To bias correct or not to bias correct? An agricultural impact modelers' perspective on regional climate model data

Patrick Laux[1,2], Reimund P. Rötter[3], Heidi Webber[4], Diarra Dieng[1], Jaber Rahimi[1], Jianhui Wei[1], Babacar Faye[6], Amit K. Srivastava[5], Jan Bliefernicht[2], Oluwafemi Adeyeri[7], Joel Arnault[1], and Harald Kunstmann[1,2]

[1]Institute of Meteorology and Climate Research (IMK-IFU), Campus Alpin, Karlsruhe Institute of Technology (KIT), Garmisch-Partenkirchen, Germany
[2]Institute of Geography, University of Augsburg, Augsburg, Germany
[3]TROPAGS, Department of Crop Sciences, University of Göttingen, Göttingen, Germany
[4]Leibniz-Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany
[5]Institute of Crop Science and Resource Conservation, University of Bonn, Bonn, Germany
[6]Institut de recherche pour le développement (IRD) ESPACE-DEV, Montpellier Cedex, France
[7]Department of Meteorology and Climate Science, Federal University of Technology, Akure, Nigeria

Correspondence: Patrick Laux (patrick.laux@kit.edu)

Abstract.

Many open questions and unresolved issues surround the topic of bias correction (BC) in climate change impact studies (CCIS). One question relates to the contribution of downscaling of climate change scenarios on the uncertainties in results obtained using impact models for agriculture. In particular, for large area or regional agricultural impact assessments, the question of bias correction is of high relevance. Relatively few studies exist looking at the quantification of the impacts of BC methods in general circulation model (GCM) and regional climate model (RCM) data on results of such impact studies. Here, we quantify the impact of different BC methods on temperature (T) and precipitation (P) from different CORDEX GCM-RCM combinations, and how the debiased T&P signal may propagate through agricultural impact models. Specifically, we estimate the impact of BC on i) an empirical P- and fuzzy rule-based algorithm to estimate the crop planting date, and ii) a mechanistic T&P-based approach to quantify the crop suitability index (CSI) for the main staple crops in West Africa (i.e. groundnut, maize, pearl millet, sorghum). Both approaches serve as a proxy for more complex process-based ecophysiological crop models. Depending on the choice of the BC method, the uncertainties in the assessment of the CSI can be more than twice as high compared to the uncertainties from the GCM-RCM model selection. Comparing the estimated CSI values with observed harvest area, it is found that BC generally improves the performance for models with low hit rates (< 30–35%), but decreases the performance for models with relatively high hit rates (> 35%). Such consequences can also be expected for process-based crop models, which are developed to operate on field-scale but are driven by coarser scale RCMs. It is concluded that such agriculturally oriented climate impact models as investigated here should be interpreted with great caution if applied without BC or relying on a single BC approach only. Rather, we suggest to include different BC approaches in the generation of climate projections for CCIS and quantify their uncertainties following a super-ensemble probabilistic assessment.

1 Introduction

Climate change will bring warmer temperatures, increased water demand, as well as more frequent extreme weather events which are expected to pose large challenges for agricultural production (IPCC, 2018). To assess the expected impacts of climate change, impact modelers depend on simulation results from climate models as input in their sector- or impact models (Rötter et al., 2011a). Together with elevated atmospheric CO_2 concentrations and changes in radiation, it is primarily the changes in temperature and precipitation (T&P) that are the most crucial for agricultural yields (Pirttioja et al., 2015). T&P are accordingly indicative for large-scale changes in regional energy- and water cycles in a region (e.g., Duan and Duan, 2020), have large impacts on crop growing periods and are major factors affecting heat and drought stress (Webber et al., 2018).

Assessing these changes on agriculture is largely the domain of climate change impact studies (CCIS) by utilizing agricultural models. In the following, we refer largely to mechanistic and process-based models, drawing examples specifically from crop models. Crop models dynamically simulate crop growth and development (for multiple crops and locations) in combination with many possible combinations of weather & climate, soil, and crop management conditions (Rötter et al., 2018; Faye et al., 2018), overcoming the limitations of agricultural experiments under field and controlled laboratory environments (Webber et al., 2018; Rötter et al., 2018). The accuracy of climate crop-yield relationships derived from crop models clearly depend on the availability and quality of the underlying yield, soil, crop management and climate input data (Parkes et al., 2019). The resulting uncertainties in projected impacts challenge their use by policy makers.

Uncertainties in crop impact modeling stem from three main sources: i) climate change model simulations and scenarios, ii) downscaling and bias correction methods, and iii) crop model simulations (Rötter et al., 2012). The uncertainties related to crop models can be further attributed to uncertainties related to model structure (e.g., Asseng et al., 2013), model parameters, and driving variables (e.g., Bindi et al., 2015; Tao et al., 2018).

Often, canopy- or field-scale models are applied in regional-scale climate change assessment studies (Van Bussel et al., 2011), leading to uncertainties either due oversimplification of the physiological processes represented and/or due errors resulting from spatial aggregation of input data (climate, soil, crop varieties, and management) (Webber et al., 2020). Baron et al. (2005), for instance, demonstrate the impact of using coarse-scale climate data in crop modelling using the field-scale model SARRA-H in West Africa. Their results indicate that forcing the crop model with spatially aggregated rainfall may cause yield overestimations up to 50% in dry latitudes of West Africa. More recent studies confirmed the need for high-resolution input data in order to simulate the spatial patterns of crop yield realistically (Zhao et al., 2015; Peng et al., 2020). In general, using a higher resolution in climate input data may reduce the simulated biases in crop yields, but may depend also on other settings applied in general circulation models (GCMs) and regional climate models (RCMs) such as e.g. the physical parameterization (e.g., Laux et al., 2019), the choice of the land surface model (e.g., Ramarohetra et al., 2015), or the land use data (e.g., Kerandi et al., 2017). Possible over- or underestimations might be corrected using statistical post-processing, i.e. bias correction (BC) approaches. However, the application of BC remains a controversially discussed issue (Ehret et al., 2012).

Despite its importance, the impact of different BC methods has rarely been considered in agricultural impact studies. The variety of BC approaches applied in hydrology seems to be much larger, ranging from simple linear methods (e.g., Hay et al., 2000; Lenderink et al., 2007) via nonlinear power transformation (Shabalova et al., 2003; Leander and Buishand, 2007) and statistical distribution- based algorithms (e.g., Piani et al., 2010) to more complex transfer functions using copulas (Laux et al., 2011; Mao et al., 2015; Maity et al., 2019). Albeit not yet well-established in impact modeling, multivariate approaches have been recently developed which may better retain the inter-variable relationships (i.e. the physical consistency between the variables as obtained by the climate models) after BC, as compared to more traditional approaches (Cannon et al., 2018; Vrac, 2018).

One common approach in agricultural impact assessment is the application of the 'delta change' method (e.g., Räisänen and Räty, 2013): Climate scenarios are generated by imposing the mean monthly future changes of the variables (e.g. T&P) upon historical observations (Wilby et al., 2004; White et al., 2011). This approach retains the variability of the local-scale weather observations while adding the expected climate signal. In addition to the 'delta change', weather generators and quantile-based distributional shifts are applied to create scenarios that alter interannual and intraseasonal climate variability (e.g., Rötter et al., 2011b; Confalonieri, 2012). Another BC approach is followed in the *Inter-Sectoral Impact Model Intercomparison Project* (ISIMIP) (e.g., Frieler et al., 2014, 2017), which is purported to allow robust bias adjustment of extreme values and is able to preserve trends across all quantiles compared to other BC methods (Lange, 2019). Oettli et al. (2011) have used the *cumulative distribution function transform* (CDF-t) (Michelangeli et al., 2009) to correct the biases of eight regional models. They showed that bias-corrected climate variables better simulate crop yield means and standard deviations, but offer no improvement in the reproduction of the year-to-year variations. More recently, the CDF-t has been applied to debias CMIP5 GCMs daily data over Africa (Famien et al., 2018).

The quantification of these uncertainties is important for the robustness of projected impacts (e.g., Tao et al., 2018). While the use of climate- and impact model ensembles to quantify uncertainty are increasingly common in international efforts such as the Agricultural Model Inter-comparison and Improvement Project (AgMIP) (Rosenzweig et al., 2013), the Modeling European Agriculture with Climate Change for Food Security (MASCUR) project (Ewert et al., 2015), or the Inter-Sectoral ISIMIP. The recent initiative Bias ADJustment of climate scenarios for Agricultural Model applications (BADJAM) compares the driving data obtained from different bias adjustments techniques and their impact on crop models (Galmarini et al., 2019). Only few of these initiatives have explored the full range of uncertainties.

This study contributes to the assessment of uncertainties related to the choice and processing of climate information, as well as their propagation in climate impact assessments. Large uncertainties exist based on the selection and the spatiotemporal resolution of the climate data (Ruane and McDermid, 2017). Simulations are available through different emission scenarios or *Representative Concentration Pathways* (RCPs) and various GCMs. Large climate model ensembles are generated by joint research initiatives such as the *Coupled Model Intercomparison Project* (CMIP), or other CMIP-endorsed MIPs such as *Coordinated Regional Climate Downscaling Experiment* (CORDEX), providing downscaled high-resolution RCM projections based on different GCMs (subsequently referred to as GCM-RCM combinations). Though becoming less common, for some crop modeling groups and climate change impact modelers, resource limitations such as computing power may limit options for full consideration of large ensembles of GCM-RCM combinations (Ruane and McDermid, 2017), requiring decisions based on reduced subsets of possible future climate projections. Yet, considerable knowledge gaps remain concerning in the impact of BC approaches on the generation of climate change scenarios for crop modeling impact studies. Given the range of uncertainties and issues, with potentially large implications for uncertainties in impact projections, we address the following research questions in this study:

- is bias correction still required with state-of-the-art high-resolution (region) climate projections? And,
- what is the range of uncertainties related to the application of different bias correction approaches in crop-climate impact research?

By quantifying how different BC approaches impact on the selection through subsetting of climate projections and simulated results, we also provide practical guidance to crop modellers on how to generate climate change scenarios to drive crop-climate modeling impact assessments. We address these research questions by quantifying the uncertainties of four different BC methods applied to a set of selected CORDEX GCM-RCM combinations, propagated through two simple crop-climate models: i) an empirical P- and fuzzy rule-based algorithm to estimate the agriculturally relevant onset of the rainy season (ORS), i.e. the crop planting date, and ii) a mechanistic T&P-based approach to quantify the crop suitability expressed as a seasonal crop suitability index (CSI) for the period of June–September (JJAS).

2 Data and Methods

2.1 CORDEX climate data

CORDEX provides a set of climate projections over Africa (Nikulin et al., 2012). To enable a large ensemble of climate projections, CORDEX-Africa simulations were carried out on a rather coarse resolution of 0.44° (approx. 50 km). Many variables are provided by CORDEX, but in this study, only T&P is considered. In climate models, T is a direct response to radiation and involves relatively simple scattering, absorption, and emission equations. In turn, P is the result of many non-linear interactions such as evaporation, convection, cloud formation. Some of them occur on scales smaller than the model grid resolution, and hence, requiring parameterization with different inherent assumptions and approximations. Furthermore, P is a more non-homogeneous variable in its spatio-temporal distribution in nature compared to T (Schleiss et al., 2011). For both reasons, P is more difficult to model accurately, i.e. uncertainties in modeling precipitation are relatively higher than for temperature.

It is found that many of the CORDEX GCM-RCM combinations significantly improve the representation of precipitation compared to that from their boundary conditions (i.e. the driving GCMs), however, biases still may exist, particularly in the diurnal cycle. Only a small subset of models have a reasonable representation of the phase of the diurnal cycle, while for the majority of the RCMs precipitation is triggered too early during the diurnal cycle (Nikulin et al., 2012). It is noted that combinations based on the non-Gregorian calendar are excluded from this study. Finally, a set of nine CORDEX GCM-RCM combinations is applied in this study (Table 2).

2.2 AgMERRA climate data

AgMERRA data, which has been used in numerous modeling activities for regions where observed long series of historical weather records are scarce or not accessible, is applied in this study as reference data for bias correction and spatial disaggregation for the CORDEX data. AgMERRA is a hybrid of the *Modern-Era Retrospective-analysis for Research and Applications* (MERRA, (Rienecker et al., 2011)) and various gridded and satellite products. It provides a global, daily, $0.25° \times 0.25°$ gridded climate data from 1980–2010 containing maximum, minimum, and average temperatures, precipitation, solar radiation, wind speed, and relative humidity at the maximum temperature time of day (Ruane et al., 2015).

