Author Response

Please note that all pages and line numbers refer to those in the revised manuscript. In the revised manuscript, the blue color indicates the added contents, the green color implies that the contents are updated (only for the Figures), and the red color with strikethrough indicates the deleted contents. The numbering of all figures has been updated.

Firstly, we would like to thank the editor for giving us the opportunity to modify our manuscript. We also want to thank Prof. Bárdossy and the anonymous reviewer #2 for the comments that helped us to improve the quality of the manuscript. The manuscript is revised now following the comments of the reviewers. We additionally improved it in terms of grammar and spelling. Our point-by-point response to the comments is given in the following:

Referee Comments 1

General Comments

The goal of bias correction is to provide time series which preserve the signal of the meteorological models, but are not biased with respect to the observations.

Comment 1

The methodology provides a set of realizations as time series. The evaluations are based on the mean of theses series. The mean is however not a good bias corrected series, as it has a different marginal (with reduced variance) than what was assumed. Thus the comparisons made with the mean of the realizations are misleading and do not reflect the quality of the corrected time series.

Answer

The major goal and motivation of this study is to develop a new framework for the bias correction of precipitation time series for subsequent modellers, e.g. in hydrology, based on the concept of Copulas. Our focus is on the introduction and demonstration of the methodology. The framework provides the possibility to obtain a large set of random realizations that give additional information about the *Probability Density Function* (PDF) for each corrected time step. Subsequent modellers usually request one corrected value for each time step rather than a PDF. This is mainly due to practical reasons since they cannot handle the full PDF for subsequent analyses. Therefore, in this study we take a typical value (e.g. mean, median or mode) as the estimator of the derived PDF for each time step. The mean is finally selected as estimator because it is found to perform best. The mean indeed follows a different marginal distribution with reduced variance. However, we are more interested in better correcting each time step separately instead of better matching the statistical distribution. To justify our method, its performance is compared to the quantile mapping approach.

The following changes were made in the revised manuscript:

Page 2, Line 28:
"Finally, the Copula-based approach is compared to the quantile mapping correction method. The Root Mean Square Error (RMSE) and the percentage of the corrected time steps that are closer to the observations are analyzed.The Copula-based correction derived from the mean of the sampled distribution reduces the RMSE significantly, while e.g. the quantile mapping method results in an increased RMSE for some regions."

Page 17, Line 10:
"While this stochastic bias correction method gives a full ensemble and the empirical predictive distribution of corrected WRF precipitation, for practical reasons one can choose e.g., the expectation, median or mode to get single corrected values. This can be regarded as a Copula-based regression by taking such a "typical" value as the estimator of the derived empirical predictive distribution of corrected WRF precipitation. It is noted that this "typical" value (e.g. the mean) can also be directly derived from the analytical Copula-based conditional CDF (mean regression curve, Nelsen, 1999)."

Page 18, Line 1:
"A comparison of corrected WRF data derived by the expectation, median and mode of the predictive distribution with observations indicates that the correction performs best for the expectation value (see Fig. 10). Both simulations based on the median and the mode tend to underestimate the precipitation values, thus causing a dry bias over the domain. Therefore in the following, the results are shown and analyzed for that case only."

Page 21, Line 7:
"The quantile mapping method is often used in bias correction of RCM derived precipitation (Dosio and Paruolo, 2011; Gudmundsson et al., 2012). This method corrects the bias by rescaling the values of the RCM so that the distribution of the RCM matches that of the observations. It corrects all moments of the RCM precipitation distribution under the assumption of a perfect dependence among the ranks. This full dependence assumption is limited: In our study area, the rank correlation between the datasets varies between 0.3 and 0.6 (see Fig. 15).

The quantile mapping correction has been performed for comparison to the Copula-based approach. The RMSE between the observed (REGNIE) and bias corrected modelled data is calculated for both the Copula-based correction and the quantile mapping method. The original RMSE (between REGNIE and WRF) is also computed as a reference. For the Copula-based approach, we calculated the RMSE for all the simulations with respect to the mean-, median- and mode value. The changes of the RMSE by different corrections over the study area are shown in Fig. 16. The Copula-based correction derived from the mean regression reduces the RMSE significantly with an average number of −12% over the domain. The Copula-based correction derived from the median also reduces the RMSE, but to a less degree. The correction derived from Copula-based mode regression reduces the RMSE, but results in an increase of the RMSE in some regions. The same holds true for the quantile mapping approach. This is in agreement with our previous results (see Fig. 10) that the Copula-based correction derived from the mean regression performs best. Therefore, in the following analyses, the results focus on the Copula-based mean regression approach.

To further assess the performance of the Copula-based method, additional performance measures are analyzed. The RMSE for different magnitudes of observed precipitation (i.e. a quantile RMSE analyses) is done for the selected four grid cells (see Fig. 1). The results from the validation period are shown in Fig. 17. The RMSE in different quantiles are represented by $RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$, while the subscript indicates the magnitude level. $RMSE_{0.1}$ evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From $RMSE_{0.2}$ to $RMSE_{1.0}$ the RMSE are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example, $RMSE_{1.0}$ indicates the errors in the magnitude of the 10% highest events. As it can be seen from Fig. 17, the Copula-based correction performs equally or even better in terms of the RMSE in most of the quantiles.

Furthermore, we also investigated the percentage of the corrected time steps that are closer to the observations compared to the quantile mapping method. The results are shown in Fig. 18. The values indicate the percentage of the successful corrections (i.e. closer to the observations) by the two bias correction methods. It can be seen that the results of the quantile mapping correction strongly depends on the rank correlation (see Fig. 15), while the Copula-based correction provides a stable correction efficiency over the entire domain. The average percentages of the successful correction are 55% for the Copula-based correction and 46% for the quantile mapping correction, respectively."

We shifted the text describing the comparison between the Copula-based and the quantile mapping approach from the end of section 4.3 to the new section 4.4. The comparison to the liner scaling method is removed now in the revised manuscript.

Please note that the Figure 16 in the revised manuscript is updated (compared to that in the answers to the reviewer's comments). We identified and corrected an error in our calculations of the RMSE.

Page 24, Line 25:
"When comparing to the quantile mapping correction, the Copula-based method has an improved performance in reducing the RMSE. It is also found that the Copula-based method allows for a better correction with respect to the percentage of the time steps that are closer to the observations after the corrections. The Copula-based method is able to provide a stable correction efficiency over the entire domain, even when the rank correlations between the RCM- and observed precipitation in some regions are low."

Page43:
"Figure 10. Relative bias map of mean daily precipitation for Copula-based correction by taking the expectation (left), median (middle) and the mode (right) as the estimator of the sampled distribution. The results are based on the validation period 1986-2000."

Page48:
"Figure 15. The rank correlations between RCM and REGNIE precipitation over the domain in the validation period from 1986 to 2000."

Page49:
"Figure 16. The changes of the RMSE in the validation period (1986-2000) by different bias correction methods. The green color indicates a decrease of the RMSE, while the ocher color implies an increase of the RMSE."

Page50:
"Figure 17. The 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 pixels are the same as in Fig. 13. The black solid line indicates the errors without correction. The results are derived from the validation period from 1986 to 2000."

Page51:
"Figure 18. The percentage of the corrections that are closer to the observations. Left: Copula-based correction (mean regression); Right: quantile mapping correction. The results are derived from the validation period from 1986 to 2000."

Comment 2

It is unclear for me why the authors did not use the normal copula for the description of the dependence. There is no reason to assume tail dependence, and there are no evaluations with respect to extremes.

Answer

Initially, we did not consider the Gaussian Copula in this study due to the following reason: The Gaussian Copula is able to describe a similar dependence structure, but shows slightly higher densities in the lower and upper tails compared to the Frank Copula. Based on findings from our previous works (Laux et al., 2011, & Vogl et al., 2012), we excluded the Gaussian Copula to be a suitable candidate in our study. We now followed the suggestions of the reviewer to include the Gaussian Copula as a potential candidate in our study. All the results are recalculated. The relative bias map (Fig. 11 & 12) in the revised manuscript are updated.

The following changes were made in the revised manuscript:

Page 16, Line 11:
"Four Copulas (Clayton, Frank, Gumbel and Gaussian) are investigated by applying the Goodness-of-fit tests described in section 3.4. Figure 7 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 study area, 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 72% of the grid cells is modelled by the Frank, 20% by the Clayton, 7% by the Gaussian and only 0.09% by the Gumbel Copula.

In order to assess for the annual variability of the dependence structures between REG- NIE and WRF precipitation time series, the Copula functions are identified for the different seasons separately. The corresponding results are shown in Fig. 8.

While for spring, autumn and winter the Copulas that have no pronounced tail dependence (the Frank and Gaussian Copula) dominate (spring 49% (Frank) + 22% (Gaussian) = 71%, autumn 53% + 24% = 77% and winter 63% + 28% = 91%), in summer the Clayton Copula provides the best fit for most of the grid cells (62%). For all seasons the Gumbel Copula is only selected for few grid cells with a maximum number of hits in spring (5% of the grid cells). In general the differences are most prominent for winter and summer (see Fig. 8)."

Page 30:
"Table 1. Theoretical Copula functions used in this study."

Page 40:
"Figure 7. Identified Copula functions between REGNIE and WRF precipitation in the calibration period (1971 to 1985) with positive pairs."

Page 41:
"Figure 8. 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)."

Page 44:
"Figure 11. 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."

Page 45:
"Figure 12. 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)."

Comment 3 & 4

It is not clear how the authors handle zero precipitation. Further how the problem of different zero precipitations for the model and the observations is treated? How is the distribution of the lower dry probability modified?

It is unclear to me why the authors did not use truncated copulas as suggested in Bárdossy and Pegram (2009) which is referenced in the paper. This approach could help to avoid several problems with the zeros.

Answer

The strategy and procedure how to handle the different cases, including zero precipitation is described now in more detail in section 3.6 of the revised manuscript.

The following changes were made in the revised manuscript:

Page 12, Line 20:
"The implementation of a bias correction for precipitation (a discrete variable) is more complex than a bias correction of continuous variable, e.g. temperature. In general four cases have to be distinguished, namely (0,0), (0,1), (1,0), and (1,1), where 0 denotes a dry day and 1 indicates a wet day (see Fig. 3). A threshold of rainfall amount of 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). Therefore, the four cases are defined as follows:

1. (1,1): REGNIE and WRF precipitation ≥ 0.1 mm
2. (0,1): REGNIE < 0.1 mm, while WRF ≥ 0.1 mm
3. (1,0): REGNIE ≥ 0.1 mm and WRF < 0.1 mm
4. (0,0): Both REGNIE and WRF < 0.1 mm

Different approaches exist in the literature to account for the intermittent nature of rainfall. For example the truncated Copula suggested in Bárdossy and Pegram (2009) and the Copula-based mixed model described in Serinaldi (2008). Both methods are able to produce time series that statistically hold the same proportion of the four different cases (0,0), (0,1), (1,0) and (1,1). These methods allow for the correction of the total number of dry days, but do not allow to correct individual events in the (0,1) and (1,0) cases.

