108
	BIC (Radar)	1817	1116	1122	1112
	AIC (Gauge)	2081	1390	1388	1380
	BIC (Gauge)	2092	1400	1398	1389

HESSD
9, 937–982, 2012

HESSD
9, 937–982, 2012

Table 4. Comparison of radar precipitation derived with the standard Z/R-relationship, the DWD RADOLAN-correction, pseudo-observations simulated with the *Maximum Theta* approach to gauge observations for Garmisch-Partenkirchen, Oberammergau and Wielenbach (positive pairs, summer 2005–2009), best performance highlighted in bold.

	Garmisch-P.			Oberammergau			Wielenbach		
Methods	r	RMSE (mm h^{-1})	NSE	r	RMSE (mm h^{-1})	NSE	r	RMSE (mm h^{-1})	NSE
Standard Z/R	0.64	1.74	0.20	0.60	1.82	0.19	0.53	1.72	0.201
RADOLAN	**0.66**	1.66	0.23	0.62	1.71	0.28	**0.58**	1.64	0.27
Copula based	0.65	**1.50**	**0.41**	**0.63**	**1.60**	**0.37**	0.57	**1.59**	**0.31**

Table A1. Validation of the *Maximum Theta* and *Multiple Theta* approach for the 31 stations (summer 2008, wet days only).

ID	Station Name	Multiple Theta r	Multiple Theta RMSE ($mm\,h^{-1}$)	Multiple Theta NSE	Maximum Theta r	Maximum Theta RMSE ($mm\,h^{-1}$)	Maximum Theta NSE
1	Bernbeuren-Prachtsried	0.80	1.69	0.24	0.78	1.60	0.33
2	Diessen	0.76	1.73	0.23	0.78	1.62	0.32
3	Deisenhofen	0.74	1.92	0.19	0.74	1.73	0.34
4	Ettal	0.73	1.75	0.24	0.73	1.60	0.37
5	Garmisch-Partenkirchen	0.66	1.81	0.25	0.66	1.72	0.33
6	Gilching	0.68	1.94	0.20	0.69	1.86	0.27
7	Griesen	0.65	1.62	0.25	0.67	1.54	0.33
8	Halblech	0.70	1.98	0.20	0.70	1.88	0.27
9	Hindelang	0.62	2.12	0.15	0.61	1.98	0.26
10	Hohenpeissenberg	0.47	6.12	-0.97	0.65	6.03	-0.91
11	Kaufbeuren	0.77	1.36	0.27	0.73	1.17	0.46
12	Kochel	0.76	1.60	0.29	0.77	1.53	0.35
13	Kohlgrub, Bad	0.77	1.87	0.22	0.77	1.71	0.35
14	Kraftisried	0.67	1.97	0.17	0.68	1.90	0.24
15	Kreuth	0.66	2.10	0.15	0.67	1.91	0.30
16	Krün	0.62	2.13	0.21	0.62	2.02	0.29
17	Lenggries	0.69	1.94	0.22	0.70	1.69	0.41
18	Maisach	0.67	1.75	0.19	0.68	1.66	0.27
19	Marktoberdorf	0.79	1.93	0.24	0.79	1.78	0.35
20	München	0.74	1.55	0.26	0.71	1.48	0.33
21	Oberammergau	0.76	1.69	0.27	0.76	1.59	0.36
22	Oberschleissheim	0.69	2.01	0.19	0.68	1.85	0.31
23	Oy	0.65	1.77	0.21	0.64	1.59	0.36
24	Schwangau	0.69	1.63	0.24	0.68	1.53	0.34
25	Seeg	0.70	2.00	0.20	0.70	1.91	0.27
26	Schäftlarn	0.79	1.66	0.23	0.79	1.55	0.33
27	Steingaden	0.74	2.22	0.20	0.75	2.01	0.34
28	Schwaben	0.63	1.98	0.15	0.63	1.91	0.20
29	Schlehdorf	0.75	2.08	0.21	0.76	1.94	0.31
30	Vilgertshofen	0.70	1.96	0.20	0.71	1.87	0.27
31	Wielenbach	0.74	1.71	0.26	0.74	1.68	0.29