2.3 SPAM agricultural data

SPAM (Spatial Production Allocation Model) data is used in this study as a reference to compare the spatial patterns of crop suitability. Data on global agricultural production are usually available as statistics at administrative units, which makes them less suitable for validation of spatially distributed CCIS output due to their aggregation level. SPAM, therefore, uses a cross-entropy approach to make plausible estimates of crop distribution within disaggregated units (IFPRI, 2019). The data is provided with a 10 km grid-cell resolution and contains four variables which are calculated by the model: physical area, harvest area, production, and yield, altogether for 42 crops, split by the rainfed and irrigated production system, as well as the combination of both. In this study, we apply SPAM 2017 v1.1 Sub-Saharan Africa rainfed data on the harvested area, for groundnut, maize, pearl millet, and sorghum. The data is a snapshot for the year 2007 only.

In this study, the 2007 SPAM data is compared to suitability maps for the period 2006–2100 (assuming steady climate). Moreover, crop suitability considers only climatic variables but neglects edaphic, as well as, societal constraints.

2.4 Subsetting of possible climate scenarios

For some crop modeling groups and climate change impact modelers resource limitations such as computing power may limit options for full consideration of very large climate ensembles (Ruane and McDermid, 2017), thus requiring decisions based on reduced subsets of possible future climate projections. Subsetting is one possibility to overcome this limitation, although this leads to a loss of information and begs the question of what type of information should be maintained (Ruane and McDermid, 2017; Pirttioja et al., 2019).

In this study, we apply the *Representative Temperature and Precipitation (T&P) GCM Subsetting Approach.* The approach helps to group the single members of the climate scenarios into five basic classes of climate changes (i.e. relatively cool/wet, cool/dry, hot/wet, hot/dry, and normal). These classes of projected climate change are calculated as location-specific T&P changes in terms of their deviation from the ensemble median, i.e. their anomalies. The 'normal' class is calculated to include all models whose T&P changes are within $\pm 0.5 \times \sigma$ (standard deviation) of the median (Ruane and McDermid, 2017). Here, it is applied to demonstrate the potential impact of BC on the simulated T&P regimes in West African based on the CORDEX GCM-RCM combinations. We adapted the procedure in that we consider the statistics of the entire year, and not the growing season of specific crops exclusively as we do not want to constrain our analyses to a specific crop at this point. Rather we aim to contrast the different climate scenarios more generally. In case impact modelers are interested in specific crops, we suggest to consider crop-specific planting windows as well as the growing periods of the crops and their varieties, respectively. For future climate projections, it should be noted that planting windows may change remarkably. As the daily maximum temperature (T_{max}) plays a crucial role for crop limiting conditions in agricultural impact models, we consider T_{max} instead of daily mean temperature (T_{mean}). The period considered in the CTRL run is 1980–2005, and for the RCP4.5 scenario is 2006–2100.

2.5 Spatial disaggregation and bias correction of GCM-RCM combinations

Spatial disaggregation of the CORDEX GCM-RCM combinations is performed first before applying different bias correction approaches. The spatial disaggregation consists of the distance-weighted average (DWA) remapping of the CORDEX climate data to the grid resolution of AgMERRA, i.e. from 0.44° to 0.25°. The underlying assumption is that the similarity of any unknown value decreases with the distance to a known value. The weights to estimate unknown values are assigned that they increase to the inverse of the distance to each known point. More technical information on the spatial disaggregation is given in Appendix A.

BC methods, in general, derive transfer functions between model simulations and observations for a baseline period and apply these functions for future climate projections, in which no reference data is available. BC is crucial because CORDEX GCM-RCM combinations may deviate from the observed climatic data. Such biases may propagate through agricultural impact models.

Bias correction methods of varying complexity have been developed and applied to the output of different RCMs driven by different large-scale boundary conditions, i.e. different CORDEX GCM-RCM combinations, based on different emission scenarios. This study is limited to the application of BC methods for T&P because these two variables are the main atmospheric drivers for most climate impact models. To enable performance evaluation, the transfer functions are typically derived for a calibration period and are evaluated for an independent period (cross-validation). More technical information on the setup and the different BC approaches is given in Appendix B.

Table 1. List of applied bias correction (BC) approaches to the spatially disaggregated CORDEX GCM-RCM combinations. More information about the BC methods is given in Appendix B.

Name of BC method	Abbreviation	Variable	Reference
Linear Scaling	LS	T, P	Hay et al. (2000)
Power Law Transformation	PLT	T, P	Leander and Buishand (2007)
Theoretical Distribution Mapping	DM_{theo}	T, P	Themeßl et al. (2011); Fang et al. (2015)
Empirical Distribution Mapping	DM_{emp}	T, P	Themeßl et al. (2011); Fang et al. (2015)

Four different BC methods for T&P, borrowed from hydrological impact assessment studies, are used to estimate their potential consequences in agricultural impact assessment. Two established approaches with different complexity, both based on threshold-values of meteorological parameters, are applied for demonstration: i) For the estimation of the onset of the rainy season (section 2.6), P has been debiased. LS, PLT, DM_{theo}, and DM_{emp} are applied (Table 1); ii) for the estimation of the impact of the BC on crop suitability (section 2.7), both T&P have been debiased, leading to 16 (i.e. 4×4) permutations of the 4-fold set of {T,P} BC approaches.

2.6 Fuzzy-rule based estimation of the onset of the rainy season (ORS)

Particularly in semi-arid regions with limited water availability for cropping, the determination of the planting date is of crucial importance for food production (Laux et al., 2008). For West Africa, there are many definitions in use, most of them are threshold-based and apply P observation data, (e.g., Dodd and Jolliffe, 2001). Such threshold-based approaches are implemented in a few crop models to automatically estimate planting dates, given as agricultural management input data to the crop models.

Due to the sternness of threshold-based ORS approaches, a new fuzzy rule-based algorithm has been developed (Laux et al., 2008). It consists of three different criteria, i) accounting for sufficient soil moisture to initiate germination, ii) excluding single wet days which may often occur during the pre-monsoonal period and could be misinterpreted as the onset of the monsoonal rains, and iii) avoiding a dry period of seven or more consecutive days during the following month to ensure the survival of the seedlings. The approach has been applied across sub-Saharan Africa, e.g. for seasonal forecasting (Siegmund et al., 2015; Rauch et al., 2019) and in climate projections (Dieng et al., 2018), or coupled with process-based crop models in order to derive location-specific planting rules (Laux et al., 2010; Waongo et al., 2013). Due to the importance of the planting date for agricultural CCIS in West Africa under rainfed conditions (Faye et al., 2018), we applied the fuzzy rule-based planting approach of Laux et al. (2008) using the same parameters as used for Ghana and Burkina Faso in this study. It should be noted, that the planting rules (and consequently the planting dates) depend not only on the region but also on crop-specific needs. In this study, however, we apply the same planting rules for the entire study region since we are not interested in obtaining suitable crop-specific planting dates, but rather in the relative differences due to the different BC approaches.

2.7 Crop Suitability

Crop growing conditions vary between different crops, varieties, and for different climatic, edaphic, hydrological, and topographic conditions. Due to the lack of detailed environmental data such as ground- and surface water for irrigation and detailed edaphic parameters, climate data is often the only source of information considered. Crop suitability further describes the appropriateness of a given land area based on the growing thresholds for crops in relation to climatic conditions (FAO, 1976; Egbebiyi et al., 2019).

We apply the Ecocrop model, based on climatic growing thresholds from the FAO-Ecocrop database, which has been developed from results of a large number of field experiments across the world (Hijmans et al., 2001). The FAO-Ecocrop database consists of thresholds of monthly suitability ranges for plant species against total monthly rainfall and monthly minimum, mean and maximum air temperatures over the length of its growing season (Egbebiyi et al., 2019). The suitability of a crop in response to climate variables can be assessed through the *Crop Suitability Index* (CSI).

Table 2. List of applied CORDEX GCM-RCM combinations applied in this study. The modeling institutes are: Centre National de Recherches Météorologiques, Centre Européen de Recherche et Formation Avancée en Calcul Scientifique (CNRM-CERFACS) for #1, Irish Centre for High-End Computing (ICHEC) for #2–4, Max Planck Institute for Meteorology (MPI) for #5–8, and Norwegian Climate Centre (NCC) for #9. For the latter, T_{mean} is not available and is thus not considered in the calculation of the CSI.

ID #	Modeling Institute Abbreviation	GCM Name	RCM Name (Version)
1	CNRM-CERFACS	CNRM-CM5	CCLM4-8-17 (v1)
2	ICHEC	ICHEC-EC-EARTH	CCLM4-8-17 (v1)
3	ICHEC	ICHEC-EC-EARTH	REMO2009 (v1)
4	ICHEC	ICHEC-EC-EARTH	HIRHAM5 (v2)
5	MPI	MPI-M-MPI-ESM-LR	CCLM4-8-17 (v1)
6	MPI	MPI-M-MPI-ESM-LR	REMO2009 (v1)
7	MPI	MPI-M-MPI-ESM-LR	RCA4 (v1)
8	MPI	MPI-M-MPI-ESM-LR	UQAM-CRCM5 (v1)
9	NCC	NCC-NorESM1-M	HIRHAM5 (v1)

The Ecocrop model is considered as a computationally inexpensive but valuable tool to help to understand the spatio-temporal distribution of crop suitability over large areas and long time periods. Since fully empirical and working with coarse temporal resolutions, it cannot be applied for detailed studies dealing with crop yield predictions under various management options like process-based crop models, which are resolving the full physiological life cycle of crops. Technical details about the calculation and interpretation of the CSI are given in Appendix C.

In this study, we quantify the impact of the different BC methods on the CSI across West Africa. As CSI depends on thresholds both for T&P, it adds more complexity compared to the algorithm to estimate the ORS (see section 2.6), which depends on P thresholds exclusively. To study the impact of the BC method, we calculated the CSI for all possible combinations, i.e. the full factorial set for the univariate BC methods for T&P, respectively, resulting in 4×4 combinations.

3 Results

First, the variability of the uncorrected CORDEX GCM-RCM climate ensemble (consisting of 9 members) is illustrated for P. The ensemble mean shows large spatial differences of P across the West African domain, ranging from < 1 mm/day in the Sahara up to 16 mm/day around the coastal areas of the Gulf of Guinea (Figure 1 a)). The dispersion of the different members, expressed as the ensemble standard deviation even exceeds the magnitude of the ensemble mean, reaching values up to 25 mm/day (Figure 1 b)). This indicates the high uncertainties in the generation of climate change projections for CCIS.

In the following, we present the impact of the bias correction methods on i) the representation of the variables T&P, separately, ii) the P threshold-based approach to estimate the onset of the rainy season, and iii) the joint T&P crop suitability index.

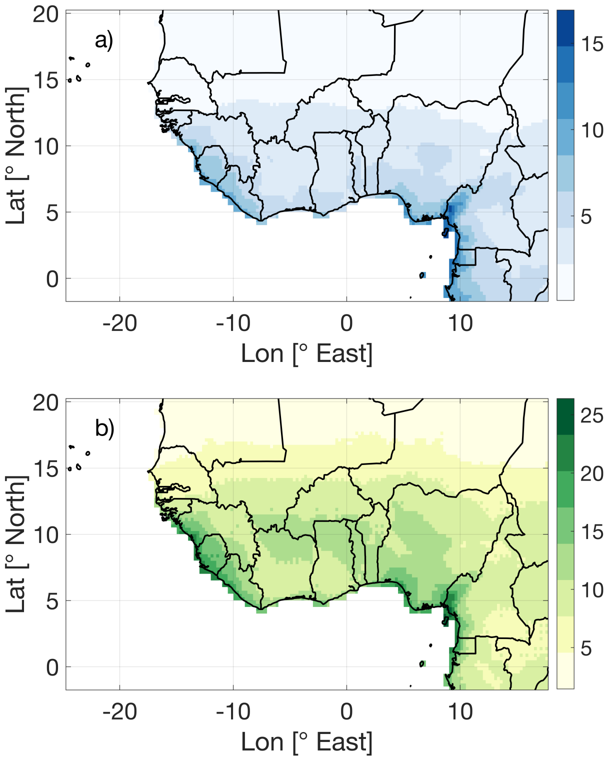


Figure 1. Uncorrected CORDEX GCM-RCM of a) ensemble mean of P [mm/day], b) ensemble standard deviation of P [mm/day] across West Africa for 2006–2100 under the RCP4.5 scenario. The ensemble consists of 9 members (Table 2).

