Hydrol. Earth Syst. Sci. Discuss., 11, 7189–7227, 2014
www.hydrol-earth-syst-sci-discuss.net/11/7189/2014/
doi:10.5194/hessd-11-7189-2014


Hydrology and Earth System Sciences
Discussions
Open Access



# Stochastic bias correction of dynamically downscaled precipitation fields for Germany through copula-based integration of gridded observation data

**G. Mao[1,2], S. Vogl[3], P. Laux[1], S. Wagner[1,2], and H. Kunstmann[1,2]**

[1]Institute of Meteorology and Climate Research (IMK-IFU), Karlsruhe Institute of Technology (KIT), Kreuzeckbahnstr. 19, 82467 Garmisch-Partenkirchen, Germany
[2]Institute of Geography, University of Augsburg, Universitätsstr. 10, 86159 Augsburg, Germany
[3]Technische Hochschule Deggendorf, Edlmairstraße 6–8, 94469, Deggendorf, Germany

Received: 17 June 2014 – Accepted: 19 June 2014 – Published: 1 July 2014

Correspondence to: G. Mao (ganquan.mao@kit.edu)

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

## Abstract

Dynamically downscaled precipitation fields from regional climate model (RCM) often cannot be used directly for local climate change impact studies. Due to their inherent biases, i.e. systematic over- or underestimations compared to observations, several correction approaches have been developed. Most of the bias correction procedures such as the quantile mapping approach employ a transfer function that based on the statistical differences between RCM output and observations. Apart from such transfer function based statistical correction algorithms, a stochastic bias correction technique, based on the concept of Copula theory, is developed here and applied to correct precipitation fields from the Weather Research and Forecasting (WRF) model. As Dynamically downscaled precipitation fields we used high resolution (7 km, daily) WRF simulations for Germany driven by ERA40 reanalysis data for 1971–2000. The REGNIE data set from Germany Weather Service is used as gridded observation data (1 km, daily) and rescaled to 7 km for this application. The 30 year time series are splitted into a calibration (1971–1985) and validation (1986–2000) period of equal length. Based on the estimated dependence structure between WRF and REGNIE data and the identified respective marginal distributions in calibration period, separately analyzed for the different seasons, conditional distribution functions are derived for each time step in validation period. This finally allows to get additional information about the range of the statistically possible bias corrected values. The results show that the Copula-based approach efficiently corrects most of the errors in WRF derived precipitation for all seasons. It is also found that the Copula-based correction performs better for wet bias correction than for dry bias correction. In autumn and winter, the correction introduced a small dry bias in the Northwest of Germany. The average relative bias of daily mean precipitation from WRF for the validation period is reduced from 10 % (wet bias) to −1 % (slight dry bias) after the application of the Copula-based correction. The bias in different seasons is corrected from 32 % (MAM), −15 % (JJA), 4 % (SON) and 28 % (DJF) to 16 % (MAM), −11 % (JJA), −1 % (SON) and −3 % (DJF), respectively. Finally,





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

the Copula-based approach is compared with linear scaling and quantile mapping correction approaches by analysing the RMSE and quantile RMSE. The results show that the Copula-based correction has improved performance in all of the quantiles, except for the extremes.

## 1 Introduction

Most climate change impact studies operate on regional and local scale. Global climate models (GCMs), however, provide climatological information only on coarse scales, usually in a horizontal resolution of 100–300 km. Since they are not able to mimic the regional and local scale climate variability, further refinement is necessary. As dynamical downscaling, regional climate models (RCMs) are capable to bridge the gap between large-scale GCM data and local-scale information to conduct climate change impact studies. Nevertheless, the RCM simulations usually do not agree well with observations even if downscaled to high spatial resolutions (Smiatek et al., 2009; Teutschbein and Seibert, 2010). Thus, they might not be useful for deriving hydrological impacts on local scales directly (Christensen et al., 2008; Bergström et al., 2001; Graham et al., 2007a, b). Therefore, further bias correction is often required. The impacts of biases on hydrological and agriculture modeling has been studied extensively (e.g., Kunstmann et al., 2004; Baigorria et al., 2007; Ghosh and Mujumdar, 2009; Ott et al., 2013). Precipitation is an important parameter in climate change impact studies (Schmidli et al., 2006). RCMs tend to generate too many wet days with small precipitation amounts (Schmidli et al., 2006; Ines and Hansen, 2006). In addition, RCMs often contain under- and overestimations of rainfall as well as incorrect representations of the seasonality (Terink et al., 2010). Therefore, several bias correction methods have been developed. These methods range from simple scaling approaches such as the linear scaling approach (e.g., Lenderink et al., 2007) and local intensity scaling (e.g., Schmidli et al., 2006) to methods like quantile mapping (e.g., Ines and Hansen, 2006). Bias correction techniques usually employ the use of a transfer function that is based on the statistical





HESSD
11, 7189–7227, 2014

differences between observed and modeled climate variables to adjust the modeled data under the assumption that these functions are stationary. A recent overview of bias correction methods for hydrological application is provided e.g. by Teutschbein and Seibert (2012) and Lafon et al. (2013).

In this study, a Copula-based stochastic bias correction method is applied which is different to the traditional transfer function approach. The main idea of this method is the identification and description of the underlying dependence structure between observed and modeled climate variables (here: precipitation) instead of the biases between them. It is well known that the traditional measures of dependence (e.g. Pearson's correlation coefficient) can only capture the strength of the linear dependence as a global parameter instead of describing the complex non-linear dependence structure between variables (Bárdossy and Pegram, 2009).

Copula-based bias correction techniques have been originally introduced by Laux et al. (2011) and Vogl et al. (2012), and are extended in this study by investigating gridded precipitation fields instead of selected stations. The Copula models are estimated for each grid cell instead of choosing the most dominant model for the whole domain. Bayesian Information Criterion (BIC) is implemented in addition to Kolmogorov–Smirnov test (K–S test) for marginal distribution Goodness-of-fit test, as the large sample size makes the K–S test highly sensitive. The performance of the correction method is analyzed for differents seasons to investigate seasonal variability. This study is based on data for a 30 year time period (1971–2000) of high resolution (7 km) dynamical downscaled precipitation fields using the Weather Research and Forecasting Model WRF-ARW (Berg et al., 2013) are used. REGNIE data from Germany Weather Service were used as the gridded observation data source. In the calibration period, only positive pairs (both REGNIE and WRF data indicate precipitation) are used to calibrate the model. Therefore, in the validation period only the days that belong to positive pairs are corrected and the other days are kept the same as the original WRF data.

The article is structured as follows: in Sect. 2 the data sets for this application are introduced. Section 3 briefly describes the basic theory of Copulas and the procedure





HESSD
11, 7189–7227, 2014

of Copula-based conditional simulations to correct RCM precipitation. Results of application of the Copula-based approach for Germany are shown in Sect. 4, followed by the discussion and conclusions (Sect. 5).

## 2 Data

In this section the data sources which are used for the application of the Copula-based bias correction method for gridded data sets is described. The newly developed approach is applied for Germany (Fig. 1) for a 30 year time period from 1971 to 2000. The RCM output and the observational data that is used in this application are both gridded data in 7 km spatial resolution and in daily scale. We split the 30 year time series into a calibration (1971–1985) and validation (1986–2000) period of equal length.

### 2.1 RCM data

Dynamically downscaled precipitation fields over Germany from a RCM simulation (Berg et al., 2013) with the non-hydrostatic WRF-ARW model (Skamarock et al., 2008) are used. For this data set, the WRF-ARW simulations are forced by ERA40 reanalysis data (Uppala et al., 2005) from 1971 to 2000 at the boundaries which implies large-scale circulation close to observations. Due to the coarse resolution of the GCM, a double-nesting approach is applied in Lambert conformal map projection. The coarse nest extends over all of Europe (42 km) and the fine nest covers Germany and the near surroundings (7 km). The model uses 40 vertical levels for both nests. For further details (e.g. parameterization schemes) on the applied WRF-ARW setup we refer to Berg et al. (2013) and the references listed therein.