In this study, we aim to an event-based correction as described in the following: The Copula-based concept focuses on the correction of the (1,1) cases, i.e. the positive pairs. 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 un- changed. For the (0,0) cases, there is no error. The errors that come from the (0,1) and (1,0) cases are not corrected by this method. To justify this strategy, we investigated the proportion of the four cases in the study area (see Fig. 4): The (1,1) cases take the highest proportion, followed by the (0,0) cases. The proportion of both the (0,1) and (1,0) cases are comparatively low. The average proportion of these cases are 40% for the (1,1) cases, 29% for the (0,0) cases, 19% for the (0,1) cases and 12% for the (1,0) cases, respectively."

Page 36:
"Figure 3. Illustration of the four cases: (0,0) indicates that both REGNIE and WRF show no rain, (0,1) stands for an observation with no precipitation but the RCM model shows a rain event, while (1,0) indicates the opposite of (0,1), (1,1) implies that both are wet."

Page 37:
"Figure 4. The proportion of the four cases over the study area for the validation time period (from 1986 to 2000)."

Comment 5

The fit of a parametric distribution followed by a fit of the copula based on the empirical distribution is statistically not correct.

Answer

We have now corrected this in our revised manuscript.

The following changes were made in the revised manuscript:

Page 10, Line 14:
"Let {r1(1), ..., r1(n)} and {r2(1), ..., r2(n)} denote the rank space values that are derived from the fitted theoretical marginal distributions."

Comment 6a

Is a parametric fit of the local precipitation distribution really needed? In my opinion the empirical distributions or a non-parametric fit would do a better job. Further this would avoid some problems with the spatial discontinuities imposed by taking different distributions.

Answer

The parametric fit of the local precipitation distribution is well-accepted in literature (please see e.g. Dupuis, 2007, Gao et al., 2007). It also allows us to provide spatially distributed differences, i.e. fitted marginal family maps between the WRF and the REGNIE data. This gives additional valuable information about the differences in their statistical properties.

The following changes were made in the revised manuscript:

Page 8, Line 23:
"Generally, both non-parametric and parametric fitting approaches for the local precipitation distribution are found in the literature (Dupuis, 2007; Gao et al., 2007; Bárdossy and Pegram, 2009; van den Berg et al., 2011). In this study a parametric fitting of the precipitation distribution is applied to allow also an illustration of the spatially distributed differences (provided as the fitted marginal distribution family maps) between WRF and REGNIE. This gives additional valuable information about their differences in statistical properties."

Comment 6b

The use of different copulas for the different locations is also causing spatial ruptures.

Answer

Using one common Copula to describe the dependence over the whole domain may not be appropriate. In some areas, this does not adequately reflect the underlying dependence structures since the dependence structure also depends e.g. on the topography and prevailing circulation patterns.

The following changes were made in the revised manuscript:

Page 9, Line 23:
"The Copulas from different families describe different dependence structures. To increase the accuracy of the description of the dependence, different types of Copulas are considered, since one common Copula might be incapable to capture the dependence structure for all grid cells over the entire study area and for all seasons."

Comment 6c

The method provides individual time series for each pixel. The simulated series are independent, thus there is no spatial coherence. This is a serious restriction of the suggested methodology, which can thus only be used for local and small scale studies.

Answer

Indeed, the method corrects the time series for each pixel independently. More details about how spatial patterns are preserved are now inserted in our revised manuscript.

The following changes were made in the revised manuscript:

Page 20, Line 5:
"Finally, in order to investigate the spatial coherence of the bias corrected precipitation fields, exemplarily the sequence of three selected days (from January 9th to 11th, 1986) are shown in Fig. 14. The left column from top to bottom are the observed precipitation fields for these three days. In the middle are the uncorrected (original) WRF simulated precipitation fields and the right column indicates the bias corrected WRF precipitation fields: While correcting the absolute precipitation values, the spatial coherence of the precipitation patterns are retained after the bias correction procedure."

Page 24, Line 18:
"By investigating the spatial coherence, the proposed method is found to be able to preserve the spatial structure of the WRF model output. This is due to the reason that the Copula-based approach is conditioned on the WRF simulation. The method is adjusting the value of the WRF precipitation according to the fitted Copula model. Even though the Copula models are estimated for each grid cell, the spatial coherence is captured by the Copula model as both the Copula families as well as the marginal distributions are also spatially clustered."

Page 47:
"Figure 14. Daily precipitation fields over Germany for the three consecutive days from January 9 to January 11, 1986."

Comment 7

The evaluation of the series is only concentrating on the mean behavior of the simulated series. There are no attempts to look at other statistics, such as variance, dry probability etc.

Answer

Besides the evaluation of the mean daily precipitation, we also analyzed mean monthly precipitation, the Root Mean Square Error (RMSE), the quantile RMSE and the percentage of the time steps that are closer to the observations after the application of our correction method.

The following changes were made in the revised manuscript:

Page 21, Line 14:
"The quantile mapping correction has been performed for comparison to the Copula-based approach. The RMSE between the observed (REGNIE) and bias corrected modelled data is calculated for both the Copula-based correction and the quantile mapping method. The original RMSE (between REGNIE and WRF) is also computed as a reference. For the Copula-based approach, we calculated the RMSE for all the simulations with respect to the mean-, median- and mode value. The changes of the RMSE by different corrections over the study area are shown in Fig. 16. The Copula-based correction derived from the mean regression reduces the RMSE significantly with an average number of −12% over the domain. The Copula-based correction derived from the median also reduces the RMSE, but to a less degree. The correction derived from Copula-based mode regression reduces the RMSE, but results in an increase of the RMSE in some regions. The same holds true for the quantile mapping approach. This is in agreement with our previous results (see Fig. 10) that the Copula-based correction derived from the mean regression performs best. Therefore, in the following analyses, the results focus on the Copula-based mean regression approach.

To further assess the performance of the Copula-based method, additional performance measures are analyzed. The RMSE for different magnitudes of observed precipitation (i.e. a quantile RMSE analyses) is done for the selected four grid cells (see Fig. 1). The results from the validation period are shown in Fig. 17. The RMSE in different quantiles are represented by $RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$, while the subscript indicates the magnitude level. $RMSE_{0.1}$ evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From $RMSE_{0.2}$ to $RMSE_{1.0}$ the RMSE are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example, $RMSE_{1.0}$ indicates the errors in the magnitude of the 10% highest events. As it can be seen from Fig. 17, the Copula-based correction performs equally or even better in terms of the RMSE in most of the quantiles.

Furthermore, we also investigated the percentage of the corrected time steps that are closer to the observations compared to the quantile mapping method. The results are shown in Fig. 18. The values indicate the percentage of the successful corrections (i.e. closer to the observations) by the two bias correction methods. It can be seen that the results of the quantile mapping correction strongly depends on the rank correlation (see Fig. 15), while the Copula-based correction provides a stable correction efficiency over the entire domain. The average percentage of the successful correction are 55% for the Copula-based correction and 46% for the quantile mapping correction, respectively.”

Page 24, Line 25:
“When comparing to the quantile mapping correction, the Copula-based method has an improved performance in reducing the RMSE. It is also found that the Copula-based method allows for a better correction with respect to the percentage of the time steps that are closer to the observations after the corrections. The Copula-based method is able to provide a stable correction efficiency over the entire domain, even when the rank correlations between the RCM- and observed precipitation in some regions are low.”

Comment 8

The bias remaining after the correction is very high. It would be interesting to know how the observed and modelled precipitation itself was changing. Was the signal captured?

Answer

In our study, after the correction the average relative bias of daily mean precipitation from WRF for the validation period is reduced from 10 % to −1 %. In different seasons, the average bias are reduced from 32% to 16% in spring (MAM), from -15% to -11% in summer (JJA), from 4% to -1% in autumn (SON) and from 28% to -3% in winter (DJF). The results show that the biases are corrected efficiently especially in winter. To investigate typical situations in detail, we also evaluated the monthly mean precipitation for four selected grid cells. It is also shown that the biases are significantly reduced.
We are not sure, whether or not the question of the reviewer is directed towards climate trends. We do not address climate trends in this study. The analyses of climate trends are beyond the scope of this study.

No changes in the revised manuscript are made.

Comment 9

The possibly biggest problem with the method is its partial inability to reflect the signal of the meteorological model. The weaker the dependence between model and observations the less the model signal is reflected after bias correction.
This is illustrated with a small example. I simulated 500 realizations of modelled precipitation with a mean of 2mm using an exponential distribution. The observed precipitation has a mean of 3 mm and follows an exponential distribution. Copulas with different degrees of dependence ranging from full dependence to independence were used.

It is assumed that the model shows a precipitation increase of 50%. The corresponding bias corrected series were simulated, and the means were compared to the original mean (3mm for the observations). The increase in precipitation varied between 0 and 50% depending on the degree of dependence. The weaker the dependence the smaller is the signal which is captured. This is unfortunately not in the sense of bias correction, where the signal should be reflected.

Answer

In order to analyze the signal-dependence relationship and inspired by the principle idea of the reviewer's example calculation, we followed the example and repeated the calculations. Apart from the results obtained by the reviewer, we found that our method is able to reflect the signal. The results are shown in the answers to the reviewer's comments. We decided not to include them in the final revised version.

Comment 10

Remark: The quantile/quantile transformation can be regarded as a special case of the suggested methodology with a fully dependent copula (rank correlation equal to one).

Answer

It is true that the quantile mapping method can be regarded as a special case of the proposed methodology as it corrects all the moments of modelled precipitation distribution under the assumption of the full dependence between the ranks. However, this full dependence assumption is limited. In our study area, the rank correlation between the datasets varies between 0.3 and 0.6. By comparing to the quantile mapping correction, the RMSE and the percentage of the corrected time steps that is closer to the observations are analyzed.

The following changes were made in the revised manuscript.

Page 21, Line 14:
"The quantile mapping correction has been performed for comparison to the Copula-based approach. The RMSE between the observed (REGNIE) and bias corrected modelled data is calculated for both the Copula-based correction and the quantile mapping method. The original RMSE (between REGNIE and WRF) is also computed as a reference. For the Copula-based approach, we calculated the RMSE for all the simulations with respect to the mean-, median- and mode value. The changes of the RMSE by different

corrections over the study area are shown in Fig. 16. The Copula-based correction derived from the mean regression reduces the RMSE significantly with an average number of −12% over the domain. The Copula-based correction derived from the median also reduces the RMSE, but to a less degree. The correction derived from Copula-based mode regression reduces the RMSE, but results in an increase of the RMSE in some regions. The same holds true for the quantile mapping approach. This is in agreement with our previous results (see Fig. 10) that the Copula-based correction derived from the mean regression performs best. Therefore, in the following analyses, the results focus on the Copula-based mean regression approach.

To further assess the performance of the Copula-based method, additional performance measures are analyzed. The RMSE for different magnitudes of observed precipitation (i.e. a quantile RMSE analyses) is done for the selected four grid cells (see Fig. 1). The results from the validation period are shown in Fig. 17. The RMSE in different quantiles are represented by $RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$, while the subscript indicates the magnitude level. $RMSE_{0.1}$ evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From $RMSE_{0.2}$ to $RMSE_{1.0}$ the RMSE are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example, $RMSE_{1.0}$ indicates the errors in the magnitude of the 10% highest events. As it can be seen from Fig. 17, the Copula-based correction performs equally or even better in terms of the RMSE in most of the quantiles.