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

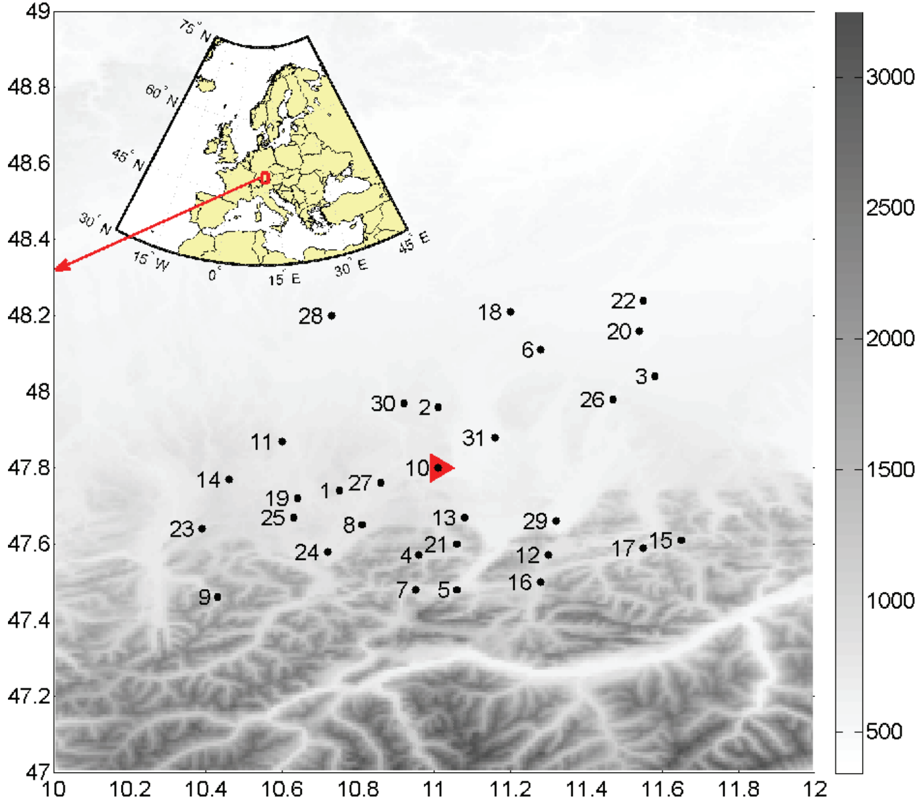

Fig. 1. Research area showing the position of the gauges and the weather radar on Mount Hohenpeissenberg (red triangle). The names of the gauge stations (black dots) can be found in Table 1.

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.