### 2.2 Observational data

As observations, we used the 1 km gridded daily data set REGNIE (DWD, 2011) from the German Weather Service (DWD). The REGNIE product is available for complete

Germany from 1951 on and the number of underlying stations is approximately 2000 stations. The statistical gridding approach of station data is based on the spatial interpolation of anomalies compared to long-term mean values. For the background climatological field a multi-linear regression approach is applied where the geographical position, elevation and wind exposure of the stations are taken into account. For the calculation of the daily precipitation fields, station values are first assigned to a grid point and divided by the background data to calculate anomalies. The anomalies are spatially interpolated using inverse distance weighted interpolations, and the results are finally multiplied by the background field. For the grid cell based bias correction the 1 km REGNIE data set is up-scaled and remapped to the 7 km WRF grid such that precipitation amounts are conserved. Also, the time period is kept the same as WRF output (1971–2000).

## 3 Methodology

In this section the fundamentals of Copula theory are briefly summarized. Details about Copula theory are given e.g. in Nelsen (1999). The basis of the Copula-based bias correction algorithm used in this study is a bivariate Copula model that allows to model the dependence structure between WRF and REGNIE data. The Copula model consists of two respective marginal distributions and a bivariate Copula function and is then used to generate bias corrected WRF data by conditional stochastic sampling. Details about the bias correction algorithm are described below. In the following section, $X$ and $Y$ refer to REGNIE and WRF data set, respectively.

### 3.1 Copula theory

Let $(X,Y)$ be a pair of random variates with a realization $(x,y)$ and the bivariate joint distribution $F_{XY}(x,y)$. Following Sklar's Theorem (Sklar, 1959), there exists an unique





HESSD
11, 7189–7227, 2014

function $C$ (Copula) such that:

$$F(x,y) = C(F_X(x), F_Y(y)) \qquad x, y \in \mathbb{R}$$
$$= C(u,v) \qquad u, v \in [0,1], \tag{1}$$

with $u = F_X(x)$ and $v = F_Y(y)$.

The Copula functions provide a functional link between the two univariate marginal distributions $F_X(x)$, $F_Y(y)$. As the Copula function allows to model the pure dependence between the two variates $X$ and $Y$, it is rather flexible to describe their relationship with full freedom to the choice of the univariate marginal distributions. This is especially advantageous in cases, where the dependence structure between the variates is too complex to be modelled by a multivariate Gaussian distribution, as it is often the case for hydrometeorological variables (Salvadori and Michele, 2007; Dupuis, 2007).

### 3.2 Copula models

As a consequence of Sklar's Theorem, each complex and unknown joint distribution $F_{XY}(x,y)$ can be estimated by assuming specific parametric functions for $F_X$, $F_Y$ and $C$ in Eq. (1). The bivariate Copula model of the variates $X$ and $Y$ consists of two univariate parametric marginal distributions ($F_X(x)$ and $F_Y(y)$) and a theoretical parametric Copula function $C_\theta(u,v)$ that can be estimated separately based on the realizations $x, y$. Figure 2 visualizes the process of estimating a Copula model with a bivariate exemplary data set, i.e. realizations $(x,y)$ of the two random variates $X$ and $Y$.

A scatter plot of the two realizations $(x,y)$ is shown in Fig. 2 (left). The Copula model for the data set consists of two marginals and the theoretical Copula. Therefore, the first step is to estimate the theoretical univariate distribution functions for the two variates $X$ and $Y$ (see Fig. 2, middle).

The next step is to estimate the theoretical Copula function $C_\theta$ (see Fig. 2, right). Finally, the unknown joint distribution $F_{XY}(x,y)$ is fully determined by the marginal distributions and the Copula function, i.e. the dependence structure itself. Figure 2 visualizes the fact that different marginal distributions and theoretical Copula functions can





HESSD
11, 7189–7227, 2014

be combined independently allowing to model highly complex interdependencies between the variables $X$ and $Y$. This is especially beneficial if these interdependencies are non-linear, asymmetric or the data show heavy-tail behaviour.

### 3.3 Marginal Goodness-of-fit test

The Copula-based modelling of the dependence between $X$ and $Y$ requires the fitting of suitable marginal distributions for both data sets (REGNIE and WRF) for each grid cell. In this study, five different distribution functions are tested (Weibull, Gamma, Normal, Generalized Pareto and Exponential). For all time series (REGNIE and WRF) the parameters of the respective distribution functions are estimated by a standard maximum likelihood estimation (MLE). The Goodness-of-fit is evaluated in a two-stage process. Firstly, a Kolmogorov–Smirnov test (K–S test) is applied (Massey, 1951). As the K–S test is highly sensitive due to the large sample sizes (Serinaldi, 2008), the null hypothesis (the sample comes from the selected distribution) is rejected in some cases for all of the candidates. In other cases there might be more than one possible candidate for the best fit. For that reason, all candidates which are accepted by the K–S test are further inspected by using the Bayesian Information Criterion (BIC) (Weakliem, 1999). If all of the candidates are rejected by the K–S test, only the BIC is relevant for the selection of the best fit.

The Bayesian Information Criterion selects the optimum within a finite set of models. It is based on the likelihood function and deals with the trade-off between the Goodness-of-fit of the model and its complexity:

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

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 the BIC suggests the best fitting distribution.





HESSD
11, 7189–7227, 2014

### 3.4 Copula Goodness-of-fit test

For the Copula Goodness-of-fit test we closely follow the approach as described in Laux et al. (2011) and Vogl et al. (2012). For reason of completeness it is briefly summarized.

Since the dependence structure, i.e. the theoretical Copula function, between $X$ and $Y$ is in general not known in advance, the empirical Copula which can be calculated from the data is analyzed (Deheuvels, 1979). Let $\{r_1(1),\ldots,r_1(n)\}$ and $\{r_2(1),\ldots,r_2(n)\}$ denote the ranks of observed and modelled rainfall from day 1 to day $n$, obtained from the original data. Then, the empirical Copula, a rank based estimator of $C_\theta$, is defined as:

$$C_n(u,v) = 1/n \sum_{t=1}^{n} \mathbf{1}\left(\frac{r_1(t)}{n} \leqslant u, \frac{r_2(t)}{n} \leqslant v\right) \tag{3}$$

with $u = F_X(x)$, $v = F_Y(y)$ and $\mathbf{1}(\ldots)$ is denoting the indicator function. A visual inspection of $C_n$ allows to choose promising candidates out of the set of available theoretical parametric Copula functions. To estimate the unknown parameter $\theta \in \mathbb{R}$ for each candidate, a standard maximum likelihood (MLE) approach is used. To decide which Copula function is able to describe the dependence structure best, different Goodness-of-fit tests (e.g., Genest and Rémillard, 2008; Genest et al., 2009) are available. In this study the Goodness-of-fit test is based on the Cramér-von Mises statistic (Genest and Favre, 2007), where the empirical Copula $C_n$ is compared to the parametric estimate $C_\theta$:

$$S_n = \frac{1}{n} \sum_{t=1}^{n} \{C_\theta(u_t,v_t) - C_n(u_t,v_t)\}^2 \tag{4}$$

The specific parametric bootstrap procedure to obtain the approximate $P$-value is described by Genest et al. (2009).





HESSD
11, 7189–7227, 2014

### 3.5 Copula-based bias correction

The Copula-based bias correction applied for this study is based on the estimation of a Copula model for each pair of observed ($X$) and modelled ($Y$) rainfall for each grid cell. As soon as this Copula model ($F_X(x), F_Y(y)$ and $C_\theta(u,v)$) is estimated, conditional random samples are generated (Salvadori et al., 2007). The procedure follows the algorithm detailed in Laux et al. (2011) and Vogl et al. (2012) to generate pseudo-observations conditioned on the modelled data. We condition by purpose on the RCM data as the method is the first step for correcting future climate projection (where no observations are available). This algorithm is based on conditional probabilities of the form:

$$C_{U|V=v}(u) = P[U \leq u | V = v] = \frac{\partial C(u,v)}{\partial v} \tag{5}$$

The complete algorithm consists of the following steps:

1. estimate the theoretical marginal distributions $F_X(x)$ and $F_Y(y)$ for observation and RCM data respectively
2. transform the time series $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ to the rank space by taking $u = F_X(x)$ and $v = F_Y(y)$
3. calculate the empirical Copula $C_n(u,v)$ as a rank based estimator for the theoretical Copula function $C_\theta(u,v)$
4. estimate the Copula parameter $\theta$ and perform Goodness-of-fit tests to identify the best theoretical Copula function $C_\theta(u,v)$
5. calculate the Copula distribution conditioned on the variate $v$ representing the RCM time series in the rank space
6. generate the pseudo-observations in the rank space for each time step by using the conditional Copula distribution





HESSD
11, 7189–7227, 2014

7. transform back the random samples to the data space by using the integral transformation.