Furthermore, we also investigated the percentage of the corrected time steps that are closer to the observations compared to the quantile mapping method. The results are shown in Fig. 18. The values indicate the percentage of the successful corrections (i.e. closer to the observations) by the two bias correction methods. It can be seen that the results of the quantile mapping correction strongly depends on the rank correlation (see Fig. 15), while the Copula-based correction provides a stable correction efficiency over the entire domain. The average percentage of the successful correction are 55% for the Copula-based correction and 46% for the quantile mapping correction, respectively."

Page 24, Line 25:

"When comparing to the quantile mapping correction, the Copula-based method has an improved performance in reducing the RMSE. It is also found that the Copula-based method allows for a better correction with respect to the percentage of the time steps that are closer to the observations after the corrections. The Copula-based method is able to provide a stable correction efficiency over the entire domain, even when the rank correlations between the RCM- and observed precipitation in some regions are low."

Referee Comments 2

General Comments

The authors have submitted a paper in which precipitation fields are rescaled based on copulas. While the idea is interesting (something similar was already published in HESS by van den Berg et al., 2011), it does not really make use of the advantages of copulas. Basically, van den Berg et al. perform a rescaling where the intra-pixel variability in a coarse scale observation is estimated based on copulas (leading to the pdf of the expected rainfall within one coarse scale pixel). In this paper, more or less the same is done, except that the mean or median of the pdf is used as estimation of precipitation value. So, the total information of the pdf is lost, and actually, one obtains a relation between the expected mean (or median) and an observation (the copula in- deed corrects for bias). Actually, what is thus obtained is a kind of non-linear regression between REGNIE and (bias-corrected) WRF. My question then is: why should one go through copulas and shouldn't any kind of non-linear regression be done immediately? Why do the authors throw away all information within the conditional CDF obtained from the copula? Why did the authors rescale the REGNIE to 7 km and didn't they keep the 1 km scale (and then follow van den Berg et al.), as it is perfectly possible to compare different scales using the proposed framework.

Comment 1

Why should one go through copulas and shouldn't any kind of non-linear regression be done immediately?

Answer

The non-linear regression is a bijection method. The Copula-based method is a stochastic bias correction method. It provides a full PDF of the corrected value for each time step. For practical reasons and the typical needs of subsequent modellers, we take the mean as estimator of the derived PDF for each time step. Different to the non-linear regression, this mean regression curve, however, is derived based on the dependence between RCM and observations. There are two advantages compared to the non-linear regression: 1) The Copula-based regression provides a higher flexibility since it can be used with different combination of marginals and Copulas; 2) The Copula-based regressions predicts the full ensemble of the possible outcomes which gives an additional quality criterion of the bias correction.

The following changes were made in the revised manuscript:

Page 25, Line 8:
"Apart from traditional approaches such as the quantile mapping, which is based on a bijection transfer function, the Copula-based stochastic bias correction technique provides the information of the full PDF for each individual time step. This additionally provides a quality criterion for the bias correction, e.g. expressed as the spread of the PDF in form of the interquantile range. Subsequent modellers using RCM derived precipitation data are potentially enabled to make use of the full PDF, especially if they are interested in other statistical moments or in estimating uncertainties arising from this approach."

Comment 2

Why do the authors throw away all information within the conditional CDF obtained from the copula?

Answer

Our approach calculates and stores the full information of the sampled PDF for each time step. Thus, we do not throw away the information about the distributions. Due to the practical reasons, we take the mean as the estimator of the sampled PDF.

The following changes were made in the revised manuscript:

Page 17, Line 10:
"While this stochastic bias correction method gives a full ensemble and the empirical predictive distribution of corrected WRF precipitation, for practical reasons one can choose e.g., the expectation, median or mode to get single corrected values. This can be regarded as a Copula-based regression by taking such a "typical" value as the estimator of the derived empirical predictive distribution of corrected WRF precipitation. It is noted that this "typical" value (e.g. the mean) can also be directly derived from the analytical Copula-based conditional CDF (Mean Regression Curve, Nelsen, 1999)."

Page 25, Line 12:
"Subsequent modellers using RCM derived precipitation data are potentially enabled to make use of the full PDF, especially if they are interested in other statistical moments or in estimating uncertainties arising from this approach."

Comment 3

Why did the authors rescale the REGNIE to 7 km and didn't they keep the 1 km scale (and then follow van den Berg et al.), as it is perfectly possible to compare different scales using the proposed framework.

Answer

The aim of our study is the bias correction of WRF precipitation fields instead of a downscaling. The REGNIE data was aggregated from 1 km to 7 km to achieve the same resolution as WRF.

The following changes were made in the revised manuscript:

Page 5, Line 10:

“To achieve the same resolution, the REGNIE data are aggregated to 7 km.”

Specific Comments

Comment 1

page 7195, line 15: why only consider parametric functions?

Answer

Please see the response to General Comment #6a of reviewer #1.

Comment 2

Figure 2 shows the density, while the paper is referring to cumulative probabilities (so c versus C)

Answer

The Figure 2 has been updated now in the revised manuscript.

The following changes were made in the revised manuscript:

Page 35:
“Figure 2. Visualisation of a bivariate Copula model consisting of two marginal distributions and a theoretical Copula function that describes the pure dependence.”

Comment 3

page 7198, lines 24-25: it is not clear to me why step 6 is needed as in the end only the mean (or median) is used: this could be determined right away instead of first a random set from which the mean (or median) is calculated.

Answer

The Copula-based correction is a stochastic method. Repeating the step 6 and step 7 can be regarded as the Monte Carlo simulation to obtain the random realizations. We take the mean of the realization as the estimation of the corrected value regarding to the practical reasons, e.g. spatial illustration. We stated now more clearly in the revised manuscript.

The following changes were made in the revised manuscript:

Page 12, Line 15:
"This provides the possibility to obtain a large set of random realizations and additionally gives the information of a probability density function (PDF) for each corrected time step. From the PDF the spread of the distribution in form of the interquantile range can e.g. be provided as an additional uncertainty criterion for the bias correction."

Page 17, Line 10:
"While this stochastic bias correction method gives a full ensemble and the empirical predictive distribution of corrected WRF precipitation, for practical reasons one can choose e.g., the expectation, median or mode to get single corrected values. This can be regarded as a Copula-based regression by taking such a "typical" value as the estimator of the derived empirical predictive distribution of corrected WRF precipitation. It is noted that this "typical" value (e.g. the mean) can also be directly derived from the analytical Copula-based conditional CDF (Mean Regression Curve, Nelsen, 1999)."

Comment 4

page 7199, last paragraph: operationally, you should also apply it to the (0,1) case, as you have no idea whether the REGNIE-data is 0 or 1.

Answer

The implementation of a bias correction for precipitation, in general, needs to investigate four cases, i.e. (0,0), (0,1), (1,0), and (1,1). Our strategy to deal with these four cases is now described in more detailed in the revised manuscript.

The following changes were made in the revised manuscript.

Page 13, Line 9:
" Different approaches exist in the literature to account for the intermittent nature of rainfall. For example the truncated Copula suggested in Bárdossy and Pegram (2009) and the Copula-based mixed model described in Serinaldi (2008). Both methods are able to produce time series that statistically hold the same proportion of the four different cases (0,0), (0,1), (1,0) and (1,1). These methods allow for the correction of the total number of dry days, but do not allow to correct individual events in the (0,1) and (1,0) cases.
  In this study, we aim to an event-based correction as described in the following: The Copula-based concept focuses on the correction of the (1,1) cases, i.e. the positive pairs. 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 un- changed. For the (0,0) cases, there is no error. The errors that come from the (0,1) and (1,0) cases are not corrected by this method. To justify this strategy, we investigated the proportion of the four cases in the study area (see Figure 4): The (1,1) cases take the highest proportion, followed by the (0,0) cases. The proportion of both the (0,1) and (1,0) cases are comparatively low. The average proportion of these cases are 40% for the (1,1) cases, 29% for the (0,0) cases, 19% for the (0,1) cases and 12% for the (1,0) cases, respectively."

Comment 5

page 7199: what are the proportions of (0,0), (1,1), (0,1) and (1,0) within the data used?, Is it also possible to show a scatterplot of observations versus RCM?

Answer

The proportions of (0,0), (0,1), (1,0) and (1,1) cases are calculated. The results are shown and discussed in the revised manuscript.
We were analyzing scatterplots for the (1,1) case at four different grid cells (please see answers to the reviewer's comments), but decided not to include them in the final revised version.

The following changes were made in the revised manuscript.

Page 13, Line 23:
"To justify this strategy, we investigated the proportion of the four cases in the study area (see Figure 4): The (1,1) cases take the highest proportion, followed by the (0,0) cases. The proportion of both the (0,1) and (1,0) cases

are comparatively low. The average proportion of these cases are 40% for the (1,1) cases, 29% for the (0,0) cases, 19% for the (0,1) cases and 12% for the (1,0) cases, respectively."

Page 37:
"Figure 4. The proportion of the four cases over the study area for the validation time period (from 1986 to 2000)."

Comment 6

page 7202", line 7: the Frank copula also allows to model negative dependence.

Answer

We stated this more clearly in the revised manuscript.

The following changes were made in the revised manuscript:

Page 10, Line 4:
"For the Clayton Copula, the formulas of positive and negative dependence are different. The parameter θ can take values $-1 < \theta < 0$ to model negative dependence. In that case the formula is
$C_\theta(u,v) = [\max(u^{-\theta} + v^{-\theta} - 1,0)]^{-1/\theta}$
For the data used in this study, negative dependences could not be found. Therefore it is θ□[0,∞] and the Clayton Copula is defined as described in Table 1."
Please note that this part has been moved from section 4.2 to section 3.4 and modified.

Comment 7

page 7203 line 10: here you actually could do the same with a non-linear regression

Answer

Please see the response to General Comment #1 of reviewer #2.

Comment 8

page 7203, line 18: does this hold if you look at the blue line?

Answer

For most of the time steps the blue line (corrected WRF precipitation) is indeed closer to REGNIE precipitation compared to the raw WRF precipitation. We also calculated the overall proportion of successful corrections (the time steps that are closer to the observations after correction) for the whole study area.

The following changes were made in the revised manuscript:

Page 22, Line 12:
"Furthermore, we also investigated the percentage of the corrected time steps that are closer to the observations compared to the quantile mapping method. The results are shown in Fig. 18. The values indicate the percentage of the successful corrections (i.e. closer to the observations) by the two bias correction methods. It can be seen that the results of the quantile mapping correction strongly depends on the rank correlation (see Fig. 15), while the Copula-based correction provides a stable correction efficiency over the entire domain. The average percentage of the successful correction are 55% for the Copula-based correction and 46% for the quantile mapping correction, respectively."

Additional changes which are not related to the suggestions of the reviewers are:

Page 4, Line 15:
"In this study, a Copula-based stochastic bias correction method is applied to correct each individual time step of a RCM simulation. This is different to the traditional transfer function based statistical correction approaches. The strategy of this method is the identification and description of the underlying dependence structures between observed and modeled climate variables (here: precipitation) and its application for bias correction. It is known that the traditional measures of dependence (e.g. Pearson's correlation coefficient) can only capture the strength of the linear dependence as a single global parameter. Alternatively, Copulas are able to describe the complex non-linear dependence structure between variables (Bárdossy and Pegram, 2009). Based on the identified dependence structure between observed and modeled precipitation and the identified respective marginal distributions, a set of realizations is finally obtained through Monte Carlo simulations. Recently, Copulas are used for various applications in hydrometeorology (e.g., Gao et al., 2007; Serinaldi, 2008; van den Berg et al., 2011; Bárdossy

and Pegram, 2012)."