3.1 Representation of T&P

The nine models adequately represent the different possible combined wetness and warming states, i.e. dry/hot, wet/hot, dry/cold, wet/cold, and normal/normal (Figure 2). Only for the CTRL period, the state “normal” is not occupied. We apply the approach for the RCP4.5 scenario. The reason for this choice is the higher number of CORDEX GCM-RCM combinations availability than for RCP26 and RCP85. Compared to the results for the CTRL, it is found that the representation of the states remains relatively stable if the RCP4.5 scenario is considered. Most of the GCM-RCM combinations remain in the same class, except for combinations #4 and #5, which shift from wet/cold and dry/cold into dry/cold and normal states, respectively. In addition, it is noteworthy that despite the same large-scale forcing GCM, different RCMs can lead to different states. For instance, the 4 MPI-ESM-LR models are represented in 3 states respectively, both under CTRL and RCP4.5. This indicates that the variability from the RCMs can be relatively higher than that from the GCMs, supporting the use of an RCM ensemble driven by fewer GCMs rather than a large GCM ensemble combined with one single RCM. However, it is worth mentioning that using multi-GCM multi-RCM combinations is always preferential (in case of sufficient resources for this kind of impact simulations).

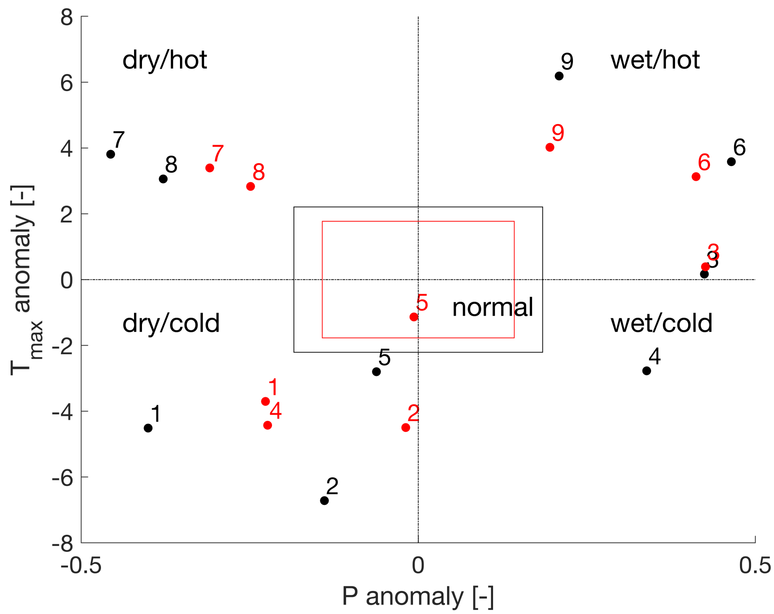


Figure 2. Anomaly plot of T&P, following the procedure of Ruane et al. (2015). Shown are the uncorrected CTRL run of 1980–2005 (black dots) vs. the uncorrected RCP4.5 runs (2006–2100) (red dots). The numbers denote the different CORDEX GCM-RCM combinations given in Table 2.

Figure 3 contrasts the impact of different bias correction methods on the mean values of daily T_{max} and P. For CORDEX GCM-RCM combination #1 it is seen that the variability (expressed as a range of the distribution, i.e. the boxes) of T (top) is lower for the uncorrected runs (both the control run (CTRL) as well as the RCP4.5 scenario run) compared to the observation (AgMERRA). As expected, the distribution of the RCP4.5 is shifted to higher values than for the CTRL. All of the applied bias correction approaches can increase the variability of the RCP4.5 scenarios (comparable to that of AgMERRA) and to lift the distribution in the order of the differences between CTRL and RCP4.5 run. This effect is less pronounced for the LS compared to the PLT and DM_{theo} and DM_{emp}.

For P (bottom), a slightly reduced variability of the distribution is found for the uncorrected CTRL and RCP4.5 runs compared to the observation. However, more values are declared as outliers in the upper tails, indicating more and higher P extremes in this CORDEX GCM-RCM run. The bias correction methods can inflate the distributions that are comparable to the observations. In addition, the BC is reducing the number of identified 'outliers', i.e. extreme values. It is however not the subject of this study to validate the results with independent observational data. Unlike T, there is no clear shift in the distribution from the CTRL to the RCP4.5 scenario, thus the median values should remain at similar levels after bias correction.

The LS bias correction is reducing the spread of the T_{max} anomaly compared to the uncorrected RCP4.5 run (Figure 4, top). The spread in the P anomaly (a-axis) remains similar compared to the uncorrected RCP4.5 run, is however generally on a small level (see scaling of the x-axis). The spread in the T_{max} anomaly is remarkably reduced and centered around zero. Altogether,

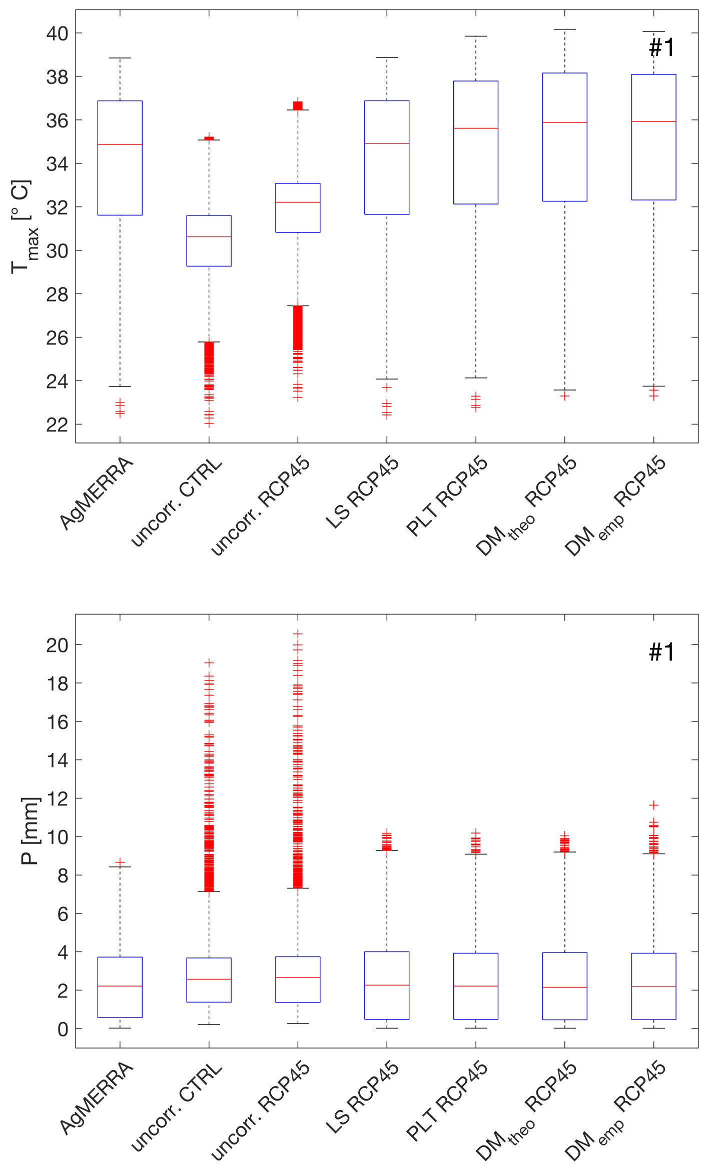
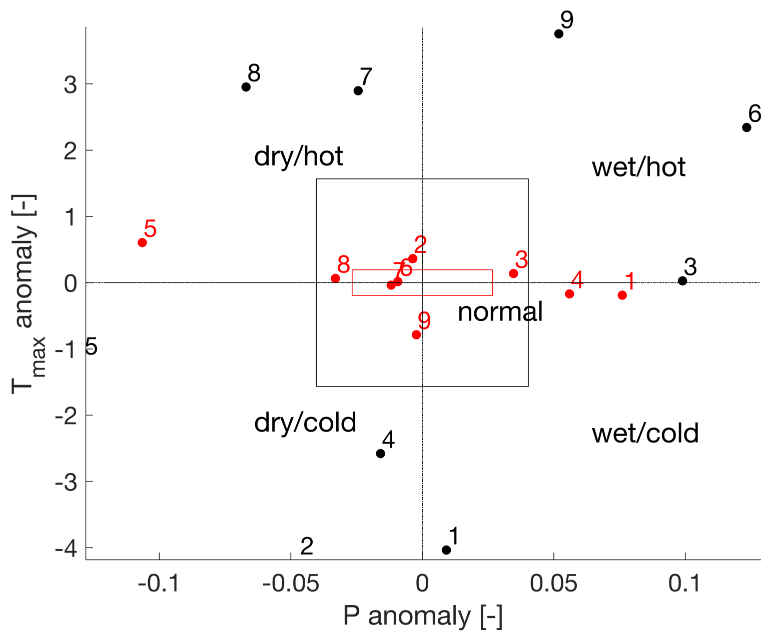


Figure 3. Boxplot of observation (AgMERRA), uncorrected CTRL (both 1980–2005), and RCP4.5 (2006–2100) run of CORDEX GCM-RCM combination #1 (see Table 2), and the bias-corrected RCP4.5 runs for T_{max} (top) and P (bottom). The boxplots represent the mean values for all grid cells in the study region. The box spans the interquartile range, the median is marked by the red vertical line inside the box, the whiskers represent the highest values, whereas the red crosses represent outliers.



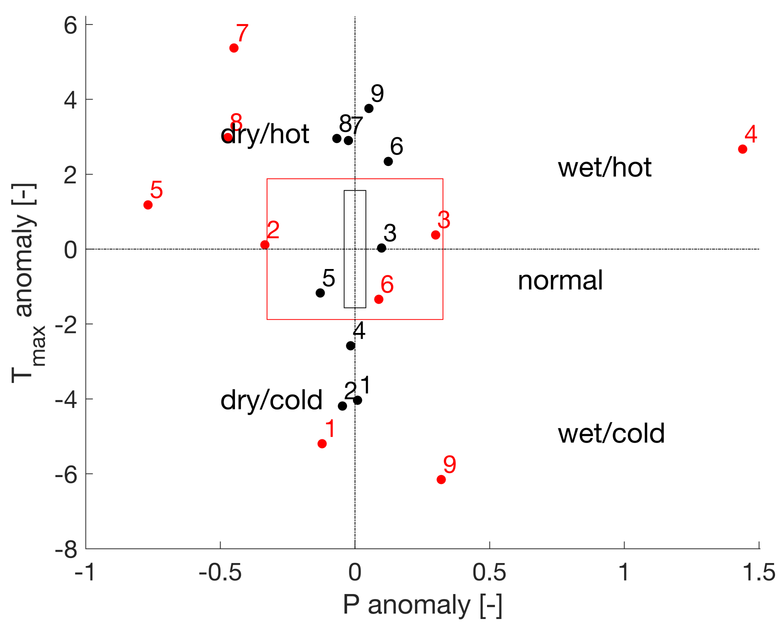


Figure 4. Same as Figure 2, but for uncorrected RCP4.5 runs (black dots) vs. bias-corrected RCP4.5 runs (red dots). The randomly selected BCs applied are the linear scaling (LS) (top) and the distribution mapping based on a fitted theoretical distribution (DM_{theo}) (bottom) for both T and P, respectively.

the bias correction leads to some changes of the wetness-temperature classification: e.g. #9 changes from wet/hot to dry/cold, #4 from dry/cold to wet/cold, and #6 from wet/hot to normal. The bias correction based on the DM_{theo} (Figure 4, bottom), in turn, is widening the spread of both the T and the P anomaly. This may also lead to regime changes (e.g. #4 changes from wet/hot to dry/cold).

3.2 Onset of the rainy season

Figure 5 demonstrates the impact of different BC methods on the calculated ORS dates in West Africa. Figure 5 a) shows the mean ORS dates (2006–2100) for the uncorrected RCP4.5 run of CORDEX GCM-RCM combination # 1 as an example. It can be seen that the ORS pattern, in general, follows the latitudinal movement of the monsoonal rains, triggered by the movement of the Intertropical Convergence Zone (ITCZ). The same model run has been bias-corrected by 4 different BC methods and the deviations between the corrected model runs and the uncorrected run (i.e., ORS (debiased RCP4.5) *minus* ORS (uncorr. RCP4.5)) are shown in Figure 5 b) to e). The spatial pattern of the deviation may vary tremendously over the West African domain, between the different CORDEX GCM-RCM combinations (not shown) but also between the different BC approaches.

Averaged for the entire domain, the deviations for #1 are ranging from 5 days for PLT (DOY 165) up to 21 days (DOY 183) in relation to the uncorrected RCP4.5 run (DOY 160).

Table 3 gives a summary of the corresponding spatial-temporal averaged statistics (i.e. the mean and standard deviations) for all CORDEX GCM-RCM combinations under consideration. In general, it is found that in comparison to the ORS dates of the observation AgMERRA (DOY 163), the mean ORS in the uncorrected RCP4.5 runs occur too early, partly up to 1 month (see #9). The bias correction methods compensate for this effect and are able to delay the mean ORS dates. For most of the GCM-RCM combinations, however, the two DM approaches overcompensate for most model runs, while LS follows more closely the overall mean ORS dates of the observation. Amongst the BC methods, the mean ORS dates may vary remarkably, ranging from DOY 131 to DOY 186, i.e. altogether 55 days, while the differences between the CORDEX runs are only 30 days (DOY 165 in #5 vs. DOY 135 in #9). The ranges in the ORS dates are higher between most of the applied BC methods (32, 34, 39, 26 days for LS, PLT, DM_{theo}, and DM_{emp}, respectively) are compared to the range between the 9 different CORDEX runs of the uncorrected RCP4.5 (30 days).