The Copula-based conditional prediction is the critical step of this bias correction approach, as it forces a certain variable (observation) to take a value when another variable (RCM) is given. To assess the uncertainty associated with this prediction, the conditional prediction process (step 6 and 7) must be repeated for a large number of times (Grégoire et al., 2008).

### 3.6 Correction strategy for continous time series

The implementation of bias correction for precipiation (a discrete variable) is more complex than a bias correction of continuous variables, e.g., temperature. In addition to general under- and overestimations, the RCM bias correction has to cope with the problem that precipitation data is zero inflated, i.e. (0,1) and (1,0) cases are possible ((0,1) case stands for a observation indicates no precipitation event but the RCM model shows a rain event, while (1,0) indicates the opposite). In general four cases have to be distinguished, namely (0,0), (0,1), (1,0), and (1,1), with 0 denotes a dry day and 1 indicates a wet day. A threshold of rainfall amount 0.1 mm per day was used to identify a wet day with respect to the usual precision of rain gauges (Dieterichs, 1956; Moon et al., 1994).

In this study, only the positive pairs (1,1) of REGNIE and WRF data are used to construct the Copula models in calibration period. Therefore, the correction of WRF data is also only applied for the (1,1) cases in validation period. In order to generate a complete bias corrected time series of WRF output, the events that are not covered by the (1,1) case are left unchanged. Note that this method does not correct for errors coming from the (0,1) and (1,0) cases while there is no error in the (0,0) case.





HESSD
11, 7189–7227, 2014

## 4 Results

In this section, details about the estimated Copula models are presented including information about the fitting of the marginal distributions and the theoretical bivariate Copula functions from the calibration period (1971–1985). Since the estimated marginal distributions reflect the statistical characteristics of RCM and observations, their differences are analyzed spatially. The fitted Copula models are applied for the validation period (1986–2000) to bias correct the WRF precipitation. It is found that the dependence structures vary intra-annually, therefore the performance of the algorithm is analyzed separately for the different seasons.

### 4.1 Estimated marginal distributions

For both REGNIE and WRF data five different distribution functions are employed for each grid cell separately: Generalized Pareto distribution (gp); Gamma distribution (gam); Exponential distribution (exp); Weibull distribution (wbl) and Normal distribution (norm). This guarantees the flexibility in selecting the most appropriate distribution for each grid cell. The Goodness-of-fit tests (K–S test and the Bayesian information criterion, see Sect. 3.3) reject the Normal distribution in all cases, while the Generalized Pareto distribution is accepted most frequently for both REGNIE and WRF (Fig. 3). The result shows a reasonable agreement of selected marginal distribution between REGNIE and WRF mainly in the Eastern and Southern parts of Germany. The patterns of the selected types follow the topography of Germany (see Fig. 1). In the Northwest of Germany, the Weibull distribution function prevails as well as in the low mountain ranges. In general, this effect is stronger for WRF while the patterns are more patchy for REGNIE.

The coincidence between REGNIE and WRF marginals is shown in the confusion matrix. Each row of the matrix represents the distribution types of REGNIE, while each column represents that of WRF (in %). The major diagonal shows the fraction of concurring marginal types. The confusion matrix for the calibration period is shown





HESSD
11, 7189–7227, 2014

in Table 1. It is found that for 42 % of grid cells, the Generalized Pareto distribution is selected for both data sources concordantly. For the Weibull distribution this holds true for 16 % of the grid cells. Since the total number of grid cells where Gamma and Exponential distribution are fitted is very low, the percentage of hits in the diagonal of the confusion matrix is small. Summing up the major diagonal gives a measure for the overall agreement. For the complete calibration series about 59 % correspond. The failures of 21 % of grid cells, where REGNIE follows the Generalized Pareto distribution and WRF follows the Weibull distribution, are predominately located in the Northwest of Germany (Fig. 3).

In order to assess for the annual variability in the precipitation time series, the marginal distributions are estimated for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF).

For both REGNIE and WRF data, the seasonal representation of the different distribution types is shown in Fig. 4. It indicates that the choice of the optimal marginal distribution clearly depends on the season. In WRF, the winter (summer) season is dominated by Exponential (Generalized Pareto). The differences for REGNIE are not that obvious since the dominant distribution type is the Generalized Pareto distribution for all seasons. For WRF data the effect of the underlying elevation on the identified distribution type is most prominent during winter and fall. In the low mountain regions the favorite marginal distribution change from fall (Weibull, Generalized Pareto) to winter (Exponential, Weibull).

The seasonal confusion matrices are shown in Table 2. The results indicate the best agreement between WRF and REGNIE (approximately 56 % of the grid cells) in summer, while in wintertime only approximately 30 % of the types agree.

As mentioned above in Sect. 3.3 the Goodness-of-fit tests follow a two-step process due to the fact that the K–S test is highly sensitive to large sample sizes. For the annual marginal distribution identification, for 99 % of the grid cells the K–S test fails and only the BIC is used for REGNIE, while the number for WRF is 68 %. Since the sample size





HESSD
11, 7189–7227, 2014

is reduced in seasonal analysis, the failures of K–S test are decreased dramatically. The results are shown in Table 3.

### 4.2 Identified Copula functions

For each grid cell the theoretical Copula function that characterizes the dependence structure between REGNIE and WRF data is identified separately. Three different one-parametric Archimedean Copulas (see Table 4) are investigated by application of the Goodness-of-fit tests described in Sect. 3.4. For the Clayton Copula the parameter $\theta$ can also take values $-1 < \theta < 0$ to model negative dependence. In that case it is

$$C_\theta(u,v) = \left[\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right]^{-\frac{1}{\theta}}. \tag{6}$$

For the data used in this study however, no negative dependence is found. Therefore it is $\theta \in [0,\infty]$ and the Clayton Copula is defined as described in Table 4. Figure 5 shows the results of the Goodness-of-fit tests for the calibration period for the complete study area. It is found that for most of the grid cells in the Southwest of Germany the Frank Copula can capture the dependence structure best, while for the Northeast of Germany the Clayton Copula provides the best fit. In total the dependence structure of 79 % of the grid cells is modelled by the Frank, 21 % by the Clayton and only 0.09 % by the Gumbel Copula.

In order to assess for the annual variability of the dependence structures between REGNIE and WRF precipitation time series, the Copula functions are identified for the different seasons separately. The corresponding results are shown in Fig. 6.

While for spring, autumn and winter the Frank Copula dominates (spring 66 %, autumn 74 % and winter 88 %), in summer the Clayton Copula provides the best fit for most of the grid cells (64 %). For all seasons the Gumbel Copula is only selected for few grid cells with spring being the season of most hits (7 % of the grid cells). In general the differences are most prominent for winter and summer (see Fig. 6).

### 4.3 Validation of the Copula-based bias correction

Based on the estimated Copula model (parametric marginal distributions and theoretical Copula functions) the conditional distribution of REGNIE conditioned on WRF is derived for each grid cell separately (see Sect. 3.5). To generate bias-corrected WRF precipitation, random samples of possible outcomes are drawn from this conditional distribution. We use a sample size of 100. The result can be interpreted as an empirical predictive distribution for corrected WRF (pseudo-observations) that is determined for all conditioning WRF precipitation values for each time step. While this stochastic bias correction method gives a full ensemble and the empirical predictive distribution of corrected WRF precipitation, for pratical reasons one can choose e.g., the expectation, median or mode to get a single corrected value.

Figure 7 examplarily shows WRF (red), REGNIE (green) and the bias-corrected WRF (blue) data for pixel 1 in Fig. 1 during wintertime 1986–1987 (positive pairs only). The box plot visualizes the spread of the generated random sample (100 members) indicating the uncertainty of the predicted bias-corrected precipitation, while the blue line shows the median of the respective emprical predictive distribution.

It can be seen from Fig. 7 that for most of the time steps the proposed Copula-based approach can sucessfully correct for biases in the modelled precipitation compared to observed values.

To investigate the spatial performance of the correction algorithm, the relative bias of RCM modelled mean daily precipitation (WRF) compared to gridded observations (REGNIE) is compared to that of the bias corrected model data (B.C. WRF) for Germany.