Page 8, Line 20:
"3.3 Marginal distribution estimation"

Page 9, Line 22:
"3.4 Copula function estimation"

Page 9, Line 27:
"In this study, four different one-parametric Copulas (see Table 1) are selected: The Gumbel Copula is able to describe an upper tail dependence structure, while the Clayton Copula allows to express higher probability in the lower tail. The Frank Copula exhibits no tail dependence, and the Gaussian Copula describes a similar dependence as the Frank Copula, but with slightly higher densities in the lower and upper tails (Venter, 2002; Schmidt, 2007)."

Page 22, Line 20:
"5 Summary and conclusions"

Page 22, Line 22:
"The strategy of this method is the identification and description of underlying dependence structures between RCM and observed precipitation and its application for bias correction."

Page 23, Line 6:
"The conditional distribution derived from fitted Copula model forms the basis of the correction procedure. It provides the possibility to access all the possible outcomes of the corrected value and additionally gives the information of a PDF for each corrected time step."

Page 24, Line 11:
"For the investigation of the spatial performance, the Copula correction based on the mean value is applied."

Manuscript prepared for Hydrol. Earth Syst. Sci. Discuss.
with version 2014/07/29 7.12 Copernicus papers of the LaTeX class copernicus.cls.
Date: 24 November 2014

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

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

Abstract

Dynamically downscaled precipitation fields from regional climate model (RCM) often cannot be used directly for ~~local~~ regional 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 is 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~~ aggregated to 7km 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 (described by the Copula function) between WRF and REGNIE data and the identified respective marginal distributions in the calibration period, separately analyzed for the different seasons, conditional distribution functions are derived for each time step in the 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, the Copula-based approach is compared ~~with~~

~~linear scaling and~~ to the quantile mapping correction ~~approaches~~ method. ~~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.~~ The Root Mean Square Error (RMSE) and the percentage of the corrected time steps that are closer to the observations are analyzed. The Copula-based correction derived from the mean of the sampled distribution reduces the RMSE significantly, while e.g. the quantile mapping method results in an increased RMSE for some regions.

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 (Graham et al., 2007a, b; Christensen et al., 2008; Bergström et al., 2001). 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 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).~~

In this study, a Copula-based stochastic bias correction method is applied to correct each individual time step of a RCM simulation. This is different to the traditional transfer function based statistical correction approaches. The strategy of this method is the identification and description of the underlying dependence structures between observed and modeled climate variables (here: precipitation) and its application for bias correction. It is known that the traditional measures of dependence (e.g. Pearson's correlation coefficient) can only capture the strength of the linear dependence as a single global parameter. Alternatively, Copulas are able to describe the complex non-linear dependence structure between variables (Bárdossy and Pegram, 2009). Based on the identified dependence structure between observed and modeled precipitation and the identified respective marginal distributions, a set of realizations is finally obtained through Monte Carlo simulations.

Recently, Copulas are used for various applications in hydrometeorology (e.g., Gao et al., 2007; Serinaldi, 2008; van den Berg et al., 2011; Bárdossy and Pegram, 2012). 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~~ individual and unevenly distributed stations. The Copula models are estimated for each single 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 ~~precipiation~~ 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. To achieve the same resolution, the REGNIE data are aggregated to 7 km. 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 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~~ summary 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 function C (Copula) such that:

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

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_θ (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 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~~Marginal distribution estimation

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. Generally, both non-parametric and parametric fitting approaches for the local precipitation distribution are found in the literature (Dupuis, 2007; Gao et al., 2007; Bárdossy and Pegram, 2009; van den Berg et al., 2011). In this study a parametric fitting of the precipitation distribution is applied to allow also an illustration of the spatially distributed differences (provided as the fitted marginal distribution family maps) between WRF and REGNIE. This gives additional valuable information about differences in their statistical properties.

In this study, five different parametric 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:

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

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.

3.4 ~~Copula Goodness-of-fit test~~Copula function estimation

The Copulas from different families describe different dependence structures. To increase the accuracy of the description of the dependence, different types of Copulas are considered, since one common Copula might be incapable to capture the dependence structure for all grid cells over the entire study area and for all seasons.

In this study, four different one-parametric Copulas (see Table 1) are selected: the Gumbel Copula is able to describe an upper tail dependence structure, while the Clayton Copula allows to express higher probability in the lower tail. The Frank Copula exhibits no tail dependence, and the Gaussian Copula describes a similar dependence as the Frank Copula, but with slightly higher densities in the lower and upper tails (Venter, 2002; Schmidt, 2007). For the Clayton Copula, the formulas of positive and negative dependence are different. The parameter θ can take values $-1 < \theta < 0$ to model negative dependence. In that case the formula is

$$C_\theta(u,v) = [max(u^{-\theta} + v^{-\theta} - 1, 0)]^{-\frac{1}{\theta}}. \quad (3)$$

For the data used in this study, negative dependences could not be found. Therefore it is $\theta \in [0, \infty]$ and the Clayton Copula is defined as described in Table 1.

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.~~ the rank space values that are derived from the fitted theoretical marginal distributions. Then, the empirical Copula, a rank based estimator of C_θ, is defined as:

$$C_n(u,v) = 1/n \sum_{t=1}^{n} \mathbf{1}\left(\frac{r_1(t)}{n} \leqslant u, \frac{r_2(t)}{n} \leqslant v\right) \quad (4)$$

with $u = F_X(x)$, $v = F_Y(y)$ and $\mathbf{1}(\ldots)$ is denoting the indicator function and n being the sample size. 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_θ:

$$S_n = \frac{1}{n}\sum_{t=1}^{n}\{C_\theta(u_t,v_t) - C_n(u_t,v_t)\}^2 \tag{5}$$

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

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~~ can be generated through Monte Carlo simulations (Gao et al., 2007; 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 conditional simulation algorithm is based on a conditional ~~probabilities~~ distribution of the form:

$$C_{U|V=v}(u) = P[U \leq u | V = v] = \frac{\partial C(u,v)}{\partial v} \tag{6}$$

The complete Copula-based bias correction 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 θ 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

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

The Copula-based conditional ~~prediction~~ simulation 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). This provides the possibility to obtain a large set of random realizations and additionally gives the information of a probability density function (PDF) for each corrected time step. From the PDF the spread of the distribution in form of the interquantile range can e.g. be provided as an additional uncertainty criterion for the bias correction.

3.6 Correction strategy for continous time series

The implementation of a bias correction for ~~precipiation~~ precipitation (a discrete variable) is more complex than a bias correction of a continuous ~~variables~~ variable, 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~~ where 0 denotes a dry day and 1 indicates a wet day (see Fig. 3). A threshold of rainfall amount of 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). Therefore, the four cases are defined as follows:

1. (1,1): REGNIE and WRF precipitation $\geq$ 0.1 mm
2. (0,1): REGNIE $<$ 0.1 mm, while WRF $\geq$ 0.1 mm
3. (1,0): REGNIE $\geq$ 0.1 mm and WRF $<$ 0.1 mm
4. (0,0): Both REGNIE and WRF $<$ 0.1 mm

Different approaches exist in the literature to account for the intermittent nature of rainfall. For example the truncated Copula suggested in Bárdossy and Pegram (2009) and the Copula-based mixed model described in Serinaldi (2008). Both methods are able to produce time series that statistically hold the same proportion of the four different cases (0,0), (0,1), (1,0), and (1,1). These methods allow for the correction of the total number of dry days, but do not allow to correct individual events in the (0,1) and (1,0) cases.

~~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 this study, we aim to an event-based correction as described in the following: the Copula-based concept focuses on the correction of the (1,1) cases, i.e. the positive pairs. 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.~~ For the (0,0) cases, there is no error. The errors that come from the (0,1) and (1,0) cases are not corrected by this method. To justify this strategy, we investigated the proportion of the four cases in the study area (see Fig. 4): the (1,1) cases take the highest proportion, followed by the (0,0) cases. The proportion of both the (0,1) and (1,0) cases are comparatively low. The average proportion of these cases are

40% for the (1,1) cases, 29% for the (0,0) cases, 19% for the (0,1) cases and 12% for the (1,0) cases, respectively.

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. 5). 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 in Table 2. 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. 5).

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. 6. 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 3. 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 is reduced in seasonal analysis, the failures of K–S test are decreased dramatically. The results are shown in Table 4.

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 1) are investigated by application of the Goodness-of-fit tests described in Sect. 3.4. For the Clayton Copula the parameter θ can also take values $-1<\theta<0$ to model negative dependence. In that case it is~~

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

~~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 1.~~ Four Copulas (Clayton, Frank, Gumbel and Gaussian) are investigated by applying the Goodness-of-fit tests described in Sect. 3.4. Figure 7 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~~ study area, 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 %~~ 72 % of the grid cells is modelled by the Frank, ~~21 %~~ 20 % by the Clayton, 7 % by the Gaussian 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. 8.

While for spring, autumn and winter the ~~Frank Copula~~ Copulas that have no pronounced tail dependence (the Frank and Gaussian Copula) dominate (~~spring 66 %, autumn 74 % and winter 88 %~~ spring 49 % (Frank) + 22 % (Gaussian) = 71 %, autumn 53 % + 24 % = 77 % and winter 63 % + 28 % = 91 %), in summer the Clayton Copula provides the best fit for most of the grid cells (~~64 %~~ 62 %). For all seasons the Gumbel Copula is only selected for few grid cells with ~~spring being the season of most hits~~ a maximum number of hits in spring (~~7 %~~

5 % of the grid cells). In general the differences are most prominent for winter and summer (see Fig. 8).

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~~ practical reasons one can choose e.g., the expectation, median or mode to get ~~a~~ single corrected ~~value~~ values. This can be regarded as a Copula-based regression by taking such a "typical" value as the estimator of the derived empirical predictive distribution of corrected WRF precipitation. It is noted that this "typical" value (e.g. the mean) can also be directly derived from the analytical Copula-based conditional CDF (mean regression curve, Nelsen, 1999).

Figure 9 exemplarily 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. 9 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 the expectation, median and mode of the predictive distribution with observations indicates that the correction performs best for the expectation value (see Fig. 10). Both simulations based on the median and the mode tend to underestimate the precipitation values, thus causing a dry bias over the domain. Therefore in the following, the results are shown and analyzed for ~~that case~~ the mean only.

Figure 11 (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. 11, 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. 12. It shows that the relative bias is even larger for different seasons. Figure 12 (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 12 (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 12 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. 12).

Figure 13 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. 13 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 performance of WRF was already satisfactory, the algorithm was still able to further improve the results.

Finally, in order to investigate the spatial coherence of the bias corrected precipitation fields, the sequence of three selected days (from January 9th to 11th, 1986) are exemplarily shown in Figure 14. The left column from top to bottom are the observed precipitation fields for these three days. In the middle are the uncorrected (original) WRF simulated precipitation fields and the right column indicates the bias corrected WRF precipitation fields: while correcting the absolute precipitation values, the spatial coherence of the precipitation patterns are retained after the application of the bias correction.