The slightly increased range of the distribution of the P observation data (AgMERRA) compared to the uncorrected CTRL and RCP4.5 runs from Figure 3 propagates through the ORS algorithm and results in a comparatively high widening of the observed ORS dates in relation to the uncorrected CTRL and RCP4.5 runs (Figure 6). Fewer values are classified as outliers after BC. This underlines the need for BC in threshold-based algorithms. All 4 applied bias correction approaches are able to reduce these differences. LS and PLT as well as DM_{theo} and DM_{emp} show similar results in terms of the width of the box and the median value. Similar results are obtained for the other CORDEX GCM-RCM combinations (not shown).

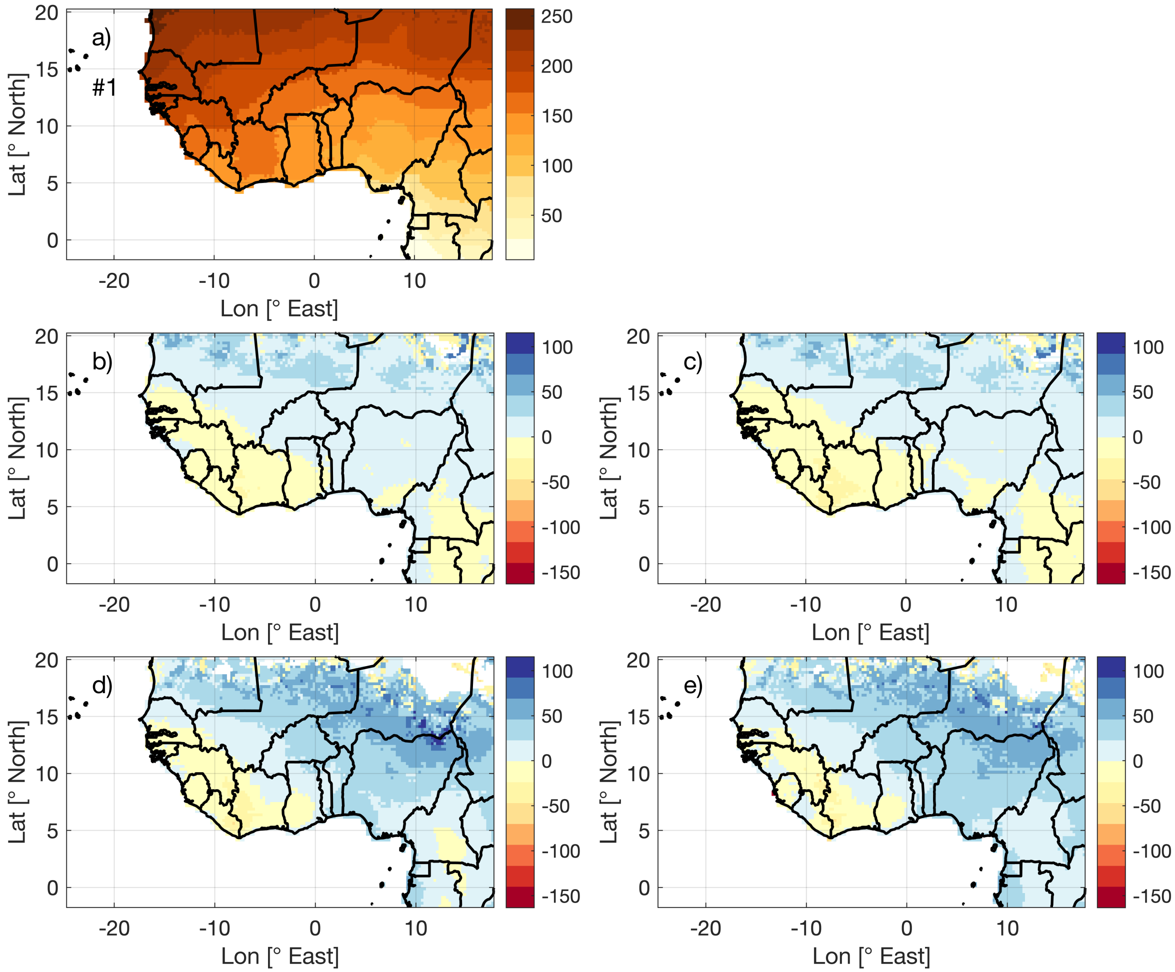


Figure 5. (a) Mean ORS date [DOY] across West Africa, for uncorrected RCP4.5 (2006–2100) run of CORDEX GCM-RCM combination #1 (see Table 22), and (b-e) ORS differences (ORS (debiased RCP4.5) *minus* ORS (uncorr. RCP4.5)). The applied BC methods are: LS (b), PLT (c), DM_{theo} (d), and DM_{emp} (e) (see Table 1). White areas denote regions, in which no ORS date could be calculated due to rainfall deficit. The values are averaged for calculated annual ORS dates for the period 2006–2100.

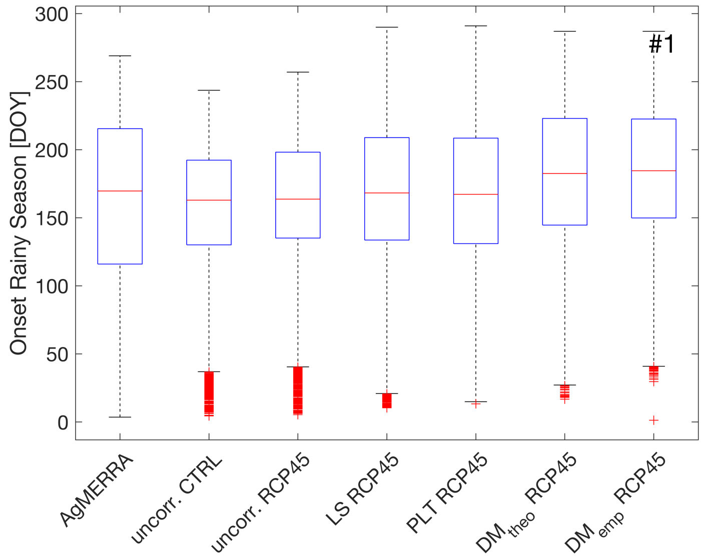


Figure 6. Boxplot of observation (AgMERRA), uncorrected CTRL (both 1980–2005) and RCP4.5 (2006–2100) run of CORDEX GCM-RCMcombination #1 (see Table 22), and bias-corrected (see 1) RCP4.5 runs for the estimated dates of the rainy season's onset [DOY]. The boxplots represent the mean values for each grid cell with the study region.

Table 3. Mean and standard deviation (in brackets) of ORS dates [DOY], calculated for the observation (AgMERRA), the uncorrected CTRL (both 1980–2005) and RCP45 (2006–2100) run of different CORDEX GCM-RCM combinations (see Table 2), and bias-corrected (see Table 1) RCP4.5 runs across the entire West African domain.

ID #	AgMERRA	uncorr. CTRL	uncorr. RCP4.5	LS RCP4.5	PLT RCP4.5	DM_{theo} RCP4.5	DM_{emp} RCP4.5
1	163 (60)	155 (48)	160 (48)	166 (56)	165 (56)	179 (56)	181 (53)
2		150 (57)	153 (57)	157 (61)	158 (60)	175 (60)	179 (55)
3		142 (67)	141 (64)	144 (61)	144 (61)	162 (65)	171 (61)
4		133 (55)	136 (60)	147 (67)	147 (67)	156 (74)	171 (65)
5		159 (59)	165 (58)	167 (59)	165 (58)	181 (57)	186 (53)
6		156 (68)	159 (67)	161 (63)	160 (64)	169 (62)	180 (58)
7		142 (60)	145 (60)	147 (63)	151 (61)	163 (59)	166 (57)
8		148 (67)	149 (63)	156 (62)	157 (60)	167 (57)	168 (54)
9		133 (61)	135 (61)	134 (54)	131 (55)	142 (68)	159 (58)
Range		26	30	32	34	39	26

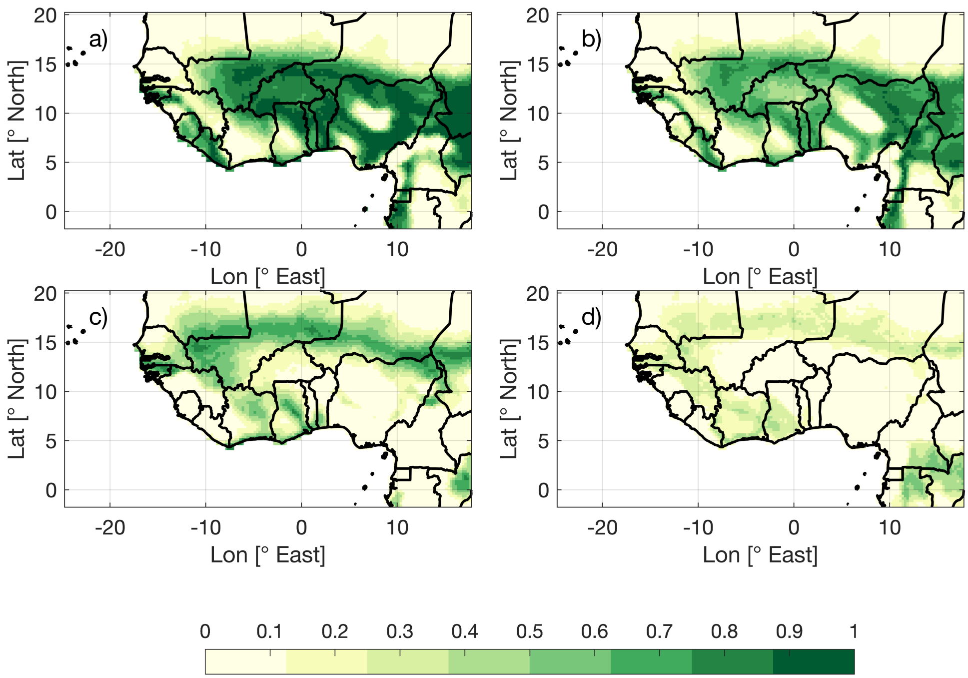


Figure 7. Spatial representation of the mean seasonal (JJAS) CSI for the uncorrected RCP4.5 run of CORDEX GCM-RCMcombination #1 (see Table 2) for the following staple crops: a) groundnut, b) maize, c) pearl millet, and d) sorghum, for the period 2006–2100. Higher CSI values indicate higher climate suitability (see Appendix C).

3.3 Crop Suitability

The crop suitability is calculated based on the different CORDEX GCM-RCM combinations and emission scenario RCP4.5 and for the following four staple crops across West Africa: groundnut, maize, pearl millet, and sorghum. It is noted again that the main focus is on the quantification of the impact of different BC approaches rather than deriving accurate projections.

Figures 7 and 8 illustrate the spatial representation of the mean seasonal (JJAS) CSI for the uncorrected and the bias-corrected RCP4.5 run of CORDEX GCM-RCM combination #1, respectively. By comparing Figures 7 and 8 it can be seen that higher CSI values are obtained after BC. This holds in general for all of the crops. Moreover, the spatial patterns change remarkably. The regions suitable for growing crops (with medium to high CSI values) are relocated to the south and the extend of the regions is shrinking. This effect is more pronounced for sorghum. After BC, the suitable region shrinks remarkably to a relatively small band of about 12° latitude.