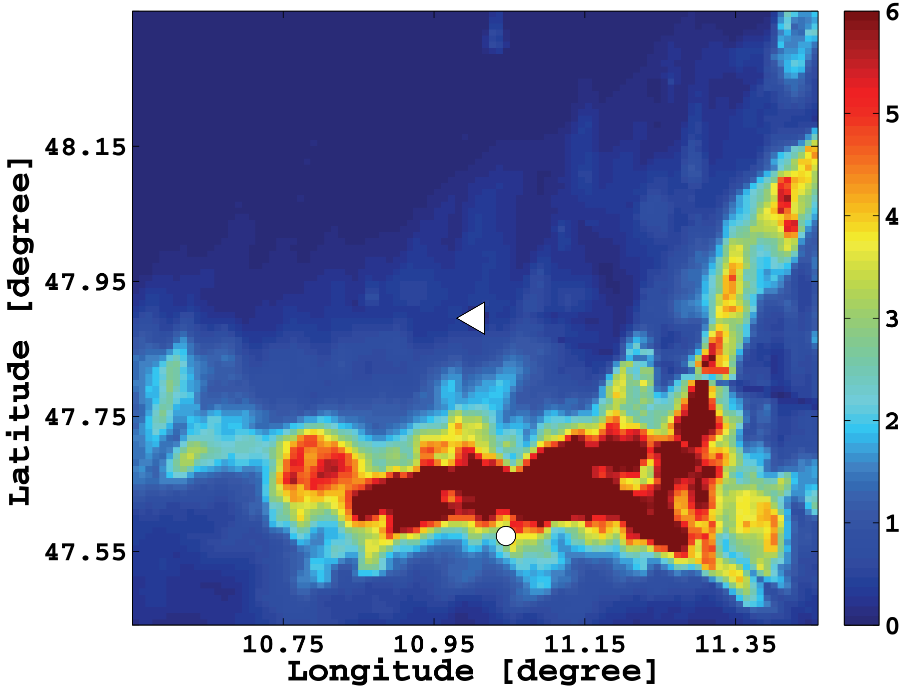

Fig. 2. Precipitation field (mm h^{-1}) derived from radar measurements (standard Z/R relationship) at Mount Hohenpeissenberg on 14 July 2008 (13:00). The position of Hohenpeissenberg (Garmisch-Partenkirchen) is indicated by a white triangle (circle).

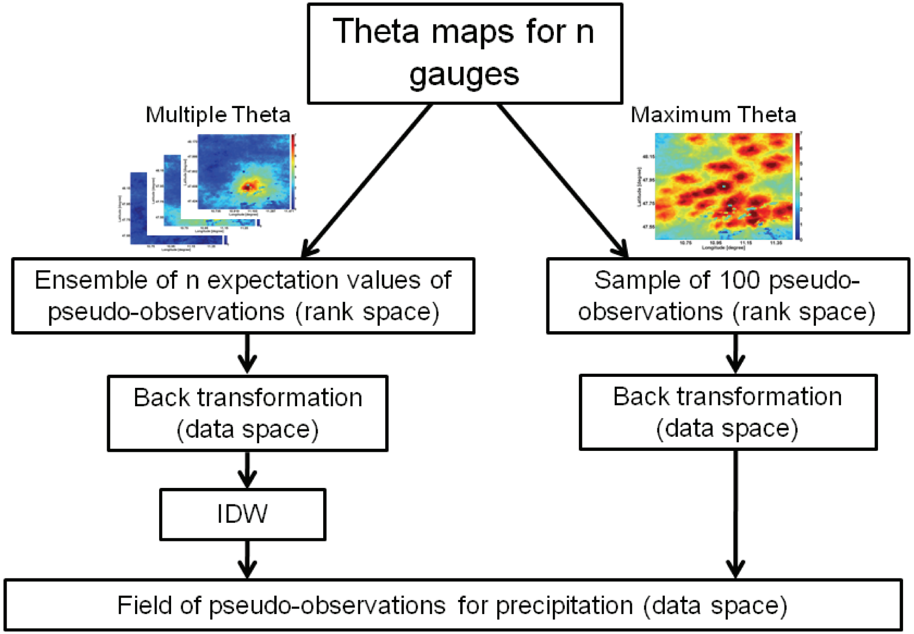

Fig. 3. Overview over the Copula based interpolation methods.