~~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. 13. The result from the validation period is shown in Fig. 17 and similar results are found for the entire study area. The RMSE in different quantiles are represented by~~ ~~$RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$~~~~, while the subscript indicates the magnitude level.~~ ~~$RMSE_{0.1}$~~ ~~evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From~~ ~~$RMSE_{0.2}$~~ ~~to~~ ~~$RMSE_{1.0}$~~ ~~the root mean square errors are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example,~~ ~~$RMSE_{1.0}$~~ ~~indicates the errors in the magnitude of the most extreme events.~~

~~As it can be seen from Fig. 17, 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 ($\mathrm{RMSE}_{1.0}$). As the Copula-based correction is only applied for (1,1) cases, the (0,1) errors are not corrected (Fig. 17, $\mathrm{RMSE}_{0.1}$). The results show that for linear scaling and the quantile mapping correction, the (0,1) errors are also not corrected.~~

4.4 Comparison of Copula-based bias correction to the quantile mapping method

The quantile mapping method is often used in bias correction of RCM derived precipitation (Dosio and Paruolo, 2011; Gudmundsson et al., 2012). This method corrects the bias by rescaling the values of the RCM so that the distribution of the RCM matches that of the observations. It corrects all moments of the RCM precipitation distribution under the assumption of a perfect dependence among the ranks. This full dependence assumption is limited: in our study area, the rank correlation between the datasets varies between 0.3 and 0.6 (see Fig. 15).

The quantile mapping correction has been performed for comparison to the Copula-based approach. The RMSE between the observed (REGNIE) and bias corrected modelled data is calculated for both the Copula-based correction and the quantile mapping method. The original RMSE (between REGNIE and WRF) is also computed as a reference. For the Copula-based approach, we calculated the RMSE for all the simulations with respect to the mean-, median- and mode value. The changes of the RMSE by different corrections over the study area are shown in Fig. 16. The Copula-based correction derived from the mean regression reduces the RMSE significantly with an average of -12% over the domain. The Copula-based correction derived from the median also reduces the RMSE, but to a less degree. The correction derived from Copula-based mode regression reduces the RMSE, but results in an increase of the RMSE in some regions. The same holds true for the quantile mapping approach. This is in agreement with our previous results (see Fig. 10) that the Copula-based correction derived from the mean regression performs best. Therefore, in the following analyses, the results focus on the Copula-based mean regression approach.

To further assess the performance of the Copula-based method, additional performance measures are analyzed. The RMSE for different magnitudes of observed precipitation (i.e. a quantile RMSE analysis) is done for the selected four grid cells (see Fig. 1). The results from the validation period are shown in Fig. 17. The RMSE in different quantiles are represented by $\mathrm{RMSE}_{0.1}$, $\mathrm{RMSE}_{0.2}$, ..., $\mathrm{RMSE}_{1.0}$, while the subscript indicates the magnitude level. $\mathrm{RMSE}_{0.1}$ evaluates the errors in the dry part of the observation distribution, implying the (0,1) errors. From $\mathrm{RMSE}_{0.2}$ to $\mathrm{RMSE}_{1.0}$ the RMSE are calculated for equally spaced probability intervals of the observed empirical CDF of wet days. For example, $\mathrm{RMSE}_{1.0}$ indicates the errors in the magnitude of the 10% highest events. As it can be seen from Fig. 17, the Copula-based correction performs equally or even better in terms of the RMSE in most of the quantiles.

Furthermore, we also investigated the percentage of the corrected time steps that are closer to the observations compared to the quantile mapping method. The results are shown in Fig. 18. The values indicate the percentage of the successful corrections (i.e. closer to the observations) by the two bias correction methods. It can be seen that the results of the quantile mapping correction strongly depends on the rank correlation (see Fig. 15), while the Copula-based correction provides a stable correction efficiency over the entire domain. The average percentages of the successful correction are 55% for the Copula-based correction and 46% for the quantile mapping correction, respectively.

5 ~~Discussion and conclusions~~Summary and conclusions

In this study, a Copula-based stochastic bias correction technique for RCM- output is introduced. The strategy of this method is the identification and description of underlying dependence structures between RCM and observed precipitation and its application for bias correction. 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~~ flexibility to construct a correction model by combining different marginal distributions and Copula functions, as many parametric univariate distribution and theoretical Copulas are available.

The conditional distribution derived from fitted Copula model forms the basis of the correction procedure. It provides the possibility to access all the possible outcomes of the corrected value and additionally gives the information of a PDF for each corrected time step.

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.

- 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. 5 and 6)~~. 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 Fig. 7 and Fig. 8)~~. 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~~ assumption of this approach is that the dependence structure between observed and modelled precipitation is stationary over the period of interest. For the investigation of the spatial performance, the Copula correction based on the mean value is applied. 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 Fig. 11 and Figure 12)~~. It is also found that the correction method performs better for overestimation ~~correction~~ rather than for underestimation ~~correction~~. ~~By applying a specific analyses with four specific grid cells, results show that the Copula-based correction provides better results than linear scaling and quantile mapping.~~ By investigating the spatial coherence, the proposed method is found to be able to preserve the spatial structure of the WRF model output. This is due to the reason that the Copula-based approach is conditioned on the WRF simulation. The method is adjusting the value of the WRF precipitation according to the fitted Copula model. Even though the Copula models are estimated for each grid cell, the spatial coherence is captured by the Copula model as both the Copula families as well as the marginal distributions are also spatially clustered.

When comparing to the quantile mapping correction, the Copula-based method has an improved performance in reducing the RMSE. It is also found that the Copula-based method allows for a better correction with respect to the percentage of the time steps that are closer to the observations after the correction. The Copula-based method is able to provide a stable correction efficiency over the entire domain, even if the rank correlations between the RCM- and observed precipitation are low.

~~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.~~

Apart from traditional approaches such as the quantile mapping, which is based on a bijection transfer function, the Copula-based stochastic bias correction technique provides the information of the full PDF for each individual time step. This additionally provides a quality criterion for the bias correction, e.g. expressed as the spread of the PDF in form of the interquantile range. Subsequent modellers using RCM derived precipitation data are potentially enabled to make use of the full PDF, especially if they are interested in other statistical moments or in estimating uncertainties arising from this approach.

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. Finally, we acknowledge support from the Deutsche Forschungsgemeinschaft for the Open Access Publishing Fund of Karlsruhe Institute of Technology.

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, Climate Research, 34, 211–222, doi:10.3354/cr00703, 2007.

Bárdossy, A. and Pegram, G.: Multiscale spatial recorrelation of RCM precipitation to produce unbiased climate change scenarios over large areas and small, Water Resources Research, 48, 1–13, doi:10.1029/2011WR011524, 2012.

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

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

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, Climate Research, 16, 101–112, doi:10.3354/cr016101, 2001.

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, Geophysical Research Letters, 35, L20 709, doi:10.1029/2008GL035694, 2008.

Deheuvels, P.: La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance, Académie Royale de Belgique, Bulletin de la Classe des Sciences, 65, 274–292, 1979.

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

Dosio, A. and Paruolo, P.: Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate, Journal of Geophysical Research, 116, D16 106, doi:10.1029/2011JD015934, 2011.

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

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

Gao, H., Wood, E. F., Drusch, M., and McCabe, M. F.: Copula-derived observation operators for assimilating TMI and AMSR-E retrieved soil moisture into land surface models, Journal of Hydrometeorology, 8, 413–429, doi:10.1175/JHM570.1, 2007.

Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, Journal of Hydrologic Engineering, pp. 347–368, doi:10.1061/(ASCE)1084-0699(2007)12:4(347), 2007.

Genest, C. and Rémillard, B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 44, 1096–1127, doi:10.1214/07-AIHP148, 2008.

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

Ghosh, S. and Mujumdar, P. P.: Climate change impact assessment: Uncertainty modeling with imprecise probability, Journal of Geophysical Research, 114, D18 113, doi:10.1029/2008JD011648, 2009.

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.

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.

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

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrology and Earth System Sciences, 16, 3383–3390, doi:10.5194/hess-16-3383-2012, 2012.

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

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, Hydrology and Earth System Sciences, 8, 1031–1045, doi:10.5194/hess-8-1031-2004, 2004.

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

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

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

Massey, F. J.: The Kolmogorov-Smirnov test for goodness of fit, Journal of the American statistical Association, 46, 68–78, doi:10.2307/2280095, 1951.

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

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

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, Journal of Hydrometeorology, 14, 1175–1193, doi:10.1175/JHM-D-12-091.1, 2013.

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

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

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

Schmidt, T.: Coping with copulas, in: Copulas - From Theory to Application in Finance, 15, pp. 3–34, Risk Books, London, 2007.

Serinaldi, F.: Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective, Stochastic Environmental Research and Risk Assessment, 23, 677–693, doi:10.1007/s00477-008-0249-z, 2008.

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.

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.

Smiatek, G., Kunstmann, H., Knoche, R., and Marx, A.: Precipitation and temperature statistics in high-resolution regional climate models: Evaluation for the European Alps, Journal of Geophysical Research, 114, D19 107, doi:10.1029/2008JD011353, 2009.

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, Hydrology and Earth System Sciences, 14, 687–703, doi:10.5194/hess-14-687-2010, 2010.

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

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

Uppala, S., Kå llberg, 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, Quart. J. R. Meteorol. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

van den Berg, M. J., Vandenberghe, S., De Baets, B., and Verhoest, N. E. C.: Copula-based downscaling of spatial rainfall: a proof of concept, Hydrology and Earth System Sciences, 15, 1445–1457, doi:10.5194/hess-15-1445-2011, 2011.

Venter, G. G.: Tails of copulas, Proceedings of the Casualty Actuarial Society, LXXXIX, 68–113, 2002.

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, Hydrology and Earth System Sciences, 16, 2311–2328, doi:10.5194/hess-16-2311-2012, 2012.

Weakliem, D. L.: A Critique of the Bayesian Information Criterion for Model Selection, Sociological Methods & Research, 27, 359–397, doi:10.1177/0049124199027003002, 1999.