Figure 9 shows the spatial distribution of the coefficient of variation (CV) for the CSI, defined as the standard deviation in relation to the mean of the population across West Africa for 4 different staple crops. The CV is thus a measure of the extent of variability, i.e. of the dispersion of a population. The population consists of 8 uncorrected CORDEX GCM-RCM

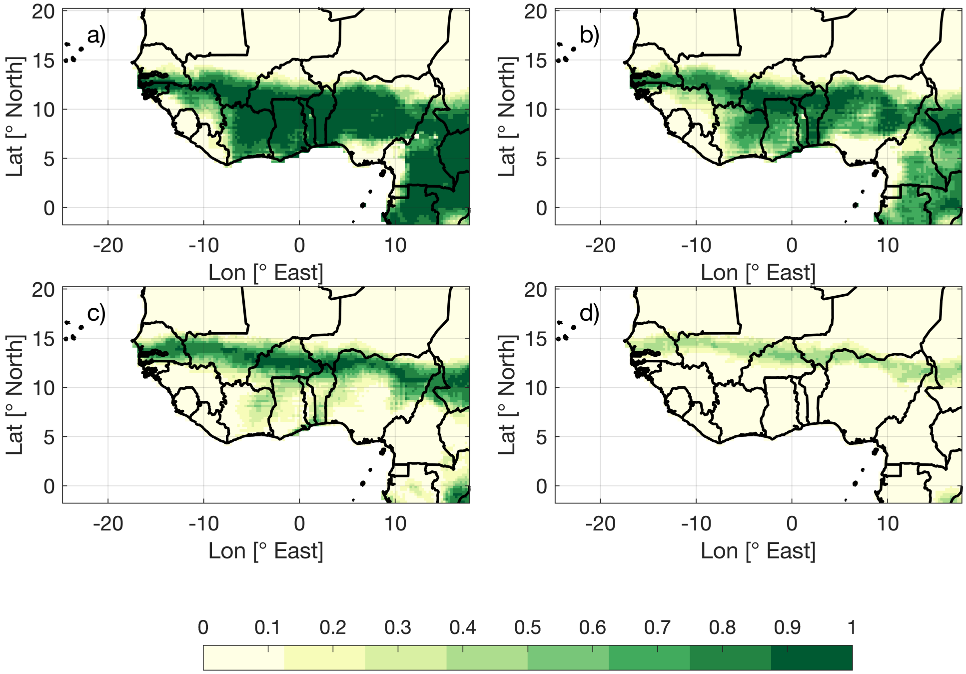


Figure 8. Same as for Figure 7, but based on bias-corrected data for T&P, both corrected by DM_{emp} (randomly chosen).

Table 4. Coefficient of variation (CV) of CSI for the 8 uncorrected CORDEX GCM-RCM runs (left column) and the 16 debiased T&P combinations, for the 4 staple crops averaged over the entire West African domain.

Crop	CV of 8 CORDEX GCM-RCM runs	CV of 16 debiased T&P combinations
Groundnut	1.63	2.58
Maize	1.38	2.61
Pearl millet	1.15	2.46
Sorghum	0.94	2.28

simulations (left column) and 16 permuted T&P runs (right column). High dispersion values greater than 5 are found among the bias-corrected runs, while dispersion is lower and spatially less extended for the corresponding uncorrected runs (Figures 9). Dispersion is greater for groundnut, maize, and pearl millet, but less pronounced for sorghum. The spatially averaged CV values for the entire West African domain is given in Table 4. It can be seen that the CV values are approximately doubled for the 16 debiased T&P combinations, compared to the values of the 9 CORDEX GCM-RCM simulations.

The results suggest that BC increases the variability, and thus has a big impact on the results in agricultural CCIS. Accordingly, impact assessments based on uncorrected or debiased climate simulations should be interpreted with great care.

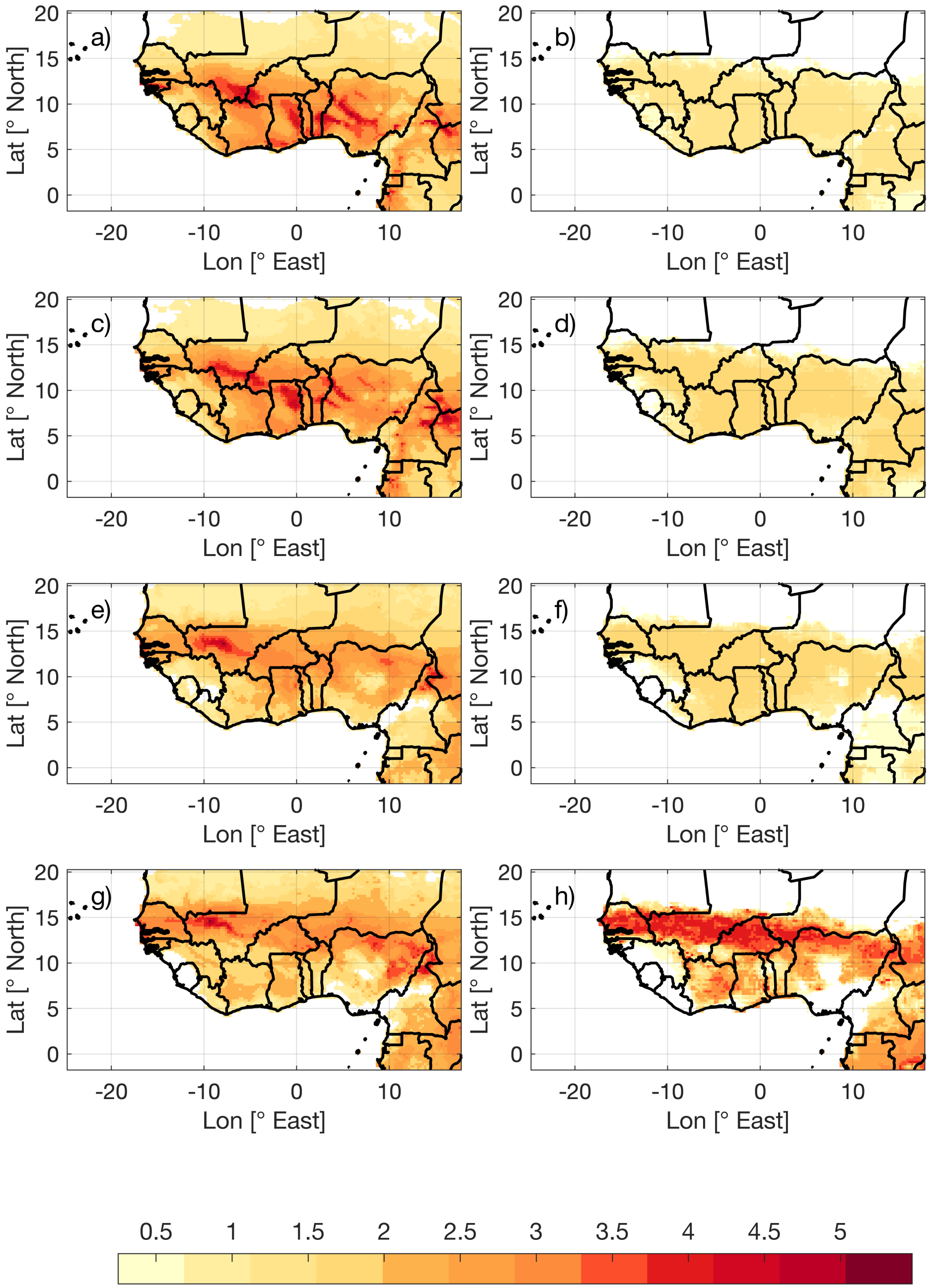


Figure 9. Coefficient of variation (CV) of CSI, for the uncorrected 8 CORDEX GCM-RCM simulations (2006–2100) under RCP4.5 scenario (left column) and all bias-corrected simulations for CORDEC GCM-RCM combination #1, i.e. the 16 permuted T&P BC runs (right column) for groundnut (top row, a) and b)), maize (second row, c) and d)), pearl millet (third row, e) and f)), and sorghum (bottom row, g) and h)). White areas denote regions, in which the mean CSI is zero due to water- or temperature limitations, leading to non-defined CV values.

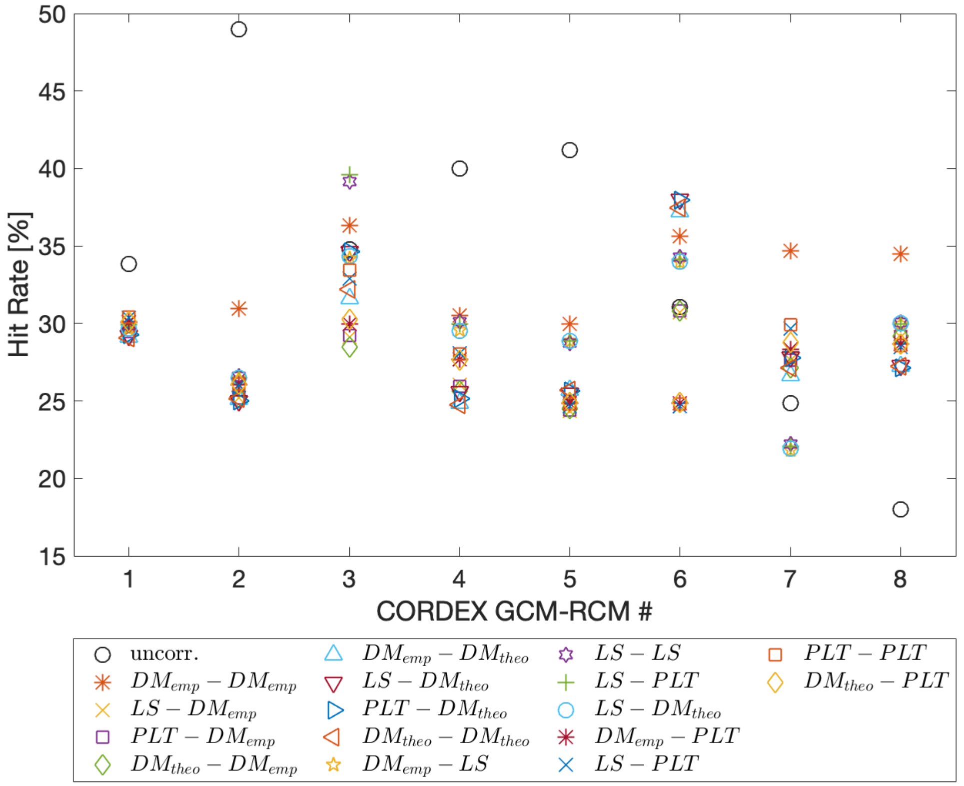


Figure 10. Hit rate [%] between crop with the maximum CSI (2006–2100) and crop with maximum harvested area (2007). The hit rates are shown for the uncorrected and bias-corrected RCP4.5 CORDEX simulations and the 16 bias-corrected T&P data, given as: *BC of T–BC of P*.

Figure 10 illustrates the hit rate between crop with maximum CSI and crop with maximum harvested area, both for uncorrected and bias-corrected RCP4.5 CORDEX simulations. It is seen that some uncorrected CORDEX GCM-RCM simulations have much higher hit rates than others, reaching hit rates up to 50% (#2), whereas the lowest hit rate is less than 20% (#8).

In cases of relatively high hit rates of the uncorrected simulations, none of the bias corrections can exceed the performance of the uncorrected simulations. This is the case for run #1 ,#2, #4, and #5. For simulations with lower hit rates, BC can remarkably increase the hit rates (e.g. runs #7 and #8). Among the 16 combinations of BC for T&P, the combination performing DM_{emp} for both T & P shows the best results.

4 Discussion

The discussion centers around the two research questions defined in this study. Moreover, new and promising avenues or alternatives in CCIS are discussed.

Is bias correction still required in CCIS using state-of-the-art high-resolution (region) climate projections?

It is well-documented that BC may impair the advantages of GCMs and RCMs by altering spatiotemporal field consistency, relations among variables, and by violating conservation principles (Ehret et al., 2012). However, for cases in which the error is predominately time-independent (i.e. systematic), BC is considered as a useful and defensible post-processing tool to transform the output from climate models in such a way that it is more suitable as input in CCIS (IPCC, 2015). In the case of impact models designed for field-scale studies driven by GCMs or RCMs, BC methods have the potential to better represent the spatial patterns of the climate and derived agricultural impacts as well as the increased variability at smaller scales. This could be confirmed by Glotter et al. (2014), who found that regardless of using raw GCM- or RCM input in CCIS for the US, the yield distributions skewed low relative to those driven by observations. Dynamical downscaling of GCMs exacerbates the skew, but applying even simple BC results in yield distributions largely consistent with those generated with historical observations. However, it is well-known that biases may differ regionally, and results from literature studies should be considered to decide whether or not BC shall be applied. For West Africa, for instance, errors in GCMs are more pronounced over the Tropical Atlantic, leading to relatively higher errors in RCMs (inherited by GCM lateral conditions) at the Guinean coast compared to the Sahel (e.g., Paxian et al., 2016; Sultan and Gaetani, 2016). Although we do not explicitly analysed the spatial distribution of the biases inherent in the climate projections, there is strong indication that the BC can have positive and negative impacts for the results obtained in CCIS (demonstrated for observed and simulated CSI across West Africa), depending on the performance level of the uncorrected CORDEX GCM-RCM data. For specific GCM-RCM combinations with high performance of the uncorrected simulations, it is found that BC, in general, can deteriorate the results. Thus, BC cannot be recommended without restrictions before application in CCIS. It is rather suggested to check the skill for uncorrected and debiased data (for the past and present) against climate reference data, before applying the BC for future climate projections. Moreover, the same performance analysis is recommended for the output variable(s) of interest after application of the impact model.

What is the range of uncertainties related to the application of different bias correction approaches in crop-climate impact research?