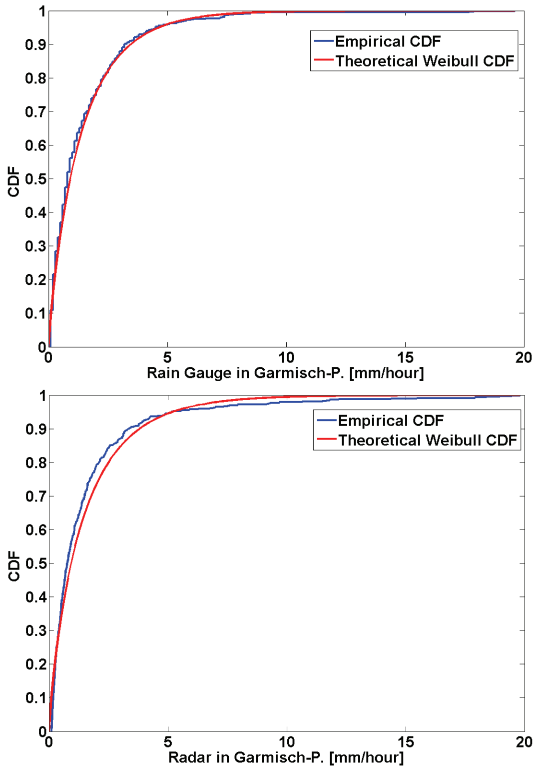

Fig. 4. Empirical and estimated CDF for the station Garmisch-Partenkirchen (top) and for the corresponding radar grid (bottom).

HESSD
9, 937–982, 2012

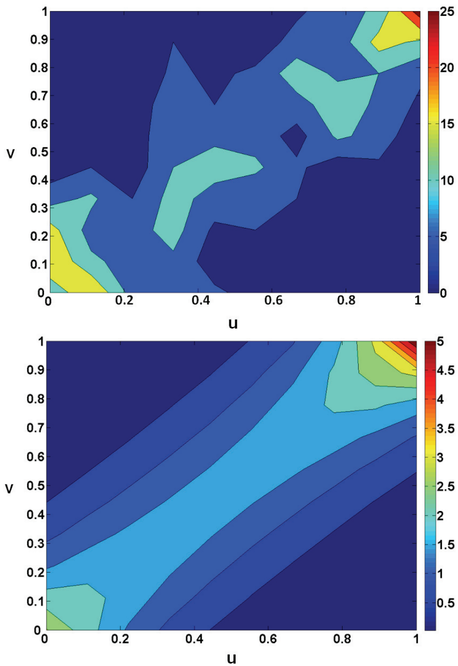

Fig. 5. Probability density function of the empirical Copula (top) and of the fitted Frank Copula (bottom) for the station Garmisch-Partenkirchen.

HESSD

9, 937–982, 2012

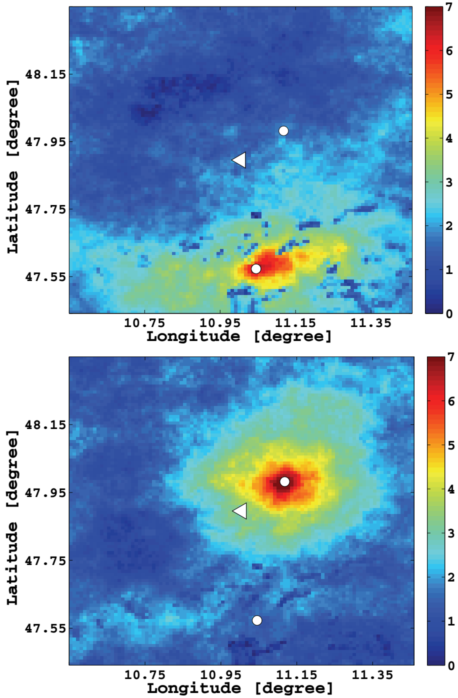

Fig. 6. Copula parameter maps for stations Garmisch-Partenkirchen (top) and Wielenbach (bottom) showing the parameter θ of the Frank Copula. The position of Hohenpeissenberg is indicated by a white triangle, Garmisch-Partenkirchen (in the South) and Wielenbach (in the North) by a white circle.