Table 1. Theoretical 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	$e^{-((-\ln(u)^{\theta})+(-\ln(v)^{\theta}))^{\frac{1}{\theta}}}$	$(-\ln(t)^{\theta})$	$[1,+\infty]$
Gaussian	$\Phi_R(\Phi^{-1}(u),\Phi^{-1}(v)); R=\begin{bmatrix}1 & \theta\\ \theta & 1\end{bmatrix}$		(-1, 1)

Table 2. 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 %

Table 3. 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 %

Table 4. 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 %

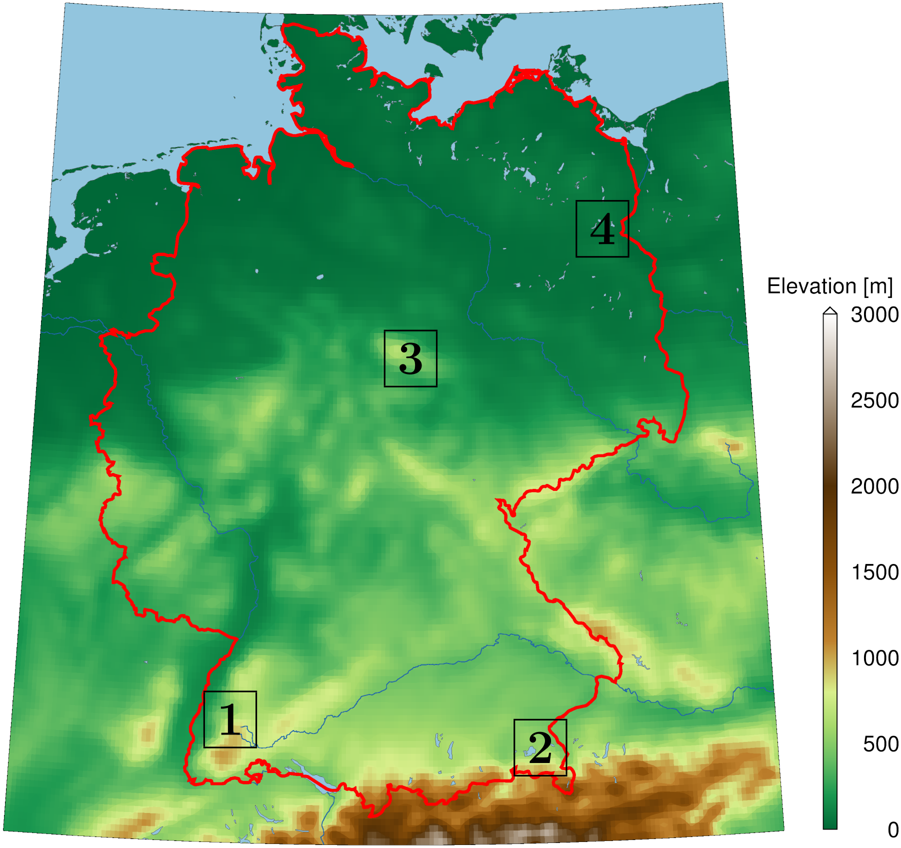

Figure 1. Terrain elevation of Germany (DEM). The numbers represent the position of the four specific grid cells for which the performance of the Copula-based algorithm is analyzed in Fig. 13.

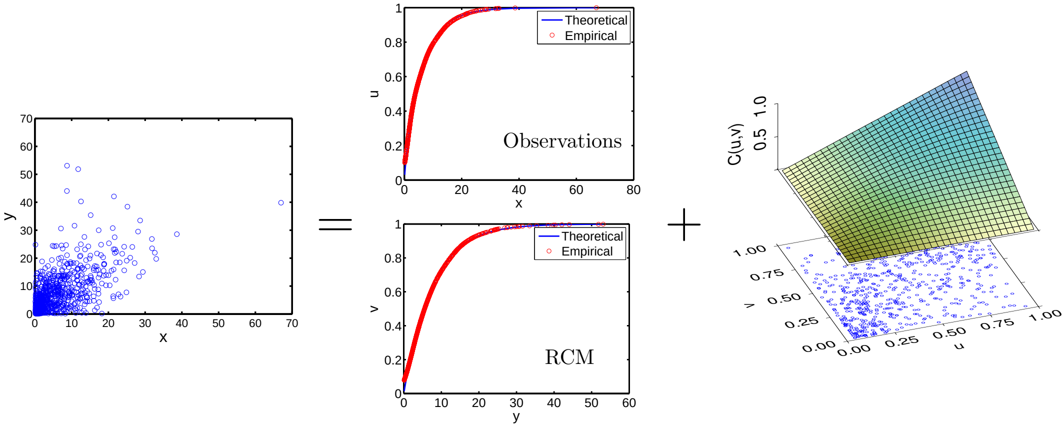

Figure 2. Visualisation of a bivariate Copula model consisting of two marginal distributions and a theoretical Copula function that describes the pure dependence.

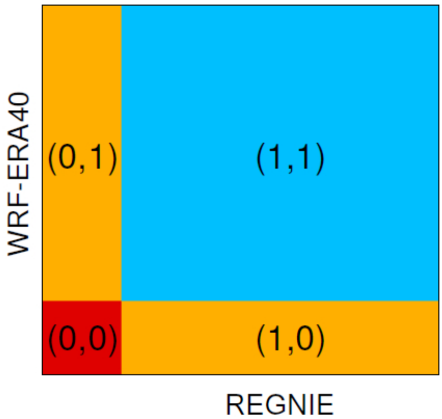

Figure 3. Illustration of the four cases: (0,0) indicates that both REGNIE and WRF show no rain, (0,1) stands for an observation with no precipitation but the RCM model shows a rain event, while (1,0) indicates the opposite of (0,1), (1,1) implies that both are wet.

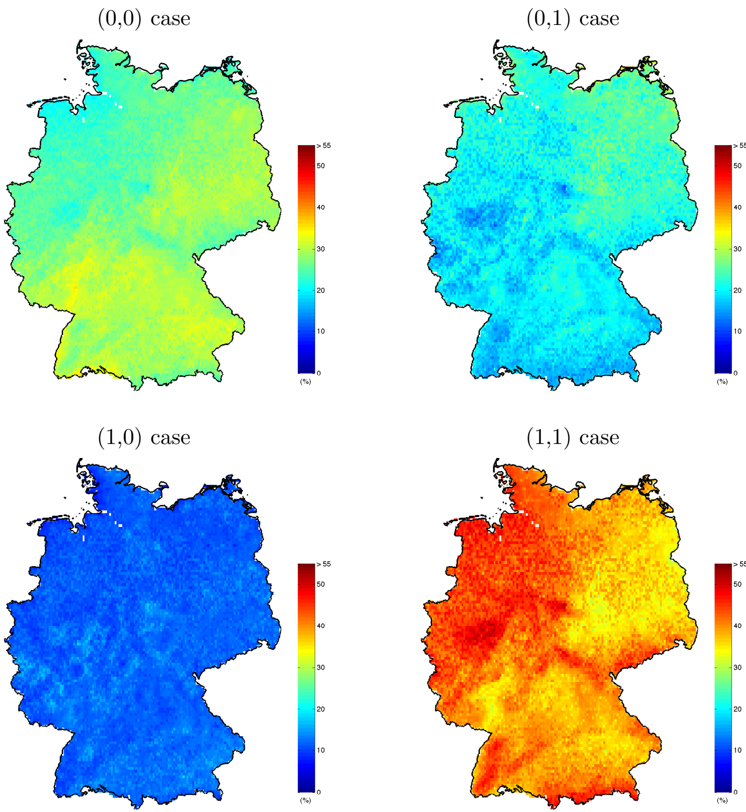

Figure 4. The proportion of the four cases over the study area for the validation time period (from 1986 to 2000).

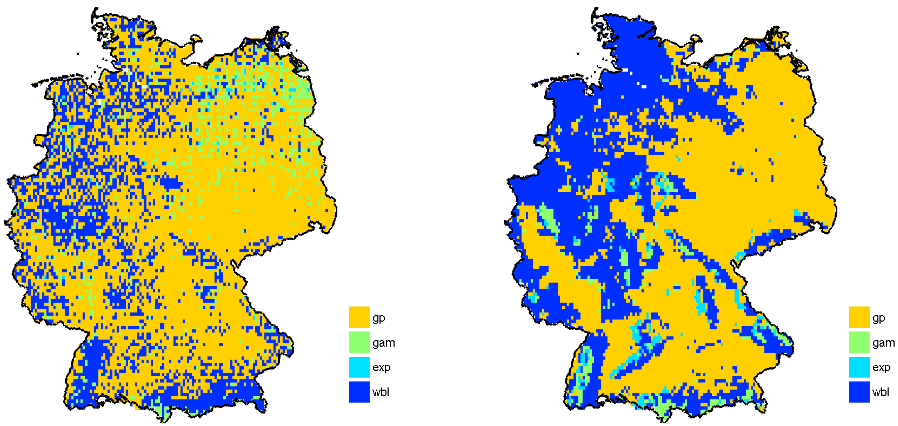

Figure 5. Estimated marginal distributions of precipitation for Germany for both REGNIE (left) and WRF (right). The results are shown for the calibration period (1971–1985) and positive pairs only.

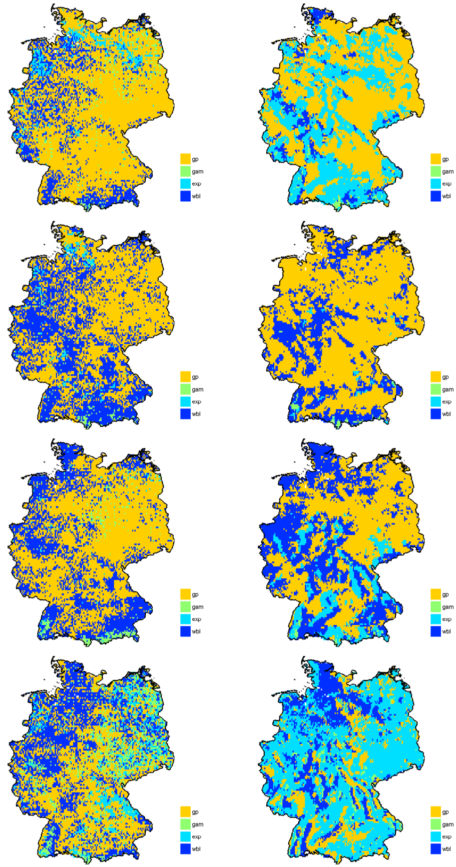

Figure 6. Estimated marginal distribution of precipitation for the different seasons for REGNIE (left column) and WRF (right column) in Germany. The results are shown for the calibration period (1971–1985) for positive pairs only. Spring (MAM), summer (JJA), autumn (SON) and winter (DJF) are illustrated from top to bottom.

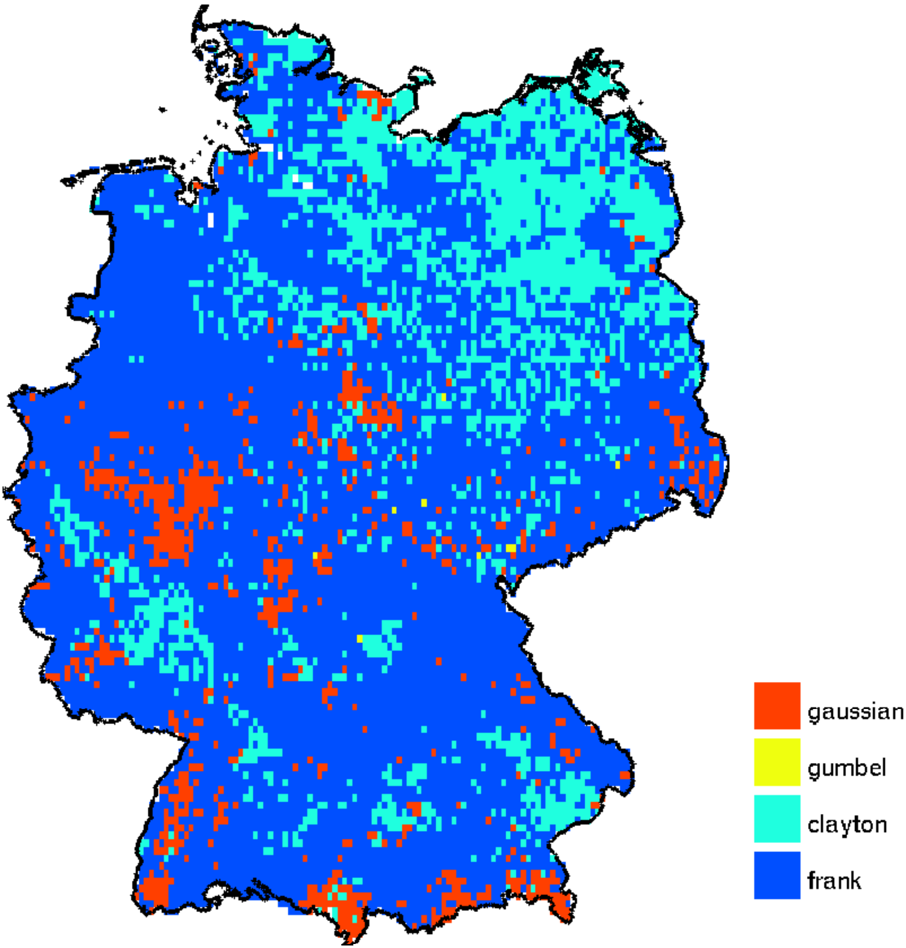

Figure 7. Identified Copula functions between REGNIE and WRF precipitation in the calibration period (1971 to 1985) with positive pairs.