As highlighted by many studies, uncertainty is *per se* high in crop-climate impact research. It is recommended to apply ‘super-ensembles’ of climate projections in climate change impact assessment studies to address the uncertainties (Tao et al., 2009a). To date, this has typically consisted of multiple GCM-RCM combinations, emission scenarios (future population growth, economic and social development, technological changes, resource depletion, environmental management), and impact models to consider all kind of inherent uncertainties. This can be confirmed by the results of this study. As an example for the uncorrected CORDEX GCM-RCM combinations for West African, we found that the variability from the RCMs can be relatively greater than that from the GCMs. Therefore, in the case of limited resources for CCIS, it is suggested to use an RCM ensemble driven by fewer GCMs rather than a large GCM ensemble combined with one single RCM instead. Additionally, it is worth noting that inherent RCM uncertainties are not considered in this study, but could be large, in particular for precipitation (e.g., Laux et al., 2016; von Trentini et al., 2019). Approaches such as the presented *Representative Temperature and Precipitation (T&P) GCM Subsetting Approach* help to quantify the uncertainty range of climate scenarios with low costs, but leave flexibility for impact modelers to decide based on their specific objectives (e.g. variables to be considered, aggregation time of the variables) (Ruane and McDermid, 2017).

Due to their intrinsic limitations, each BC method has its own uncertainties. As observed in this study, the uncertainties arising from the choice of BC approach can be more than double than those arising from the large-scale input data. This result is in line with Hagemann et al. (2011), who state that BC on the climate change signal may be larger than the signal itself. Moreover, it has been demonstrated in this study that BC can even result in state changes in the climatic regime of a climate scenario member, i.e. a specific member grouped as cool/wet can be grouped as hot/dry after BC. For impact modelers, this means also that they should classify the climate regime of certain climate scenario members after the application of BC.

Due to the above-mentioned reasons, we claim that in general BC is still needed in CCIS. However, we strongly suggest to include an ensemble of different BC methods (rather than relying on one single approach) in the 'super-ensemble' and to quantify the BC-related uncertainties. Thus, the question is not whether to bias correct or not, the question for the future is rather how to bias correct to generate more reliable climate projections.

Are there new and promising avenues or alternatives in CCIS?

One avenue for deriving more reliable climate projections could be the use of multivariate BC approaches (e.g., Vrac, 2018; Cannon, 2018). These have the potential to better retain the physical consistency between the variables as derived from the GCM-RCM combination. The multivariate approach of Cannon (2018) tested in ISIMIP has shown to be prone to artifacts (Lange, 2020). Multivariate approaches also assume that the available GCM-RCM combinations are well calibrated and thus are able to represent the joint dependency structures among the different variables of interest (Laux et al., 2019). Another avenue could be the BC of the GCMs, before applying them in RCMs, i.e. as a pre-processing correction. Studies have suggested that dynamical downscaling of debiased GCM output does offer im- provements in fidelity when reproducing historical climate (e.g., Xu and Yang, 2012). To date, no studies have evaluated the consequences in an impacts model (Glotter et al., 2014).

To circumvent issues and uncertainties related to downscaling and BC, the use of large-area crop models can be seen as an alternative. Such models (e.g. GLAM (Challinor et al., 2005) and MCWLA (Tao et al., 2009b)) can potentially cope with GCM or RCM output directly, but do not allow for detailed process descriptions, and thus cannot address all research questions field-scale models can. Nowadays, they are largely the same models applied at field and large scale in agricultural CCIS.

There are several limitations to this study, which are summarized as follows:

– The study does not address the error structures as well as physical reasons for the biases in the CORDEX data, nor does it consider the inherent issues of the applied BC methods individually. Further research, based on the understanding of the mechanisms that drive the errors and uncertainties in the projected changes is needed (Sultan and Gaetani, 2016).

– This study lacks a comprehensive validation for both the bias-corrected T&P input variables as well as the results obtained by the calculation of the ORS and CSI. Moreover, we assumed the season JJAS (122 days) as the growing cycle for all four crops over the entire study region. The results will deviate if refined location-specific assumptions will be included, however, it is not expected that this will change our main conclusions.

– Only T&P are considered in this study. It is expected that the effect of BC of further variables such as solar radiation, humidity, wind speed in CCIS might add additional uncertainty. The effect might be similar in magnitude as for T&P.

– The impact of the reference data set for the BC is not the subject of this study. We are aware that the choice of the reference data, time period for calibration & validation, interpolation strategy may impact the results. For the reference data, we refer the reader to the study of Parkes et al. (2019), who analysed how the choice of the reference data sets may influence the projected sensitivity of crop yield to weather. In this study, we solely rely on AgMERRA (Ruane et al., 2015), an alternative to be tested in the future is e.g. the WFDE5 dataset (Cucchi et al., 2020).

5 Summary and Conclusions

We investigated the impacts of four bias correction methods on the simulated temperature, precipitation, onset of the rainy season, and crop suitability index for West Africa using CORDEX GCM-RCM model simulations. Major conclusions are that:

– BC can potentially improve the performance of climate projections in agricultural impact studies. The performance gain, however, depends not only on the BC approach, but also on the applied GCM-RCM simulation, the applied impact model, the input variables and their multivariate co-variability as well as the region of interest.

– BC is an additional source of uncertainty in the entire modelling chain of a CCIS. Under certain circumstances, BC approaches can even lead to state changes of the climatic regime of climate scenarios. Uncertainties from a single BC can be larger than the signal of climate change itself. Therefore, single BC methods should be used with great caution, and their uncertainties need to be quantified. We strongly suggest the use of an ensemble of BC methods in CCIS and quantify the BC-inherent uncertainties following a super-ensemble probabilistic assessment within the modelling chain.

The findings of this study will support (regional) impact modelers who apply BC methods for generating robust regional climate change impact assessments on crop yields and other agricultural performance indicators. The bias-corrected T&P ensemble will be provided to encourage further CCIS across West Africa in agriculture, but also from other disciplines, such as hydrology and energy.

6 Data availability

The bias-corrected T&P data generated for this study is provided at PANGAEA (Data Publisher for Earth & Environmental Science), accessible under: https://doi.pangaea.de/10.1594/PANGAEA.922245. CORDEX data can be retrieved from different ESGF nodes. Please follow http://www.csag.uct.ac.za/cordex-africa/how-to-download-cordex-data-from-the-esgf/ to get more information. AgMERRA can be obtained from https://data.giss.nasa.gov/impacts/agmipcf/agmerra/. SPAM data can be obtained from https://www.mapspam.info/data/.

Appendix A: Distance-weighted average remapping

In this study, we employed the distance-weighted average (DWA) remapping method to interpolate time series for precipitation and temperature from the coarser CORDEX climate simulations to the grid of the AgMERRA data, which are used as ’pseudo-obervations’. The approach looks for the nearest four neighbors and calculates the weights w as a function of the distance, using:

$$w = \frac{1/(d+\epsilon)}{\sum_{n=1}^{4}(1/(d_n+\epsilon))},$$

where ϵ is a very small number to prevent division by zero and d is the distance between the destination grid cell and the source grid cell. d_n is the distance of the n neighbor grid cell(s). d is calculated as follows:

$$d = cos^{-1}(cos\ \Theta_d cos\ \Theta_s(cos\ \Phi_d cos\ \Phi_s + sin\ \Phi_d sin\ \Phi_s) + sin\ \Theta_d sin\ \Theta_s),$$

where Θ is the latitude,Φ is the longitude, and the subscripts d and s denote the destination and the source grids, respectively.

Appendix B: Bias Correction Approaches

Please note that in the technical description of the different bias correction methods P and T are used for precipitation and temperature, respectively. V (Variable) is used if the method can be applied more generally, and is not explicitly dedicated to P or T. The terms obs and mod denote the observed (i.e. the AgMERRA data) and modeled time series (i.e. the different CORDEX GCM-RCM combinations), respectively. For calibration, the period 1980–2005 is used, which is applied to correct the CORDEX GCM-RCM combinations of RCP4.5 for 2006–2100. The BC data is provided upon request.

B1 Linear scaling (LS) of precipitation and temperature

An additive correction of the mean value is the simplest form of bias correction of temperature T. The mean bias of the model is calculated for each month m of the period used for calibration and is then added to the model data, i.e. the RCM data of the validation period (hereinafter referred to as Method 1a, equation B1) at every time step T (here: daily). The same additive value can be potentially used for debiasing future climate projections under the assumption of stationary biases.

$$T_{corr,m}(t) = T_{mod,m}(t) + (\overline{T_{obs,m}} - \overline{T_{mod,m}}),$$

where $corr$ denotes the corrected time series. For precipitation data, instead of using a monthly additive term, monthly scaling factors are calculated. The ratio of the mean monthly sums between observed and modeled precipitation is calculated for the calibration period and then used to scale the model data for the validation period at each time step t (here: daily):

$$P_{corr,m}(t) = P_{mod,m}(t) \cdot \frac{\overline{P_{obs,m}}}{\overline{P_{mod,m}}}.$$

In the setting of spatial disaggregation and bias correction as applied in this study, this approach corresponds to the agricultural CCIS well-established 'delta change' method. Instead of impressing the 'delta' to observations, it is impressed here to the spatially disaggregated (based on AgMERRA) historical baseline simulations.

B2 Power law transformation (PLT) of precipitation and temperature

Linear correction methods as described above only adjust the first-order statistical moment, i.e. the mean values, but leave the standard deviation or variance unaffected, because both are multiplied by the same factor. As an alternative, (Shabalova et al., 2003; Leander and Buishand, 2007) used a power transformation method, which corrects also the coefficient of variation (CV). In this BC method, the variable of interest V is transformed to a corrected value V_{corr} for every time step (here: daily) by

$$V_{corr,m}(t) = \overline{a_m} \cdot V_{mod,m}(t)^{\overline{b_m}}.$$

First, the value of b is estimated such that the CV of the corrected daily variable matches that of the observed variable. This is done using a simple root-finding algorithm but the estimated values are also cross-checked using discrete values within a reasonable solution space of b. The factor a is then determined such that the mean of the transformed daily values corresponded with the observed mean. The estimation is done for every running month separately first and later averaged to obtain a long-term monthly mean for both values, i.e. $\overline{a_m}$ and $\overline{b_m}$. This allows one to generate monthly transformation functions for correcting the future climate projections. The resulting value of a depends on b, however, b only depends on the CV and it thus estimated independently from a. Main shortcoming is that this method cannot adequately correct the wet-day probability (Teutschbein and Seibert, 2012).

B3 Distribution mapping (DM) of precipitation and temperature

Another frequently used approach is distribution mapping (DM), originally developed by Sutanto et al. (2019). It is based on the concept of matching the CDFs between the observed and the modeled time series. The obtained transformation function over the full distribution is then used to debias the modeled time series:

$$F_{obs}^{-1}(F_{mod}(V_{mod})),$$

where F_{mod} is the CDF of the variable to be corrected (V_{mod}), and F_{obs}^{-1} is the inverse CDF corresponding to the observed variable (V_{obs}). The DM is thus capable to correct for the mean, variance, but also the distribution, which means that it also preserves the extreme values ((Themeßl et al., 2011; Fang et al., 2015). In literature, there exist various derivatives of this method, all of them having their pros and cons. Sometimes the mapping functions are derived for each season or month separately (Fang et al., 2015), however, in most of the cases they are derived over all values. In this study, we apply the DM over all the values (for each grid cell separately). In this study, we employ the DM based on theoretical distribution functions (DM_{theo}) using a fitted gamma and normal distribution for precipitation and temperature, respectively, as well as DM based on the empirical distribution functions (DM_{emp}).

B3.1 DM for precipitation and temperature based on theoretical distribution functions (DM_{theo})

For correcting daily precipitation, the gamma distribution function is often the first choice of a theoretical distribution function since it is able to model positively skewed data. The probability density function (PDF) for a gamma random variable, where α and β represent the shape and rate parameter, respectively, for the non-zero precipitation values:

$$f_\Gamma(p|\alpha,\beta) = x^{\alpha-1} \cdot \tfrac{1}{\beta^\alpha} \cdot e^{\frac{-p}{\beta}}, p \geq 0, \alpha, \beta > 0,$$

where $\Gamma(.)$ represents the gamma function. In our study we applied the maximum-likelihood estimator exclusively sine it is known to be more accurate and robust compared to the method of moments or least-squares (Piani et al., 2010; Teng et al., 2015).

For temperature values, the Gaussian distribution with mean μ and standard deviation σ is known to give good results (Teutschbein and Seibert, 2012):

$$f_N(t|\mu,\sigma) = \tfrac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{\frac{-(t-\mu)^2}{2\sigma^2}}, t \in \mathbb{R},$$

where $N(.)$ represents the Gaussian (Normal) function. Thus, for correction of precipitation and temperature, equation B3 is applied where the CDFs F_Γ and F_Γ^{-1} as well as F_N and F_N^{-1} are applied, respectively.