A comparison of corrected WRF data derived by expectation, median and mode of the predictive distribution with observations indicates that the correction performs best for the expectation value. Therefore in the following, the results are shown and analyzed for that case only.





HESSD
11, 7189–7227, 2014

Figure 8 (left) shows the bias between REGNIE and WRF, indicating wet biases in most of the study area. These wet biases are most prominent in high elevation areas following the topography of Germany. Wet biases are also detected in the Northeast of Germany, where the elevation is low. Dry biases are found in the alpine and pre-alpine areas in the Southeast of Germany as well as in the West of Germany. After the application of the Copula-based correction algorithm, the wet biases are corrected for most of the domain, except for a very small region in the Northeast (see Fig. 8, right). It is also found that the dry bias can also be significantly reduced, but small dry biases are introduced in some areas in the West of the domain. The average of the bias for the whole study area is reduced from 10 to −1 %.

A performance analysis with respect to seasonal variations is shown in Fig. 9. It shows that the relative bias is even larger for different seasons. Figure 9 (left) shows the relative bias between uncorrected WRF mean daily precipitation and the REGNIE data set for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF, from top to bottom). The WRF model tends to generate too much precipitation in spring and winter for the majority of grid cells in the study area. For summer and autumn, there are also regions found, where the model is too dry. These regions are mostly located in the North and in the South of Germany. This effect is found to be strongest in summer while in autumn areas with an overestimation of precipitation are still found in the Northeast and Southwest of Germany. In all cases, the bias is influenced by the underlying terrain showing an overesimation especially in regions with higher altitude. The average of the bias from spring to winter are 32, −15, 4 and 28 %, respectively. Figure 9 (right) shows the relative bias between corrected WRF mean daily precipitation and the REGNIE data set for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF, from top to bottom). It can be seen that the Copula-based correction efficiently removes most of the biases indicating a comparable performance for all seasons. Figure 9 especially for spring and winter indicates that the correction is tending to be more suitable to correct for overestimation of the rainfall. The underestimation of precipitation, that is most prominent in summer, however, is still significantly reduced. In autumn and winter the Copula-based correction reduces the rainfall amounts too much for the west of Germany, introducing a small dry bias in that region. The average bias are reduced to 16 −11, −1 and −3 respectively for different seasons from spring to winter.

In the following, it is further analyzed how well the model can reproduce the intra-annual variability of observed precipitation and how the performance for the different seasons is influenced by the Copula-based correction algorithm.

To investigate typical situations in detail, the results are shown for four specific grid cells in the study area (see Fig. 1): Grid cells 1 and 3 are selected as they show the highest wet bias between WRF and the REGNIE. Grid cell 2 is located in the region where a dry bias was generated by the WRF in summer and autumn and a wet bias was generated in winter. Grid cell 4 represents a case where the agreement between uncorrected model data and REGNIE observations is already good (see Fig. 9).

Figure 10 shows mean monthly precipitation derived for the validation period (1986–2000) for the selected grid cells 1–4 (see Fig. 1 for their exact locations). The number of the respective grid cell is noted in the upper left corner.

The results for grid cell 1 in Fig. 10 confirm the fact that the RCM model results strongly overestimate the precipitation amount in that case. The annual variability of the observations is in general reproduced, except for a strong increase of the mean precipitation in August that is not found in the observations. This behaviour is found also for grid cell 3 indicating a relatively too dry summer season. For grid cells 1 and 3, the Copula-based correction is found to be able to correct for the overestimation of precipitation amounts as well as for the effect of a too strong decrease of precipitation in August. However, the correction is introducing a slight underestimation mainly during summer and autumn instead. For grid cell 2, the correction shows a good performance by decreasing the rainfall amounts when the RCM is overestimating, and increasing the amounts when the RCM has underestimated it. The correction reduced the wet bias efficiently, while the dry bias is corrected less efficiently. The effectiveness of this correction is also highlighted by an analysis of the results for grid cell 4. Even if the





HESSD
11, 7189–7227, 2014

performance of WRF was already satisfactory, the algorithm was still able to further improve the results.

The performance of the proposed bias correction approach is also assessed by comparing to two standard correction methods, i.e. linear scaling and quantile mapping correction. The root mean square error (RMSE) between the observed (REGNIE) and bias corrected modelled data (B.C. WRF) is calculated for different bias correction methods. The original RMSE (between REGNIE and WRF) is also computed as a reference. To assess the performance for more specific properties, e.g., RMSE for different magnitude of observed precipitation, a quantile RMSE analysis is done for four grid cells the same as in Fig. 10. The result from the validation period is shown in Fig. 11 and similar results are found for the entire study area. The RMSE in different quantiles are represented by $\text{RMSE}_{0.1}$, $\text{RMSE}_{0.2}$, ..., $\text{RMSE}_{1.0}$, while the subscript indicates the magnitude level. $\text{RMSE}_{0.1}$ evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From $\text{RMSE}_{0.2}$ to $\text{RMSE}_{1.0}$ the root mean square errors are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example, $\text{RMSE}_{1.0}$ indicates the errors in the magnitude of the most extreme events.

As it can be seen from Fig. 11, the RMSE (a global evaluation) for all of the four grid cells are reduced by the Copula-based bias correction, while for the other two methods the RMSE are not decreased so efficiently. The quantile mapping correction even increased the RMSE in grid cells 2 and 4. The performance is even better for the proposed approach in the quantile RMSE analysis. The Copula-based method provides an equal or better correction for RMSE in almost every quantile except the part of the most extreme values ($\text{RMSE}_{1.0}$). As the Copula-based correction is only applied for (1,1) cases, the (0,1) errors are not corrected (Fig. 11, $\text{RMSE}_{0.1}$). The results show that for linear scaling and the quantile mapping correction, the (0,1) errors are also not corrected.

## 5 Discussion and conclusions

In this study, a Copula-based stochastic bias correction technique for RCM output is introduced. Copulas are able to capture the non-linear dependencies between variables (here: between RCM and gridded observed precipitation) including a reliable description of the dependence structure in the tails of the joint distribution. This is not possible e.g. by using a Gaussian approach or methods based on the Pearson's correlation coefficient. Yet, another albeit more practical advantage of this approach is that the univariate marginal distributions can be modeled independently from the dependence function, i.e. the Copula. This provides more flexiblity to construct a correction model by combining different marginal distributions and Copula functions, as many parametric univariate distribution and theoretical Copulas are available.

This study is an extension of the two former studies of Laux et al. (2011) and Vogl et al. (2012) by applying the Copula-based bias correction technique to high resolution RCM precipitation output and a gridded observation product. Compared to those two studies, this study is based on a framework to:

- Work on a grid cell base and to estimate the Copula model (marginal distributions and Copula function) for each grid cell separately rather than selecting e.g. the most dominant model. Therefore, the statistical characteristics of observed (REGNIE) and modelled data (WRF) and their dependence structure is visualized spatially and analyzed for the first time.

- Implement the Bayesian Information Criterion in addition to the Kolmogorov–Smirnov test for the marginal Goodness-of-fit test. From previous studies we found that very large sample sizes may bias the result of the K–S test, leading to the rejection of the null hypothesis (the sample comes from the selected distribution) most of the time.





HESSD
11, 7189–7227, 2014

– Estimate the Copula model for every season separately. Thus, different precipitation geneses types are not masked by the same models. This, in general, leads to stronger dependencies and robuster models.

Positive REGNIE and WRF pairs of fifteen years daily precipitation in calibration period (1971–1985) are used to establish the Copula models. The results indicate discrepancies between the fitted marginal distributions of REGNIE and WRF-EAR40 (see Figs. 3 and 4). The estimated marginal distributions for WRF show distinct spatial (strongly related to the orography of the domain) and seasonal patterns (clear differences between summer and winter, similar patterns for spring and fall season). The distributions are more scattered for the REGNIE data.

For the dependence function it is found that the fitted Copula families vary both in space and time (seasonally) (see Figs. 5 and 6). The fact that different dependence structures exist for the different seasons indicates that the method corrects for different dominating precipitation types, i.e. convective and stratiform precipitation.

The assumptions of this approach is that the dependence structure between observed and modelled precipitation is stationary over the period of interest. The validation results show that the proposed approach successfully corrected the errors in RCM derived precipitation, even when a slight dry bias might be introduced by this correction (see Figs. 8 and 9). It is also found that the correction method performs better for overestimation correction rather than for underestimation correction. By applying a specific analysis with four specific grid cells, results show that the Copula-based correction provides better results than linear scaling and quantile mapping.