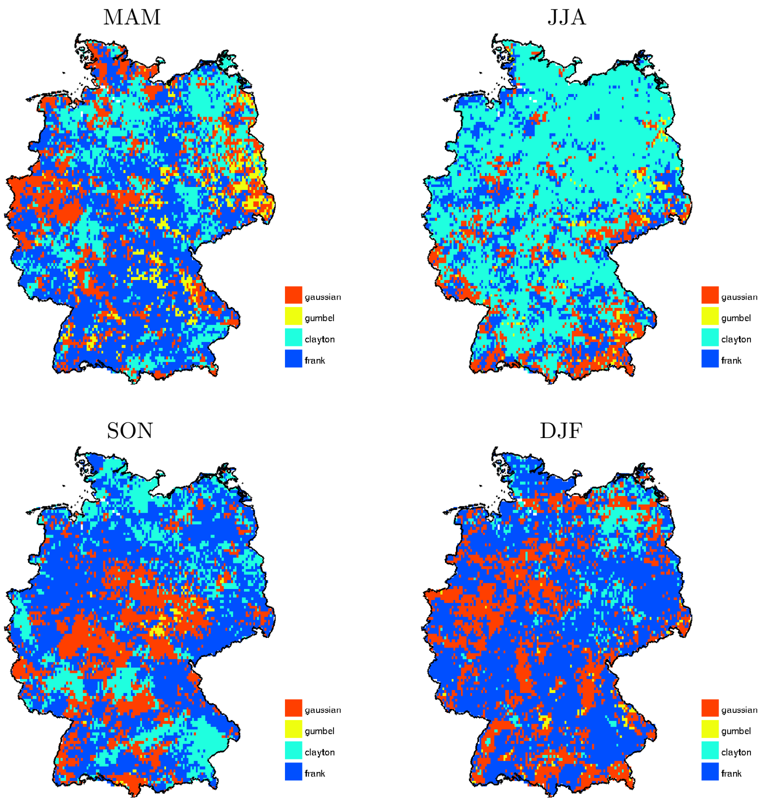

Figure 8. Fitted Copula functions between REGNIE and WRF precipitation (calibration period (1971–1985), positive pairs only). The Copulas are identified for the different seasons (spring – MAM, summer – JJA, autumn – SON, winter – DJF).

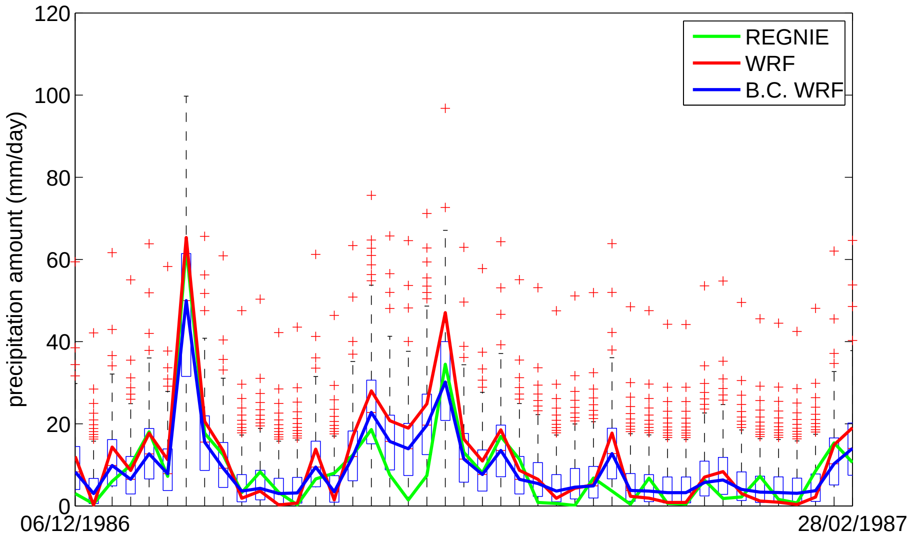

Figure 9. Comparison of bias-corrected WRF data (blue) with the original WRF data (red) and REGNIE (green) in winter 1986–1987 (positive pairs only) for pixel 1 in Fig. 1. For each time step 100 realisations are drawn from the conditional distribution visualized by the box-whiskers (boxes are defined 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.

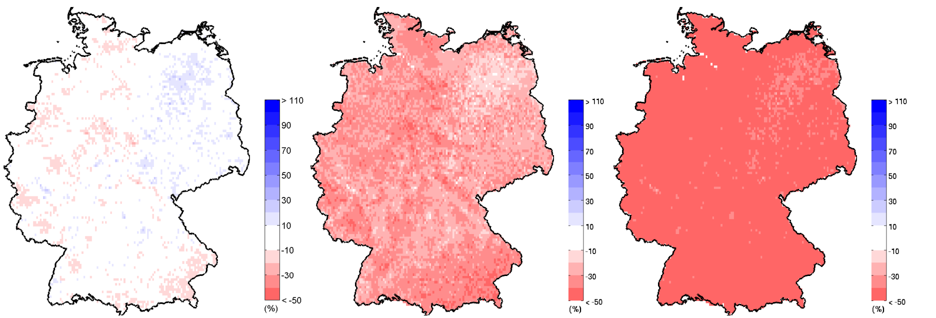

Figure 10. Relative bias map of mean daily precipitation for Copula-based correction by taking the expectation (left), median (middle) and the mode (right) as the estimator of the sampled distribution. The results are based on the validation period 1986-2000.

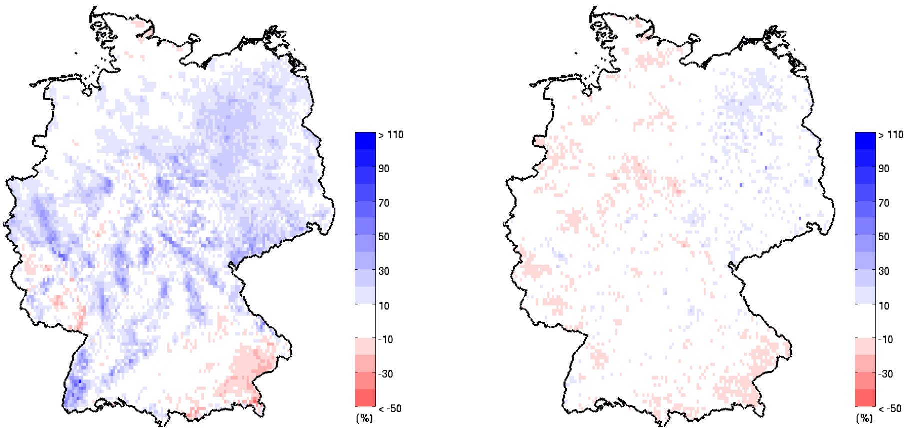

Figure 11. Relative bias of mean daily precipitation for uncorrected (left) and corrected WRF precipitation field (right). The results are based on the validation period 1986-2000.

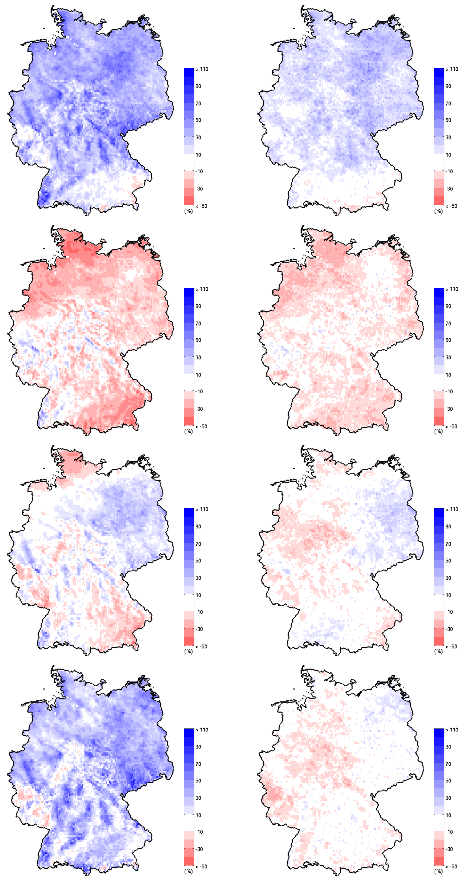

Figure 12. Relative bias between uncorrected (left) and corrected (right) WRF mean daily precipitation and the REGNIE data set in Germany for the different seasons (spring–MAM, summer–JJA, autumn–SON, winter–DJF, from top to bottom). The results are derived for the validation time period (1986-2000).

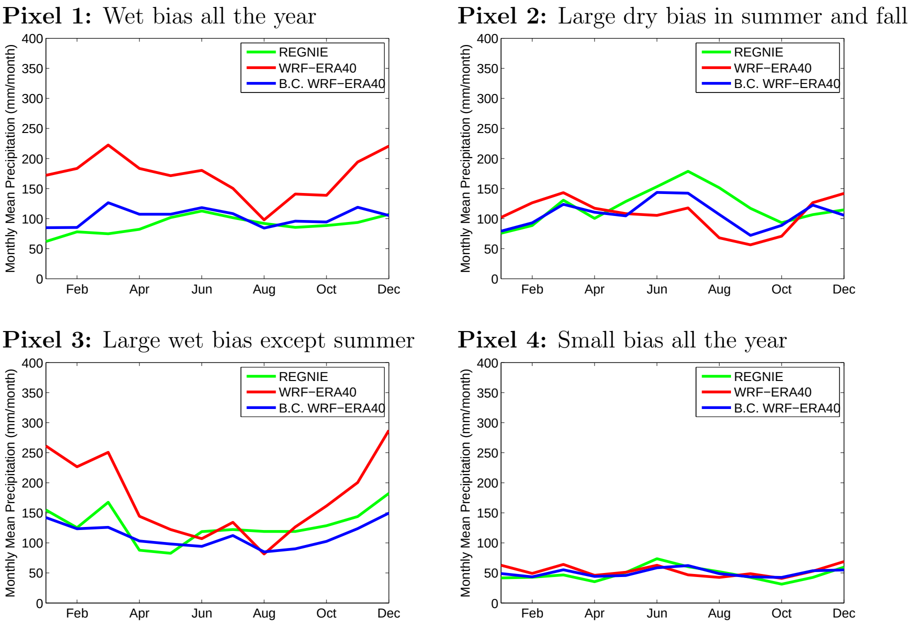

Figure 13. Comparison of bias corrected WRF mean monthly precipitation (blue) with REGNIE (green) and original WRF data (red) for the selected four pixel 1–4 in the validation period from 1986 to 2000. The number of the respective grid cell is noted in the upper left corner of each plot.