B3.2 DM for precipitation and temperature based on the empirical distribution functions (DM_{emp})

The DM method can also be applied without any assumption about the underlying distribution function, i.e. a transformation functions are derived based on the empirical distributions. It is therefore generally applicable to all variables of interest and can correct for the mean, variance, full distribution (quantiles), i.e. also extremes, but also corrects intrinsically the wet-day frequencies. In literature, it is also referred to as quantile mapping (QM). The adjustment can be written in terms of the empirical distribution ($ecdf$) and its inverse ($ecdf^{-1}$) (see (e.g., Fang et al., 2015; Themeßl et al., 2011)):

$$V_{corr}(t) = ecdf_{obs}^{-1}(ecdf_{mod}(V_{mod}(t))).$$

Appendix C: Crop suitability

C1 Calculation of CSI

The suitability of a crop in response to climate variables is assessed by the *Crop Suitability Index* (CSI), following the approach of Hijmans et al. (2001). Based on the total monthly rainfall amounts as well as the monthly mean air temperature, the CSI applies the amount and mean of P&T, respectively, over the length of the growing season for 5 selected staple crops in West Africa. The threshold values for the growing season can be obtained from FAO (1976). Please note that we applied the default Ecocrop parameters in this study and did not validate the threshold values for West Africa.

Based on these threshold values for non-suitable, marginal, and optimal conditions for T&P (see Figure C1), the CSI values are calculated. For optimum (non-suitable) conditions, the CSI is restricted to 1 (0), for marginal conditions the values are linearly interpolated, leading to values ranging from 0 to 1, which can be classified as follows: $0 < 0.25$ (not suitable), $0.25 < 0.5$ (marginally suitable), $0.5 < 0.75$ (suitable), and $0.75 < 1$ (highly suitable) (Ramirez-Villegas et al., 2013).

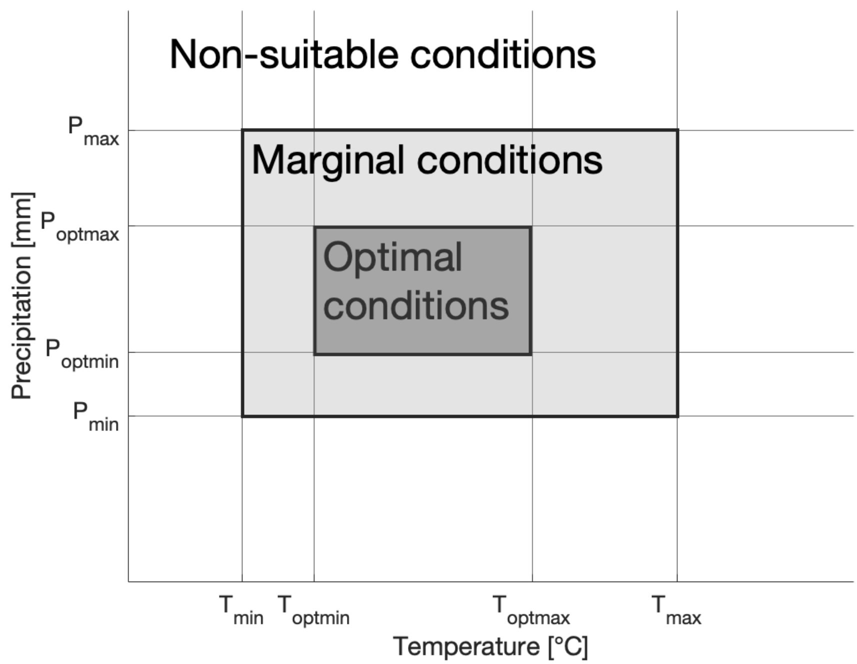


Figure C1. Schematic diagram of CSI, depending on threshold values for T&P. Threshold values for marginal (light grey) and optimal (dark grey) conditions vary for the different crops. T&P values outside the range of marginal conditions are non-suitable conditions for growing.

For the calculation of the CSI, both the start of the growing season as well as the length of the vegetation cycle is important. Both depend on the crop under consideration and variety.

For the planting date, we used the values, obtained by the ORS approach of Laux et al. (2008) for the period 2006–2100 and the RCP4.5 scenario averaged over the entire domain and the different CORDEX RCM-GCM runs. Therefore, we assumed mid-May until June for planting and accounted for the T&P conditions from June until the end of September (JJAS, 122 days in total) as a rough guess for the growing cycle for all crops. It is noted that both the start and the length of the growing cycle are not expected to follow the actual ones since we are only interested in the relative differences due to different BC approaches applied to the CORDEX RCM-GCM simulations. Note that #9 (see Table 2) is excluded from the CSI calculations, because T_{mean} is not available.

C2 Validation of CSI

SPAM (Spatial Production Allocation Model) data is used in this study as a reference to compare the spatial patterns of crop suitability. The SPAM data were re-sampled to $0.25° \times 0.25°$ following the distance-weighted average (DWA) remapping method (see Figure C2) in order that they align with the spatially disaggregated and bias-corrected climate data. For each grid cell, the crop with the maximum harvest area is assigned as a dominant crop. The underlying assumption is that the harvested area of the different crops is correlated to their climatic suitability values. Thus, maximum harvested are is seen as a proxy for the potential suitability of growing crops at a specific location and is therefore applied as an approximate validation of crop suitability (see section 2.7). The grid-wise number of hits between crop with maximum harvest area and maximum CSI values, obtained by different BC methods, is counted.

Author contributions. Conceptualization, methodology, data curation, original draft preparation: P.L.; review and editing: all others; project administration, H.K.; funding acquisition, H.K. and P.L.

Competing interests. The authors declare that no competing interests are present.

Acknowledgements. We acknowledge financial support from the *German Federal Ministry of Education and Research* (BMBF) through the *West African Science Center for Climate Change and Adapted Land Use* (WASCAL) project, the *African Union* (AU) through the *Upscaling climate-smart agriculture and land use practices to enhance production systems in West-Africa* (UPSCALERS) project, the *Joint Programming Initiative on Agriculture, Food Security and Climate Change* (FACCE-JPI) through the *Advanced tools for breeding BARley for Intensive and SusTainable Agriculture under climate change scenarios* project, and the *German Research Foundation* (DFG) through the *FOR 2936: Climate Change and Health in Sub-Saharan Africa* project. We acknowledge valuable discussions from Thomas Gaiser and Seyni Salack on the topic of agricultural impact analyses.

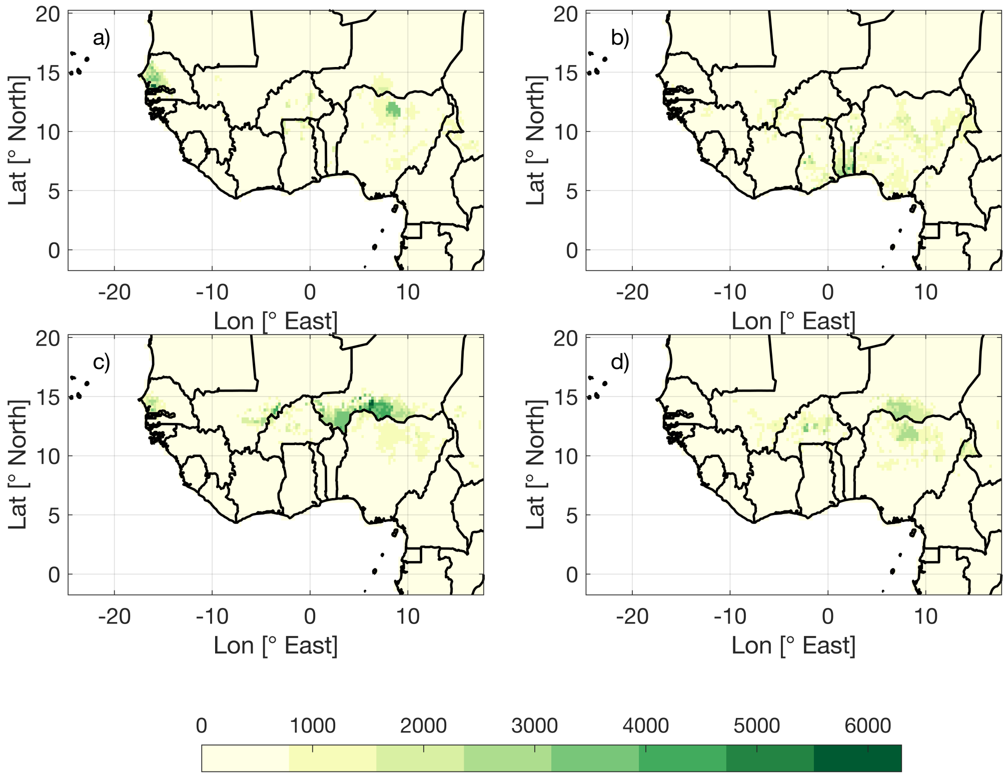


Figure C2. Harvested area (ha) under rainfed conditions for a) groundnut, b) maize, c) pearl millet, and d) sorghum. The data is obtained from the SPAM 2017 v1.1 Sub-Saharan Africa dataset and regridded using distance-weighted average (DWA) remapping method.

References

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P. J., Rotter, R. P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P. K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A. J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L. A., Ingwersen, J., Izaurralde, R. C., Kersebaum, K. C., Mueller, C., Kumar, S. N., Nendel, C., O'Leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M. A., Shcherbak, I., Steduto, P., Stoeckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J. W., Williams, J. R., and Wolf, J.: Uncertainty in simulating wheat yields under climate change, Nature Climate Change, 3, 827–832, https://doi.org/10.1038/nclimate1916, 2013.

Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S., and Dingkuhn, M.: From GCM grid cell to agricultural plot: Scale issues affecting modelling of climate impact, Philosophical Transactions of the Royal Society B: Biological Sciences, 360, 2095–2108, https://doi.org/10.1098/rstb.2005.1741, 2005.

Bindi, M., Palosuo, T., Trnka, M., and Semenov, M.: Modelling climate change impacts on crop production for food security, Climate Research, 65, 3–5, https://doi.org/10.3354/cr01342, 2015.

Cannon, A. J.: Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables, Climate Dynamics, 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2018.

Cannon, A. J., Seibert, J., Stahl, K., Kohn, I., Hakala, K., and Meyer, J.: Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments, Hydrology and Earth System Sciences Discussions, pp. 1–22, https://doi.org/10.5194/hess-2018-317, 2018.

Challinor, A. J., Wheeler, T. R., Slingo, J. M., and Hemming, D.: Quantification of physical and biological uncertainty in the simulation of the yield of a tropical crop using present-day and doubled CO2 climates., Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 360, 2085–2094, https://doi.org/10.1098/rstb.2005.1740, 2005.

Confalonieri, R.: Combining a weather generator and a standard sensitivity analysis method to quantify the relevance of weather variables on agrometeorological models outputs, Theoretical and Applied Climatology, https://doi.org/10.1007/s00704-011-0510-0, 2012.

Cucchi, M., Weedon, G., Amici, A., Bellouin, N., Lange, S., Schmied, H. M., Hersbach, H., and Buontempo, C.: WFDE5: bias adjusted ERA5 reanalysis data for impact studies, Earth System Science Data Discussions, pp. 1–32, https://doi.org/10.5194/essd-2020-28, 2020.

Dieng, D., Laux, P., Smiatek, G., Heinzeller, D., Bliefernicht, J., Sarr, A., Gaye, A. T., and Kunstmann, H.: Performance Analysis and Projected Changes of Agroclimatological Indices Across West Africa Based on High-Resolution Regional Climate Model Simulations, Journal of Geophysical Research: Atmospheres, 123, 7950–7973, https://doi.org/10.1029/2018JD028536, 2018.

Dodd, D. E. and Jolliffe, I. T.: Early detection of the start of the wet season in semiarid tropical climates of Western Africa, International Journal of Climatology, 21, 1251–1262, https://doi.org/10.1002/joc.640, 2001.

Duan, Q. and Duan, A.: The energy and water cycles under climate change, National Science Review, 7, 553–557, https://doi.org/10.1093/nsr/nwaa003, 2020.

Egbebiyi, Crespo, and Lennard: Defining Crop–climate Departure in West Africa: Improved Understanding of the Timing of Future Changes in Crop Suitability, Climate, 7, 101, https://doi.org/10.3390/cli7090101, 2019.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Ewert, F., Rötter, R. P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K. C., Olesen, J. E., van Ittersum, M. K., Janssen, S., Rivington, M., Semenov, M. A., Wallach, D., Porter, J. R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P. P., Bartošová, L., and Asseng, S.: Crop modelling for integrated assessment of risk to food production from climate change, Environmental Modelling and Software, https://doi.org/10.1016/j.envsoft.2014.12.003, 2015.