The proposed algorithm is based on the identification and description of the dependence structure between observed and modelled data which is represented by a Copula model. Apart from traditional approaches such as linear scaling and quantile mapping, which are based on a bijection transfer function, this method corrects the biases dynamically and offers the possibility to estimate the uncertainties inherently.

*Acknowledgements.* The authors acknowledge funding from China Scholarship Council (CSC), the Bavarian State Ministry of the Environment and Public Health (reference number

VH-ID:32722/TUF01UF-32722) and the Federal Ministry of Education and Research as part of the research project LUCCi (reference number 01LL0908C). The WRF simulation were carried out at HLRS at the University of Stuttgart within the project High resolution regional climate modeling for Germany using WRF. We would also extend our appreciation to the German Weather Service (DWD) for the REGNIE data set.

The service charges for this open access publication have been covered by a Research Centre of the Helmholtz Association.

## References

Baigorria, G. A., Jones, J. W., Shin, D.-W., Mishra, A., and O'Brien, J. J.: Assessing uncertainties in crop model simulations using daily bias-corrected Regional Circulation Model outputs, Clim. Res., 34, 211–222, doi:10.3354/cr00703, 2007. 7191

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. 7192

Berg, P., Wagner, S., Kunstmann, H., and Schädler, G.: High resolution regional climate model simulations for Germany: part I validation, Clim. Dynam., 40, 401–414, doi:10.1007/s00382-012-1508-8, 2013. 7192, 7193

Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Pettersson, A., and Rummukainen, M.: Climate change impacts on runoff in Sweden assessments by global climate models , dynamical downscaling and hydrological modelling, Clim. Res., 16, 101–112, doi:10.3354/cr016101, 2001. 7191

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, L20709, doi:10.1029/2008GL035694, 2008. 7191

Deheuvels, P.: La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance, B. Cl. Sci. Ac. Roy. Belg., 65, 274–292, 1979. 7197

Dieterichs, H.: Frequency of dry and wet spells in san salvador, Geofisica pura e applicata, 33, 267–272, 1956. 7199

Dupuis, D. J.: Using copulas in hydrology: Benefits, cautions, and issues, J. Hydrol. Eng., 12, 381–393, doi:10.1061/(ASCE)1084-0699(2007)12:4(381), 2007. 7195

DWD: Abteilung Hydrometeorologie: REGNIE (REGionalisierte NIEderschlage): Verfahrensbeschreibung und Nutzeranleitung, interner Bericht im DWD, Offenbach, 2011. 7193

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, doi:10.1061/(ASCE)1084-0699(2007)12:4(347), 2007. 7197

Genest, C. and Rémillard, B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, Ann. I. H. Poincaré-Pr., 44, 1096–1127, doi:10.1214/07-AIHP148, 2008. 7197

Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: a review and a power study, Insur. Math. Econ., 44, 199–213, doi:10.1016/j.insmatheco.2007.10.005, 2009. 7197

Ghosh, S. and Mujumdar, P. P.: Climate change impact assessment: uncertainty modeling with imprecise probability, J. Geophys. Res., 114, D18113, doi:10.1029/2008JD011648, 2009. 7191

Graham, L. P., Andréasson, J., and Carlsson, B.: Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River, Climatic Change, 81, 293–307, doi:10.1007/s10584-006-9215-2, 2007a. 7191

Graham, L. P., Hagemann, S., Jaun, S., and Beniston, M.: On interpreting hydrological change from regional climate models, Climatic Change, 81, 97–122, doi:10.1007/s10584-006-9217-0, 2007b. 7191

Grégoire, V., Genest, C., and Gendron, M.: Using copulas to model price dependence in energy markets, Energy Risk, 5, 62–68, 2008. 7199

Ines, A. V. M. and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, Agr. Forest Meteorol., 138, 44–53, doi:10.1016/j.agrformet.2006.03.009, 2006. 7191

Kunstmann, H., Schneider, K., Forkel, R., and Knoche, R.: Impact analysis of climate change for an Alpine catchment using high resolution dynamic downscaling of ECHAM4 time slices, Hydrol. Earth Syst. Sci., 8, 1031–1045, doi:10.5194/hess-8-1031-2004, 2004. 7191

Lafon, T., Dadson, S., Buys, G., and Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods, Int. J. Climatol., 33, 1367–1381, doi:10.1002/joc.3518, 2013. 7192

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. 7192, 7197, 7198, 7207

Lenderink, G., Buishand, A., and van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, Hydrol. Earth Syst. Sci., 11, 1145–1159, doi:10.5194/hess-11-1145-2007, 2007. 7191

Massey, F. J.: The Kolmogorov–Smirnov test for goodness of fit, J. Am. Stat. Assoc., 46, 68–78, doi:10.2307/2280095, 1951. 7196

Moon, S.-E., Ryoo, S.-B., and Kwon, J.-G.: A Markov chain model for daily precipitation occurrence in South Korea, Int. J. Climatol., 14, 1009–1016, doi:10.1002/joc.3370140906, 1994. 7199

Nelsen, R. B.: An Introduction to Copulas, Springer, New York, 1999. 7194

Ott, I., Duethmann, D., Liebert, J., Berg, P., Feldmann, H., Ihringer, J., Kunstmann, H., Merz, B., Schaedler, G., and Wagner, S.: High resolution climate change impact analysis on medium sized river catchments in Germany: an ensemble assessment, J. Hydrometeorol., 14, 1175–1193, doi:10.1175/JHM-D-12-091.1, 2013. 7191

Salvadori, G. and Michele, C. D.: On the use of copulas in hydrology: theory and practice, J. Hydrol. Eng., 12, 369–380, doi:10.1061/(ASCE)1084-0699(2007)12:4(369), 2007. 7195

Salvadori, G., Michele, C. D., Kottegoda, N. T., and Rosso, R.: Extremes in Nature: An Approach Using Copulas, Springer, New York, 2007. 7198

Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, Int. J. Climatol., 26, 679–689, doi:10.1002/joc.1287, 2006. 7191

Serinaldi, F.: Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective, Stoch. Env. Res. Risk A., 23, 677–693, doi:10.1007/s00477-008-0249-z, 2008. 7196

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., Huang, X., Wang, W., and Powers, J.: A description of the advanced research WRF version 3, Technical Report 459, NCAR, Boulder, Colorado, USA, 2008. 7193

Sklar, A.: Fonctions de répartition à n dimensions et leurs marges, Publications de l'Institut Statistique de l'Université de Paris, 8, 229–231, 1959. 7194

Smiatek, G., Kunstmann, H., Knoche, R., and Marx, A.: Precipitation and temperature statistics in high-resolution regional climate models: evaluation for the European Alps, J. Geophys. Res., 114, D19107, doi:10.1029/2008JD011353, 2009. 7191

Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, Hydrol. Earth Syst. Sci., 14, 687–703, doi:10.5194/hess-14-687-2010, 2010. 7191

Teutschbein, C. and Seibert, J.: Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies, Geogr. Compass, 4, 834–860, doi:10.1111/j.1749-8198.2010.00357.x, 2010. 7191

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods, J. Hydrol., 456–457, 12–29, doi:10.1016/j.jhydrol.2012.05.052, 2012. 7192

Uppala, S., Kållberg, P., Simmons, A., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. Roy. Meteor. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, 2005. 7193

Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H.: Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields, Hydrol. Earth Syst. Sci., 16, 2311–2328, doi:10.5194/hess-16-2311-2012, 2012. 7192, 7197, 7198, 7207

Weakliem, D. L.: A critique of the bayesian information criterion for model selection, Sociol. Method. Res., 27, 359–397, doi:10.1177/0049124199027003002, 1999. 7196

**Table 1.** Confusion matrix between REGNIE and WRF for the different distribution types.

| | | WRF | | | |
|---|---|---|---|---|---|
| | | gp | gam | exp | wbl |
| REGNIE | gp | 42.04 % | 1.27 % | 1.55 % | 20.79 % |
| | gam | 4.92 % | 0.5 % | 0.18 % | 2.44 % |
| | exp | 0.27 % | 0 % | 0 % | 0.23 % |
| | wbl | 7.14 % | 1.94 % | 0.79 % | 15.93 % |