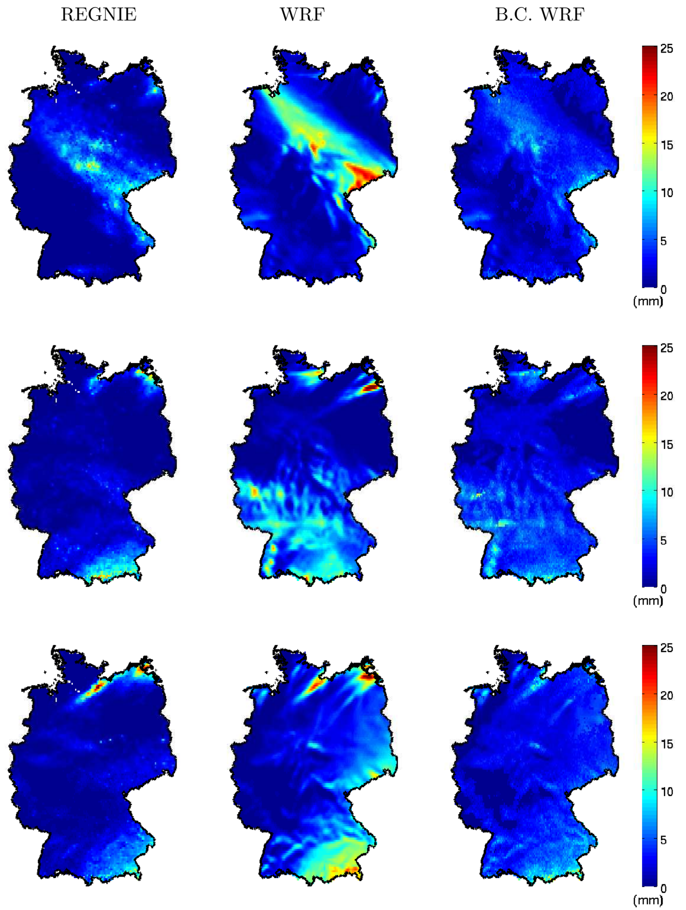

Figure 14. Daily precipitation fields over Germany for the three consecutive days from January 9 to January 11, 1986.

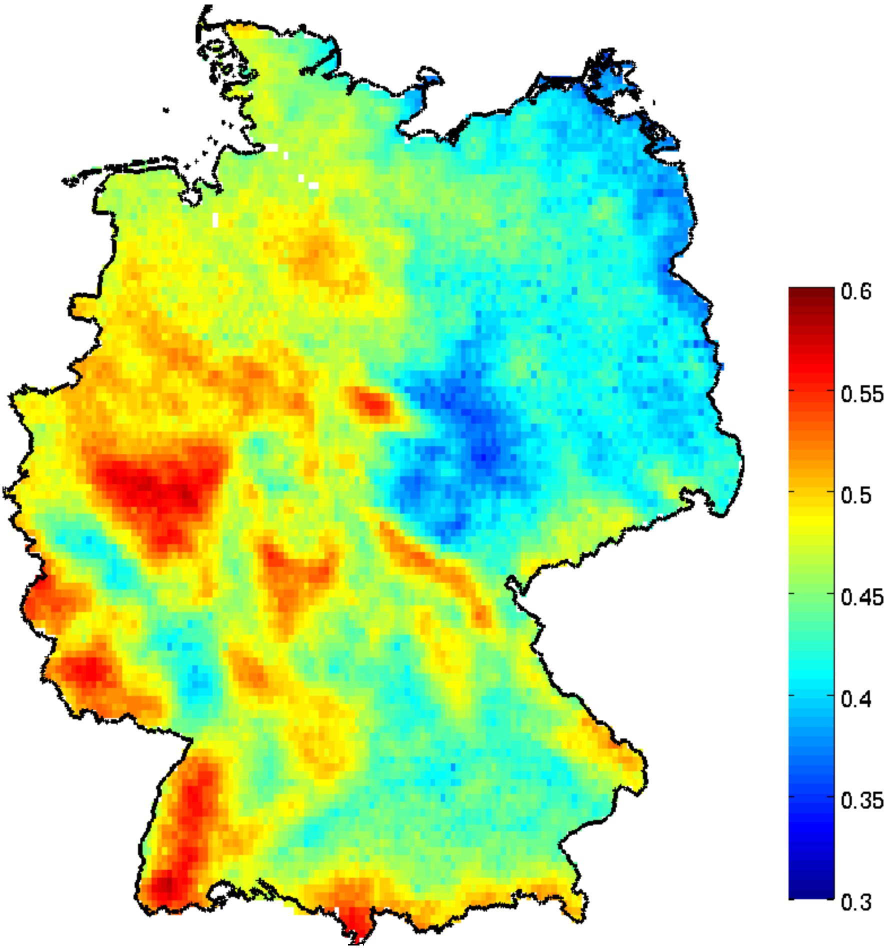

Figure 15. The rank correlations between RCM and REGNIE precipitation over the domain in the validation period from 1986 to 2000.

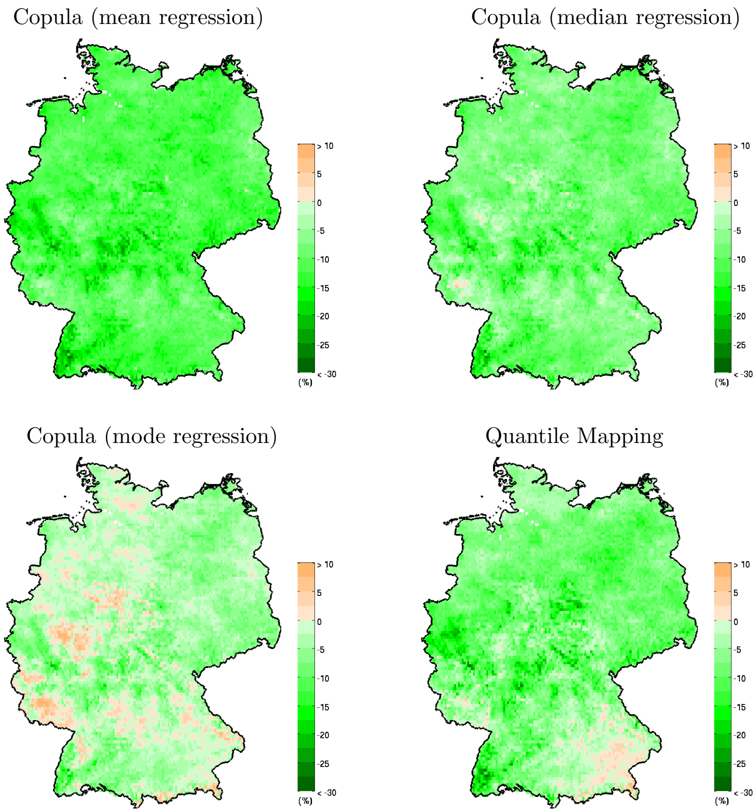

Figure 16. The changes of the RMSE in the validation period (1986-2000) by different bias correction methods. The green color indicates a decrease of the RMSE, while the ocher color implies an increase of the RMSE.

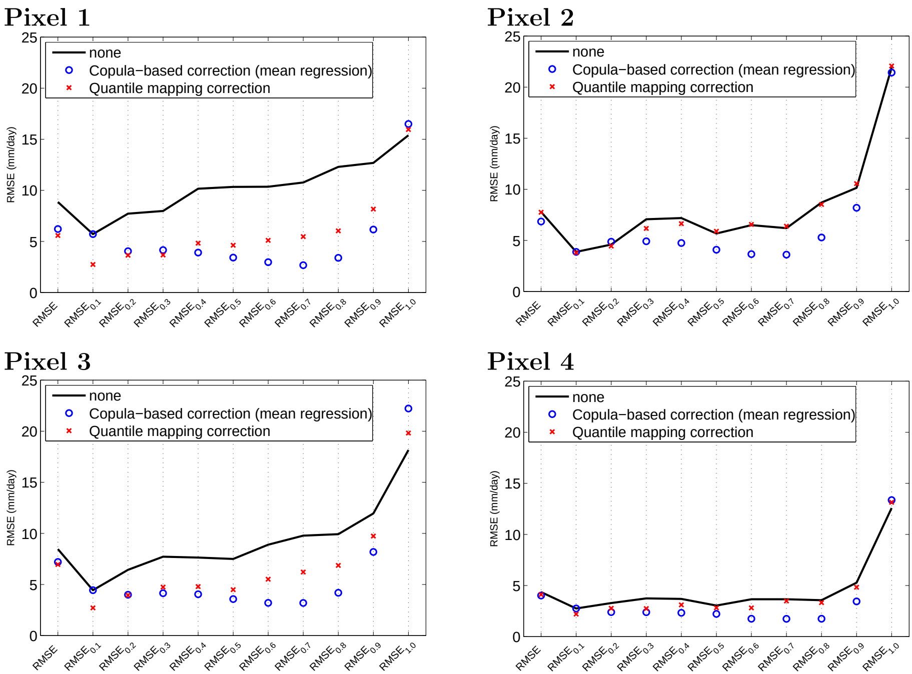

Figure 17. The root mean square errors (RMSE) and the root mean square errors for specific probability intervals ($RMSE_{0.1}$, $RMSE_{0.2}$, ..., $RMSE_{1.0}$) for different methods. The selected four pixels are the same as in Fig. 13. The black solid line indicates the errors without correction. The results are derived from the validation period from 1986 to 2000.

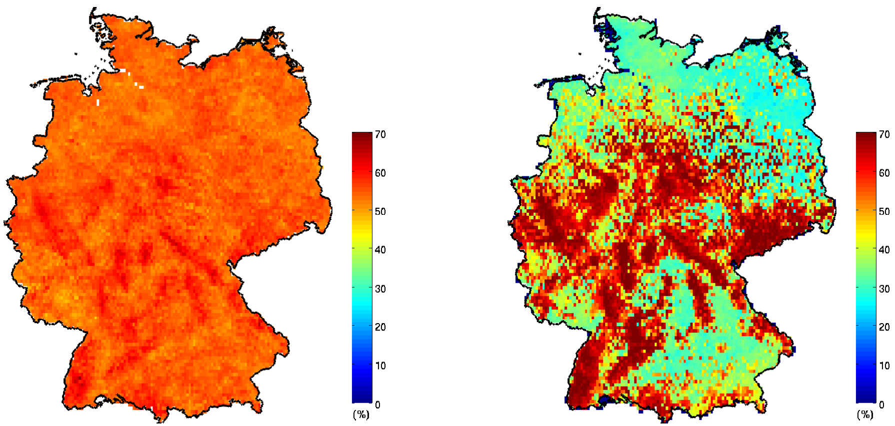

Figure 18. The percentage of the corrections that are closer to the observations. Left: Copula-based correction (mean regression); Right: quantile mapping correction. The results are derived from the validation period from 1986 to 2000.