HESSD
9, 937–982, 2012

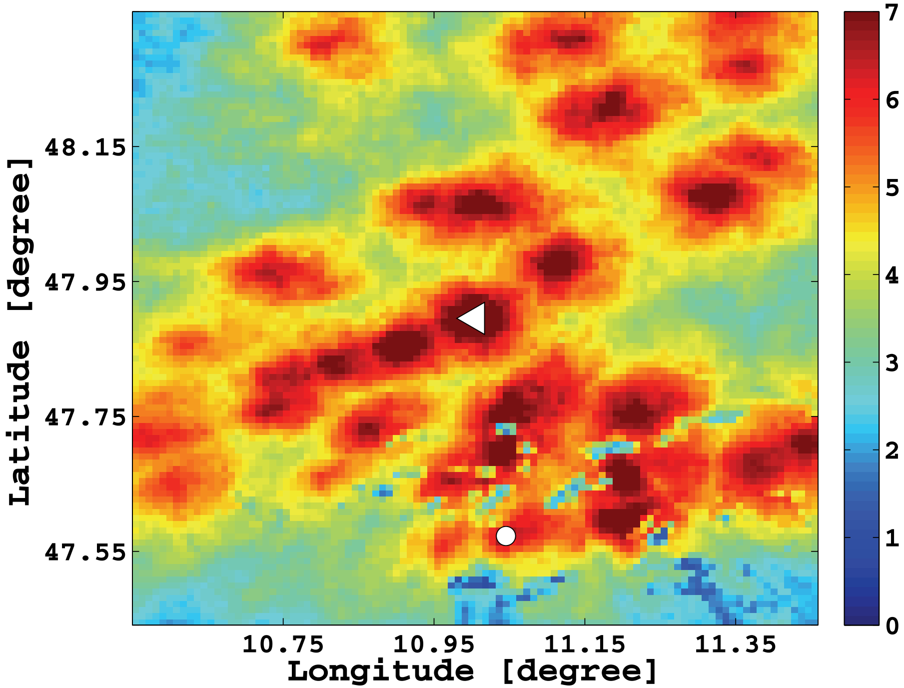

Fig. 7. Maximum Theta map derived from all gauge stations in the domain. The position of Hohenpeissenberg (Garmisch-Partenkirchen) is indicated by a white triangle (circle).

HESSD
9, 937–982, 2012

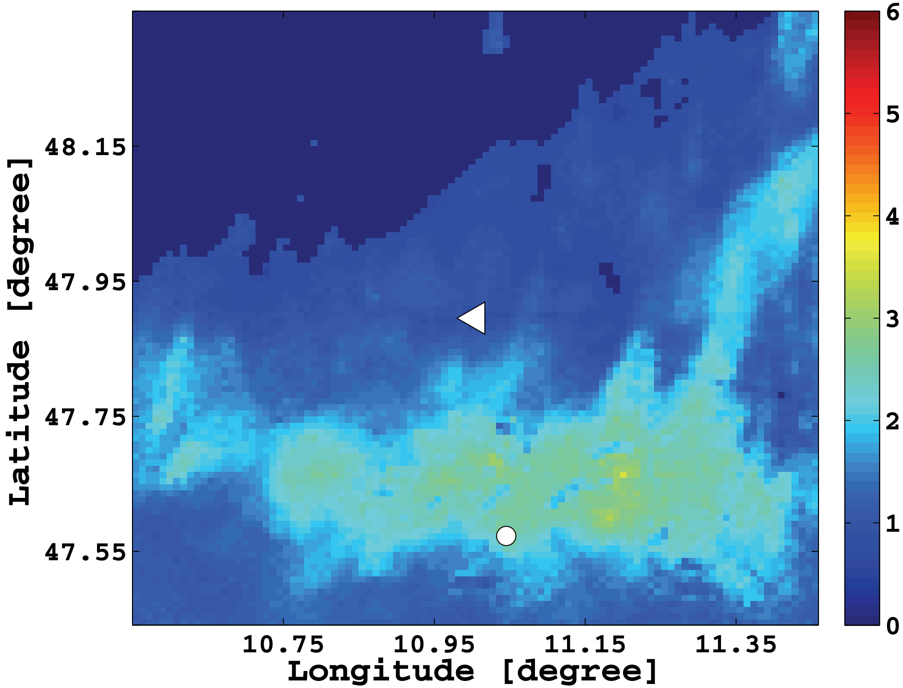

Fig. 8. Precipitation field (mm h^{-1}) in the data space derived with the *Multiple Theta* method for 14 July 2008 (13:00). The position of Hohenpeissenberg (Garmisch-Partenkirchen) is indicated by a white triangle (circle).