HESSD
11, 7189–7227, 2014

**Table 2.** Seasonal confusion matrix of fitted REGNIE and WRF precipitation distribution.

| MAM | | WRF | | | |
|---|---|---|---|---|---|
| | | gp | gam | exp | wbl |
| REGNIE | gp | 39.57 % | 0.29 % | 25.68 % | 3.89 % |
| | gam | 2.32 % | 0.12 % | 1.32 % | 0.18 % |
| | exp | 2.68 % | 0.02 % | 3.03 % | 0.14 % |
| | wbl | 8.88 % | 0.56 % | 7.81 % | 3.51 % |
| **JJA** | | **WRF** | | | |
| | | gp | gam | exp | wbl |
| REGNIE | gp | 42.3 % | 0.09 % | 0.39 % | 11.58 % |
| | gam | 0.72 % | 0.14 % | 0.04 % | 0.83 % |
| | exp | 1.74 % | 0 % | 0 % | 0.81 % |
| | wbl | 26.4 % | 0.62 % | 0.61 % | 13.73 % |
| **SON** | | **WRF** | | | |
| | | gp | gam | exp | wbl |
| REGNIE | gp | 35.43 % | 0.08 % | 6.36 % | 18.83 % |
| | gam | 1.55 % | 0.29 % | 0.95 % | 1.14 % |
| | exp | 0.51 % | 0 % | 0.15 % | 0.41 % |
| | wbl | 11.23 % | 0.29 % | 4.88 % | 17.9 % |
| **DJF** | | **WRF** | | | |
| | | gp | gam | exp | wbl |
| REGNIE | gp | 8.92 % | 1.25 % | 24.66 % | 7.12 % |
| | gam | 2.18 % | 0.27 % | 7.65 % | 1.21 % |
| | exp | 1.44 % | 0.48 % | 8.08 % | 1.12 % |
| | wbl | 6 % | 0.89 % | 16.42 % | 12.31 % |





HESSD
11, 7189–7227, 2014

HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.





**Table 3.** The proportion of grid cells for both REGNIE and WRF that K–S test failed and only BIC is used in Goodness-of-fit procedure.

| | Spring | Summer | Autumn | Winter |
|---|---|---|---|---|
| REGNIE | 25.83 % | 10.86 % | 38.38 % | 56.13 % |
| WRF | 0.31 % | 10.61 % | 12.26 % | 3.88 % |

HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.





**Table 4.** Theoretical Archimedean Copula functions used in this study.

| Copulas | $C_\theta(u,v)$ | Generator $\varphi_\theta(t)$ | Parameter $\theta \in$ |
|---|---|---|---|
| Clayton | $(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$ | $\frac{1}{\theta}(t^{-\theta}-1)$ | $(0,+\infty)$ |
| Frank | $-\frac{1}{\theta}\ln(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1})$ | $-\ln(\frac{e^{-\theta t-1}}{e^{-\theta}-1})$ | $(-\infty,+\infty)\backslash\{0\}$ |
| Gumbel | $e^{-((-\ln(u)^{\theta})+(-\ln(v)^{\theta}))^{\frac{1}{\theta}}}$ | $(-\ln(t)^{\theta})$ | $[1,+\infty]$ |

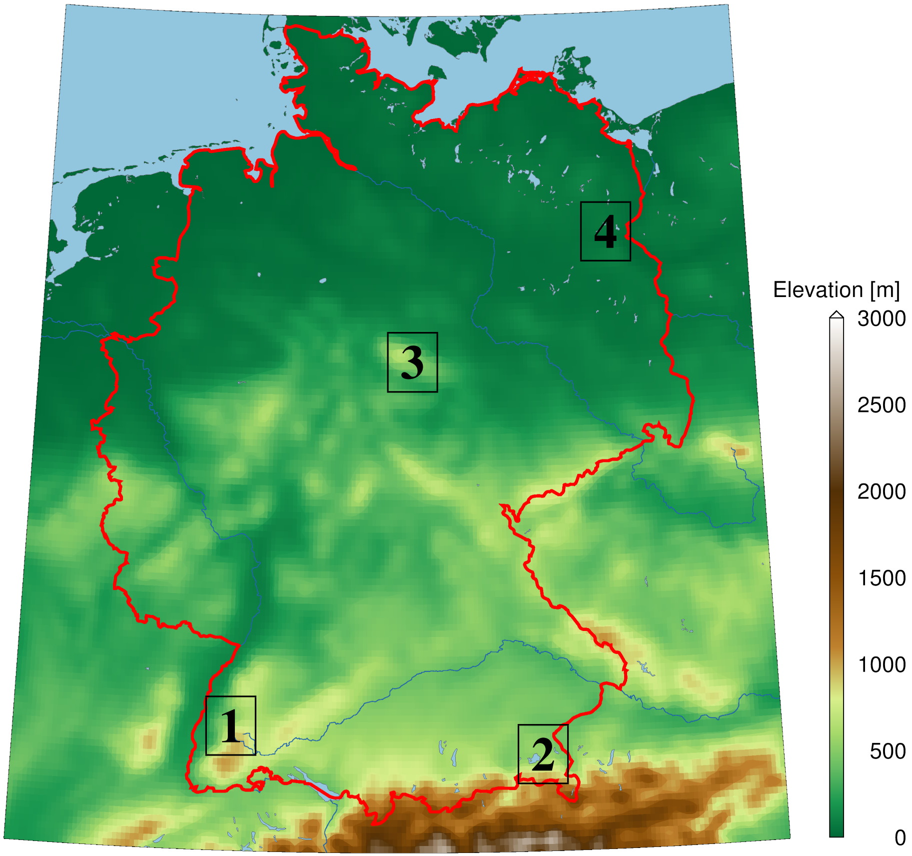


**Figure 1.** Terrain elevation of Germany (DEM). The numbers represent the position of the four specific grid cells for which the performance of the Copula-based algorithm is analyzed in Fig. 10.





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

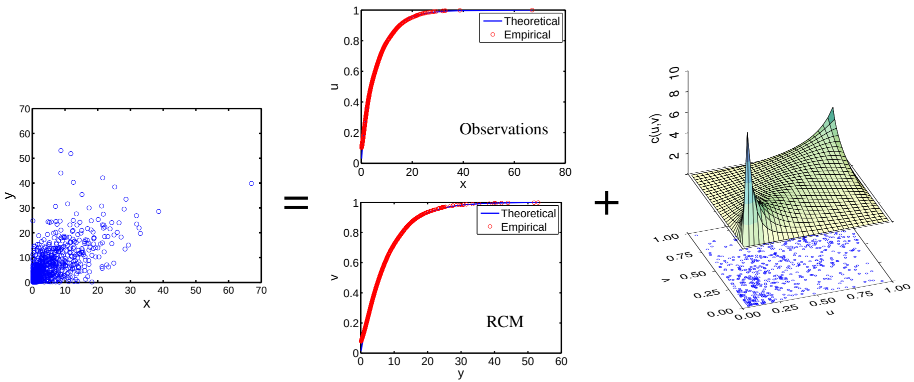


**Figure 2.** Visualisation of a bivariate Copula model consisting of two marginal distributions and a theoretical Copula function that describes the pure dependence.





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

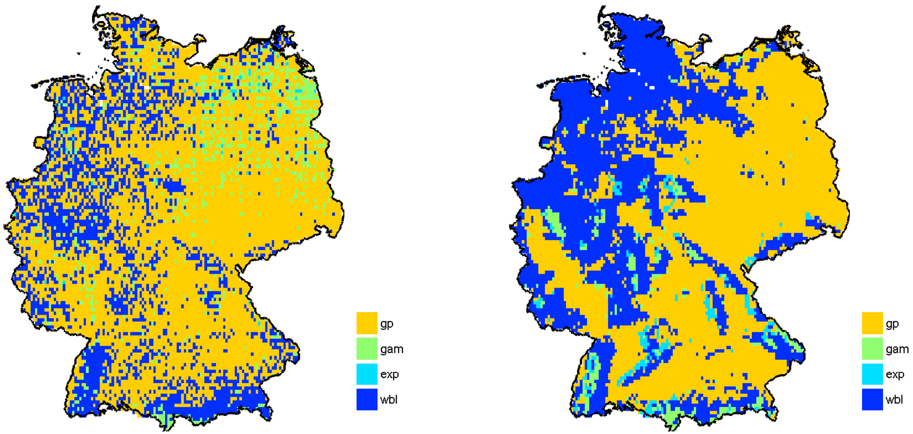


**Figure 3.** Estimated marginal distributions of precipitation for Germany for both REGNIE (left) and WRF (right). The results are shown for the calibration period (1971–1985) and positive pairs only.





HESSD
11, 7189–7227, 2014

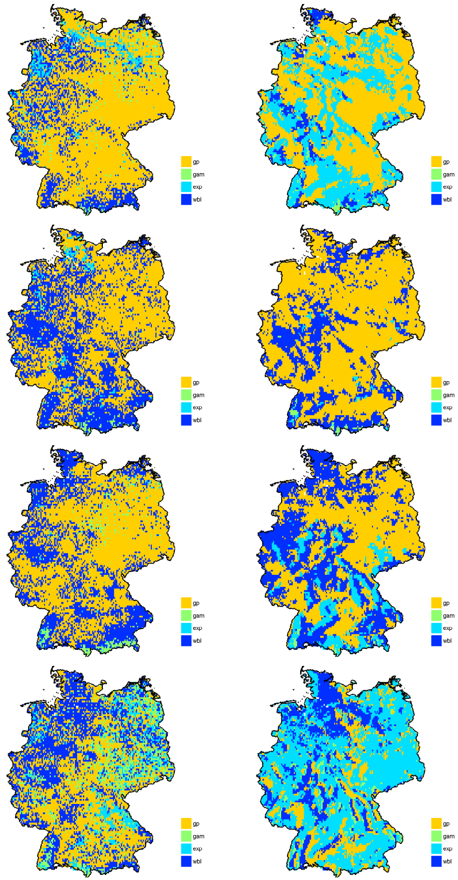


**Figure 4.** Estimated marginal distribution of precipitation for the different seasons for REGNIE (left column) and WRF (right column) in Germany. The results are shown for the calibration period (1971–1985) for positive pairs only. Spring (MAM), summer (JJA), autumn (SON) and winter (DJF) are illustrated from top to bottom.





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

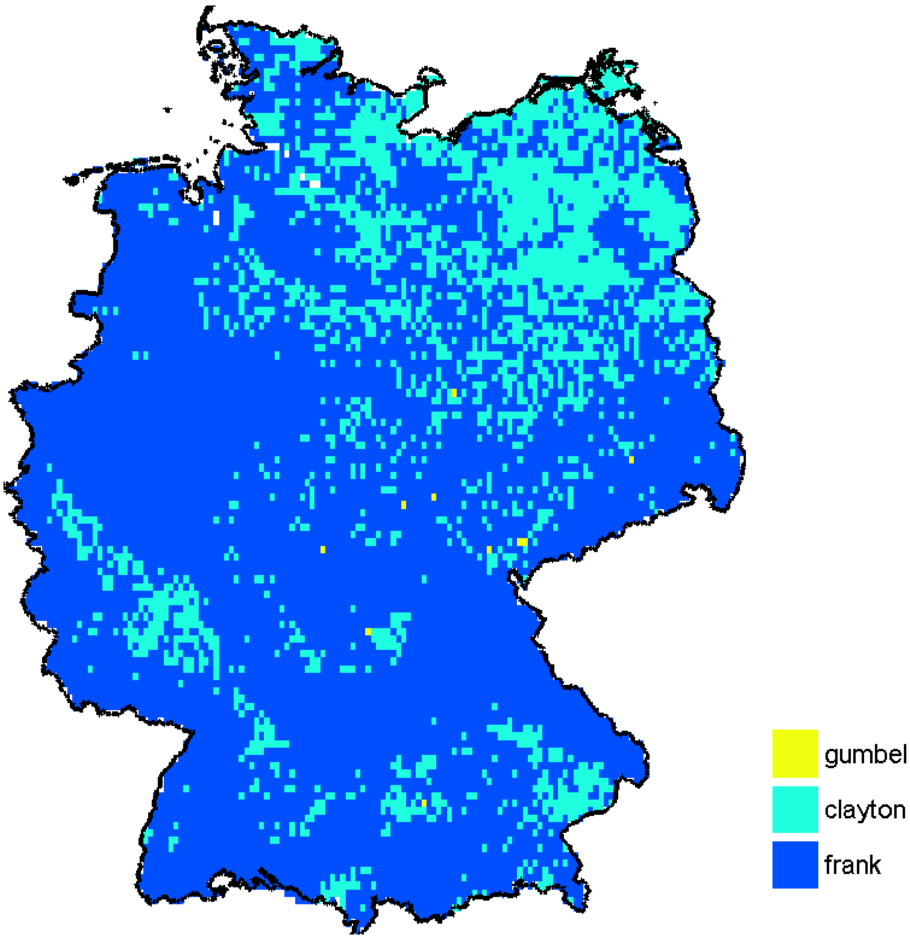


**Figure 5.** Identified Copula functions between REGNIE and WRF precipitation in the calibration period (1971 to 1985) with only positive pairs.





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

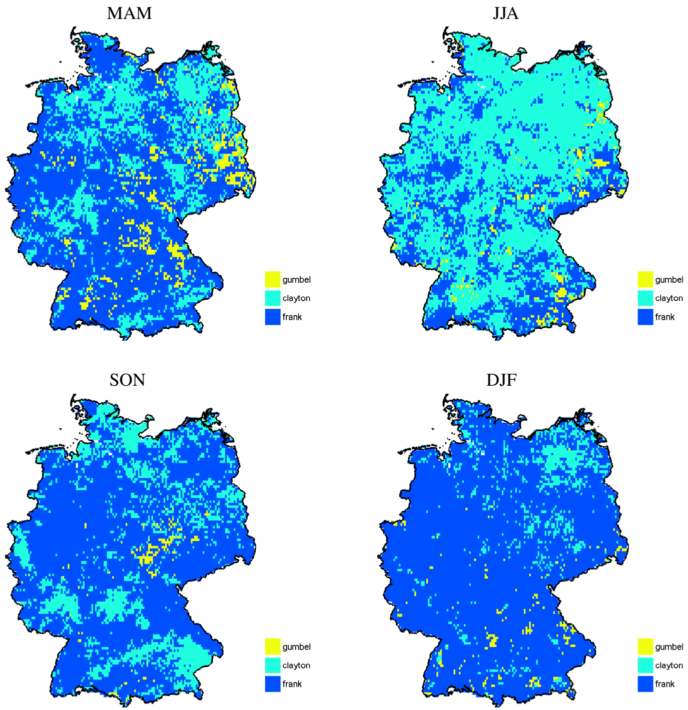


**Figure 6.** Fitted Copula functions between REGNIE and WRF precipitation (calibration period (1971–1985), positive pairs only). The Copulas are identified for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF, from left to right and from top to bottom).





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

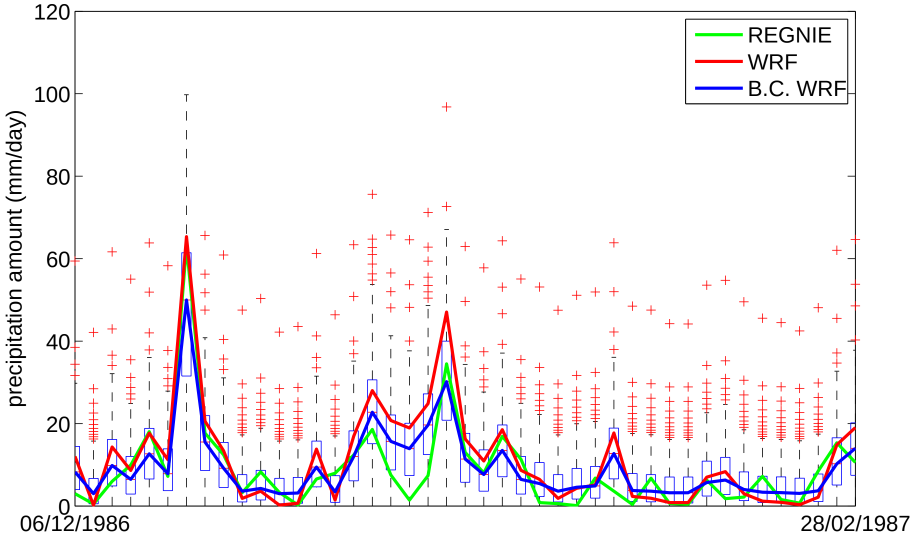


**Figure 7.** Comparison of bias-corrected WRF data (blue) with the original WRF data (red) and REGNIE (green) in winter 1986–1987 (positive pairs only) for pixel 1 in Fig. 1. For each time step 100 realisations are drawn from the conditional distribution visualized by the box-whiskers (boxes are defined by the lower $Q1$ and the upper quartile $Q3$). The length of the whiskers is determined by $1.5 \cdot (Q3 - Q1)$ and outliers, i.e. data values beyond the whiskers are marked by crosses.