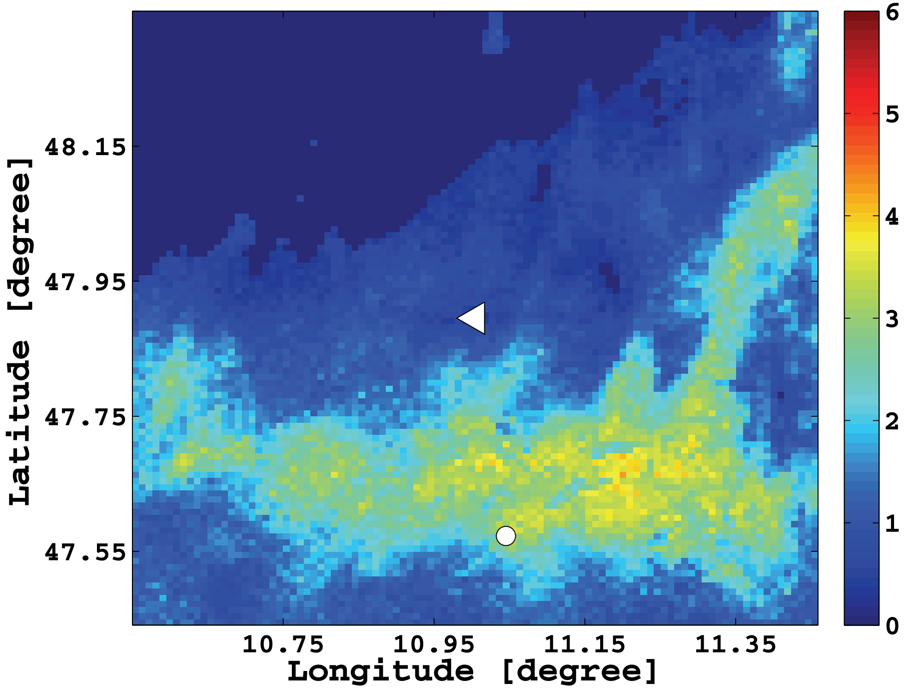

Fig. 9. Precipitation field (mm h^{-1}) in the data space derived with the *Maximum Theta* method for 14 July 2008 (13:00). The position of Hohenpeissenberg (Garmisch-Partenkirchen) is indicated by a white triangle (circle).

HESSD
9, 937–982, 2012

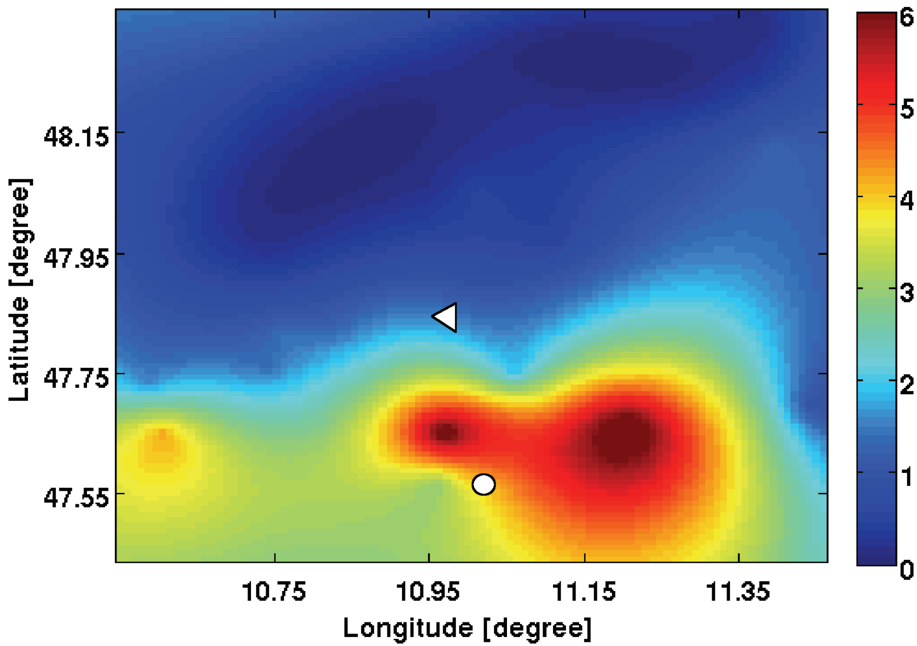

Fig. 10. Interpolated rainfall field (mm h^{-1}) derived from the observed precipitation values on 14 July 2008, (13:00). The observations from the 31 gauge stations in the radar domain are interpolated by application of an Ordinary Kriging approach. The position of Hohenpeissenberg (Garmisch-Partenkirchen) is indicated by a white triangle (circle).

HESSD
9, 937–982, 2012

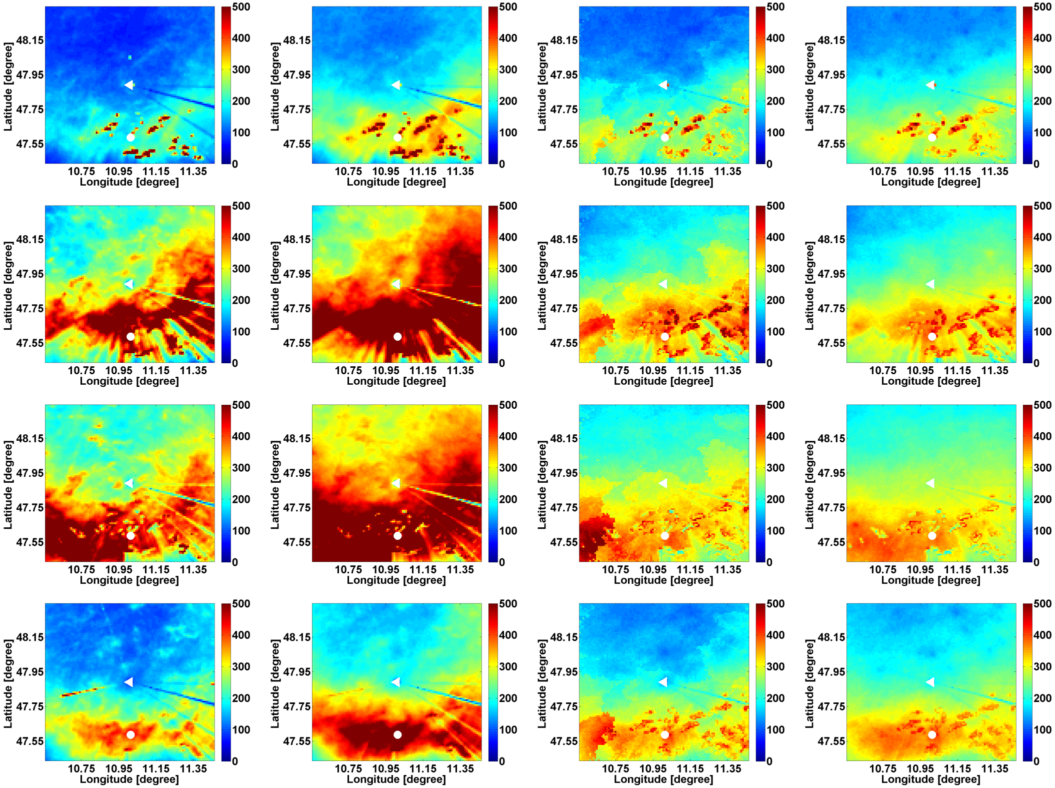

Fig. A1. Accumulated rainfall for summer (mm) obtained by radar (standard Z/R relationship), RADOLAN-correction method, *Maximum Theta*, and *Multiple Theta* approach (from left to right) for the years 2005–2008 (from top to bottom).

HESSD
9, 937–982, 2012

Copula-based assimilation of radar and gauge information

S. Vogl et al.