HESSD
11, 7189–7227, 2014

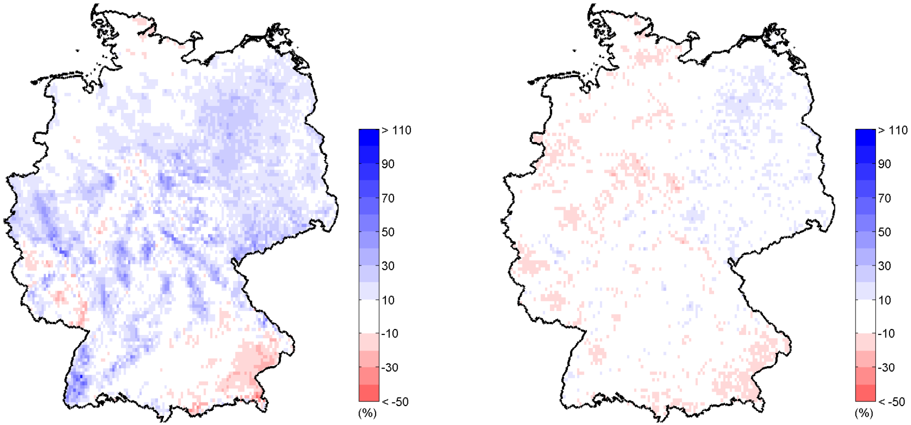


**Figure 8.** Relative bias of mean daily precipitation for uncorrected (left) and corrected WRF precipitation field (right). The results are based on the validation period 1986–2000.





HESSD
11, 7189–7227, 2014

Copula-based bias correction for RCM precipitation

G. Mao et al.

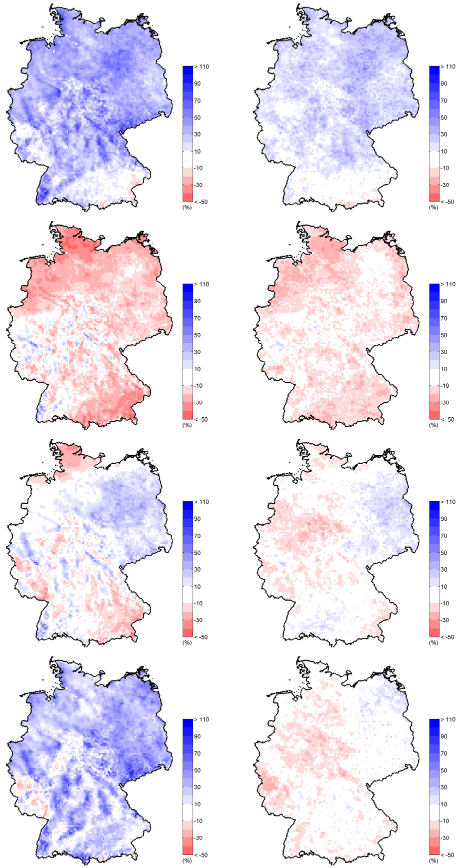

**Figure 9.** Relative bias between uncorrected (left) and corrected (right) WRF mean daily precipitation and the REGNIE data set in Germany for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF, from top to bottom). The results are derived for the validation time period (1986–2000).





HESSD
11, 7189–7227, 2014

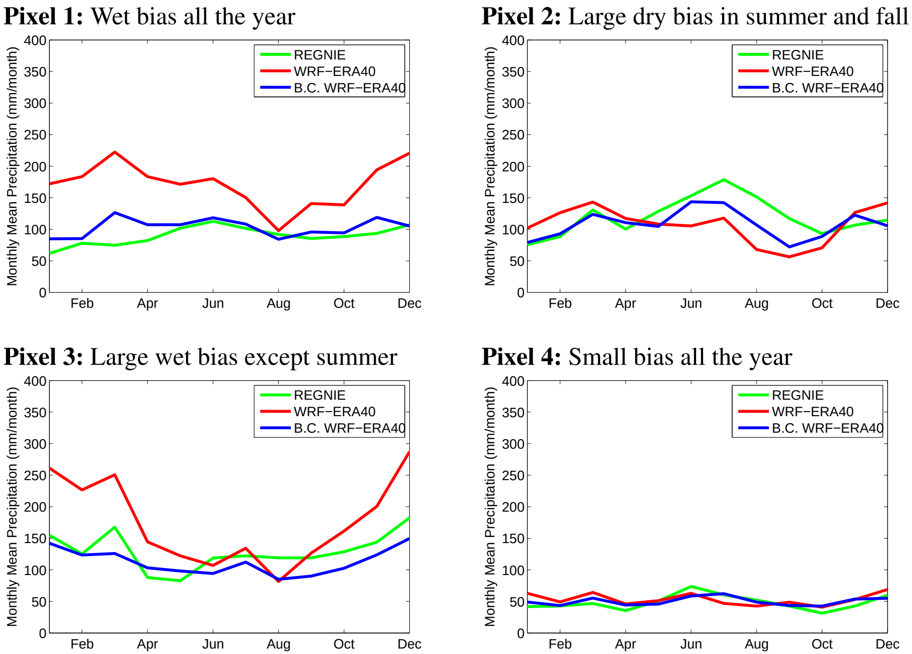


**Figure 10.** Comparison of bias corrected WRF mean monthly precipitation (blue) with REGNIE (green) and original WRF data (red) for the selected four pixel 1–4 in the validation period from 1986 to 2000. The number of the respective grid cell is noted in the upper left corner of each plot.

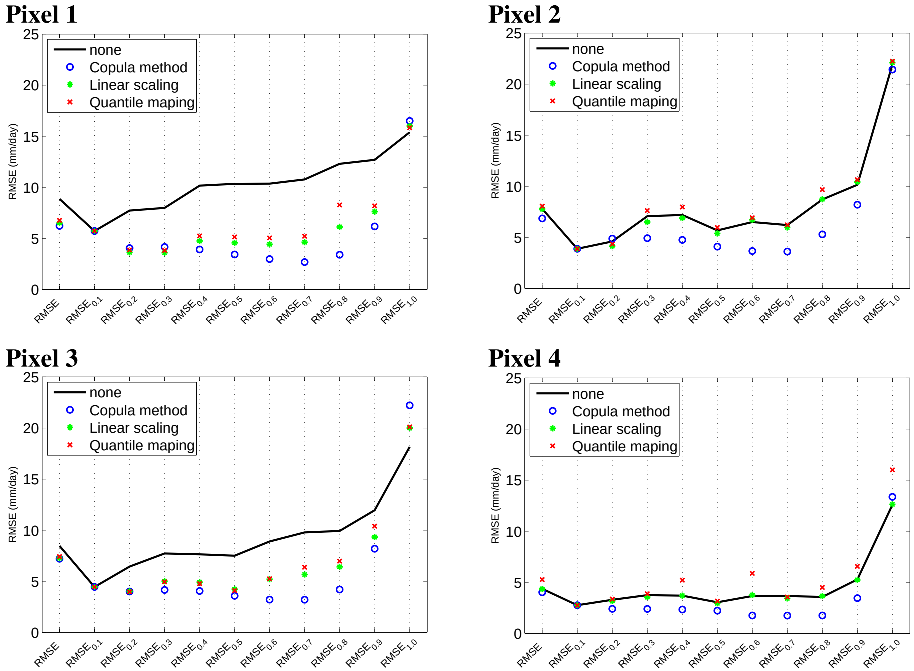


**Figure 11.** Root mean square errors (RMSE) and the root mean square errors for specific probability intervals ($RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$ for different methods). the selected four pixel are the same as Fig. 10. The black solid line indicates the errors without correction. The results are derived from the validation period from 1986 to 2000.





HESSD
11, 7189–7227, 2014