Famien, A. M., Janicot, S., Ochou, A. D., Vrac, M., Defrance, D., Sultan, B., and Noël, T.: A bias-corrected CMIP5 dataset for Africa using the CDF-t method – a contribution to agricultural impact studies, Earth System Dynamics, 9, 313–338, 2018.

Fang, G. H., Yang, J., Chen, Y. N., and Zammit, C.: Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China, Hydrology and Earth System Sciences, 19, 2547–2559, https://doi.org/10.5194/hess-19-2547-2015, 2015.

FAO: A framework for land evaluation. FAO Soils Bulletin No. 32, Tech. rep., 1976.

Faye, B., Webber, H., Naab, J. B., MacCarthy, D. S., Adam, M., Ewert, F., Lamers, J. P., Schleussner, C. F., Ruane, A., Gessner, U., Hoogenboom, G., Boote, K., Shelia, V., Saeed, F., Wisser, D., Hadir, S., Laux, P., and Gaiser, T.: Impacts of 1.5 versus 2.0 $^{\circ}$c on cereal yields in the West African Sudan Savanna, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aaab40, 2018.

Frieler, K., Levermann, A., Elliott, J., Heinke, J., Arneth, A., Bierkens, M. F. P., Ciais, P., Clark, D. B., Deryng, D., Döll, P., Falloon, P., Fekete, B., Folberth, C., Friend, A. D., Gellhorn, C., Gosling, S. N., Haddeland, I., Khabarov, N., Lomas, M., Masaki, Y., Nishina, K., Neumann, K., Oki, T., Pavlick, R., Ruane, A. C., Schmid, E., Schmitz, C., Stacke, T., Stehfest, E., Tang, Q., Wisser, D., Huber, V., Piontek, F., Warszawski, L., Schewe, J., Lotze-Campen, H., and Schellnhuber, H. J.: The relevance of uncertainty in future crop production for mitigation strategy planning, Earth System Dynamics Discussions, https://doi.org/10.5194/esdd-5-1075-2014, 2014.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., SuzGBRi, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., Van Vliet, M., Biber, M. F., Betts, R. A., Leon Bodirsky, B., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5ğ€°C global warming - Simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geoscientific Model Development, 10, 4321–4345, https://doi.org/10.5194/gmd-10-4321-2017, 2017.

Galmarini, S., Cannon, A., Ceglar, A., Christensen, O., Noblet-Ducoudré, N. d., Dentener, F., Doblas-Reyes, F. J., Dosio, A., Gutierrez, J. M., Iturbide, M., M.Jury, Lange, S., Loukos, H., Maiorano, A., Maraun, D., McGinnis, S., Nikulin, G., Riccio, A., Sanchez, E., Solazzo, E., Toreti, A., Vrac, M., and Zampieri, M.: Adjusting climate model bias for agricultural impact assessment: How to cut the mustard, Climate Services, https://doi.org/10.1016/j.cliser.2019.01.004, 2019.

Glotter, M., Elliott, J., Mcinerney, D., Best, N., Foster, I., and Moyer, E. J.: Evaluating the utility of dynamical downscaling in agricultural impacts projections, https://doi.org/10.1073/pnas.1314787111, 2014.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, Journal of Hydrometeorology, 12, 556–578, https://doi.org/10.1175/2011JHM1336.1, 2011.

Hay, L. E., Wilby, R. L., and Leavesley, G. H.: A Comparison of Delta Change and Downscaled Gcm Scenarios for Three Mounfainous Basins in the United States1, JAWRA Journal of the American Water Resources Association, 36, 387–397, https://doi.org/10.1111/j.1752-1688.2000.tb04276.x, 2000.

Hijmans, R. J., Guarino, L., Cruz, M., and Rojas, E.: Computer tools for spatial analysis of plant genetic resources data : 1 . DIVA-GIS, Plant Genetic Resources Newsletter, https://doi.org/10.1016/j.compositesa.2008.10.018, 2001.

IFPRI: Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 2.0, https://doi.org/10.7910/DVN/PRFF8V, https://doi.org/10.7910/DVN/PRFF8V, 2019.

IPCC: Workshop Report of the Intergovernmental Panel on Climate Change Workshop on Regional Climate Projections and their Use in Impacts and Risk Analysis Studies, Tech. Rep. September, 2015.

IPCC: Summary for Policymakers. In: Global Warming of 1.5°C, Tech. rep., 2018.

Kerandi, N. M., Laux, P., and Kunstmann, H.: Performance of the WRF model to simulate the seasonal and interannual variability of hydrometeorological variables in East Africa: a case study for the Tana River basin in Kenya, Theoretical and Applied Climatology, 130, 401–418, https://doi.org/10.1007/s00704-016-1890-y, 2017.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geoscientific Model Development, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lange, S.: ISIMIP3b bias adjustment fact sheet Observational dataset Bias adjustment and statistical downscaling method, Tech. rep., 2020.

Laux, P., Kunstmann, H., and Bárdossy, A.: Predicting the regional onset of the rainy season in West Africa, International Journal of Climatology, 28, 329–342, https://doi.org/10.1002/joc.1542, 2008.

Laux, P., Jäckel, G., Tingem, R. M., and Kunstmann, H.: Impact of climate change on agricultural productivity under rainfed conditions in Cameroon. A method to improve attainable crop yields by planting date adaptations, Agricultural and Forest Meteorology, 150, 1258–1271, https://doi.org/10.1016/j.agrformet.2010.05.008, 2010.

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, https://doi.org/10.5194/hess-15-2401-2011, 2011.

Laux, P., Nguyen, P. N. B., Cullmann, J., Van, T. P., and Kunstmann, H.: How many RCM ensemble members provide confidence in the impact of land-use land cover change?, International Journal of Climatology, https://doi.org/10.1002/joc.4836, 2016.

Laux, P., Kerandi, N., and Kunstmann, H.: Physics parameterization selection in RCM and ESM simulations revisited: New supporting approach based on empirical copulas, Atmosphere, 10, https://doi.org/10.3390/atmos10030150, 2019.

Leander, R. and Buishand, T. A.: Resampling of regional climate model output for the simulation of extreme river flows, Journal of Hydrology, 332, 487–496, https://doi.org/10.1016/j.jhydrol.2006.08.006, 2007.

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, https://doi.org/10.5194/hess-11-1145-2007, 2007.

Maity, R., Suman, M., Laux, P., and Kunstmann, H.: Bias correction of zero-inflated RCM precipitation fields: A copula-based scheme for both mean and extreme conditions, Journal of Hydrometeorology, pp. 18–0126, https://doi.org/10.1175/JHM-D-18-0126.1, 2019.

Mao, G., Vogl, S., Laux, P., Wagner, S., and Kunstmann, H.: Stochastic bias correction of dynamically downscaled precipitation fields for Germany through copula-based integration of gridded observation data, Hydrology and Earth System Sciences, 11, 7189–7227, https://doi.org/10.5194/hessd-11-7189-2014, 2015.

Michelangeli, P. A., Vrac, M., and Loukos, H.: Probabilistic downscaling approaches: Application to wind cumulative distribution functions, Geophysical Research Letters, https://doi.org/10.1029/2009GL038401, 2009.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., van Meijgaard, E., Samuelsson, P., Sylla, M. B., and Sushama, L.: Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations, Journal of Climate, https://doi.org/10.1175/JCLI-D-11-00375.1, 2012.

Oettli, P., Sultan, B., Baron, C., and Vrac, M.: Are regional climate models relevant for crop yield prediction in West Africa?, Environmental Research Letters, 6, https://doi.org/10.1088/1748-9326/6/1/014008, 2011.

Parkes, B., Higginbottom, T. P., Hufkens, K., Ceballos, F., Kramer, B., and Foster, T.: Weather dataset choice introduces uncertainty to estimates of crop yield responses to climate variability and change, Environmental Research Letters, 14, 124 089, https://doi.org/10.1088/1748-9326/ab5ebb, 2019.

Paxian, A., Sein, D., Panitz, H. J., Warscher, M., Breil, M., Engel, T., Tödter, J., Krause, A., Cabos Narvaez, W. D., Fink, A. H., Ahrens, B., Kunstmann, H., Jacob, D., and Paeth, H.: Bias reduction in decadal predictions of West African monsoon rainfall using regional climate models, Journal of Geophysical Research, https://doi.org/10.1002/2015JD024143, 2016.

Peng, B., Guan, K., Tang, J., Ainsworth, E. A., Asseng, S., Bernacchi, C. J., Cooper, M., Delucia, E. H., Elliott, J. W., Ewert, F., Grant, R. F., Gustafson, D. I., Hammer, G. L., Jin, Z., Jones, J. W., Kimm, H., Lawrence, D. M., Li, Y., Lombardozzi, D. L., Marshall-Colon, A., Messina, C. D., Ort, D. R., Schnable, J. C., Vallejos, C. E., Wu, A., Yin, X., and Zhou, W.: Towards a multiscale crop modelling framework for climate change adaptation assessment, https://doi.org/10.1038/s41477-020-0625-3, 2020.

Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99, 187–192, https://doi.org/10.1007/s00704-009-0134-9, 2010.

Pirttioja, N., Carter, T. R., Fronzek, S., Bindi, M., Hoffmann, H., Palosuo, T., Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P., Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M. F., Dumont, B., Ewert, F., Ferrise, R., François, L., Gaiser, T., Hlavinka, P., Jacquemin, I., Kersebaum, K. C., Kollas, C., Krzyszczak, J., Lorite, I. J., Minet, J., Minguez, M. I., Montesino, M., Moriondo, M., Müller, C., Nendel, C., Öztürk, I., Perego, A., Rodríguez, A., Ruane, A. C., Ruget, F., Sanna, M., Semenov, M. A., Slawinski, C., Stratonovitch, P., Supit, I., Waha, K., Wang, E., Wu, L., Zhao, Z., and Rötter, R. P.: Temperature and precipitation effects on wheat yield across a European transect: A crop model ensemble analysis using impact response surfaces, Climate Research, 65, 87–105, https://doi.org/10.3354/cr01322, 2015.

Pirttioja, N., Palosuo, T., Fronzek, S., Räisänen, J., Rötter, R. P., and Carter, T. R.: Using impact response surfaces to analyse the likelihood of impacts on crop yield under probabilistic climate change, Agricultural and Forest Meteorology, 264, 213–224, https://doi.org/10.1016/j.agrformet.2018.10.006, 2019.

Räisänen, J. and Räty, O.: Projections of daily mean temperature variability in the future: Cross-validation tests with ENSEMBLES regional climate simulations, Climate Dynamics, https://doi.org/10.1007/s00382-012-1515-9, 2013.

Ramarohetra, J., Pohl, B., and Sultan, B.: Errors and uncertainties introduced by a regional climate model in climate impact assessments: Example of crop yield simulations in West Africa, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/12/124014, 2015.

Ramirez-Villegas, J., Jarvis, A., and Läderach, P.: Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum, Agricultural and Forest Meteorology, https://doi.org/10.1016/j.agrformet.2011.09.005, 2013.

Rauch, M., Bliefernicht, J., Laux, P., Salack, S., Waongo, M., and Kunstmann, H.: Seasonal Forecasting of the Onset of the Rainy Season in West Africa, Atmosphere, 10, 528, https://doi.org/10.3390/atmos10090528, 2019.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, Journal of Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Rosenzweig, C., Jones, J., Hatfield, J., Ruane, A., Boote, K., Thorburn, P., Antle, J., Nelson, G., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., and Winter, J.: The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies, Agricultural and Forest Meteorology, 170, 166–182, https://doi.org/10.1016/j.agrformet.2012.09.011, 2013.

Rötter, R. P., Carter, T. R., Olesen, J. E., and Porter, J. R.: Crop-climate models need an overhaul, https://doi.org/10.1038/nclimate1152, 2011a.

Rötter, R. P., Palosuo, T., Pirttioja, N. K., Dubrovsky, M., Salo, T., Fronzek, S., Aikasalo, R., Trnka, M., Ristolainen, A., and Carter, T. R.: What would happen to barley production in Finland if global warming exceeded 4°C? A model-based assessment, European Journal of Agronomy, https://doi.org/10.1016/j.eja.2011.06.003, 2011b.

Rötter, R. P., Höhn, J. G., and Fronzek, S.: Projections of climate change impacts on crop production: A global and a Nordic perspective, Acta Agriculturae Scandinavica A: Animal Sciences, 62, 166–180, https://doi.org/10.1080/09064702.2013.793735, 2012.

Rötter, R. P., Hoffmann, M. P., Koch, M., and Müller, C.: Progress in modelling agricultural impacts of and adaptations to climate change, Current Opinion in Plant Biology, https://doi.org/10.1016/j.pbi.2018.05.009, 2018.

Ruane, A. C. and McDermid, S. P.: Selection of a representative subset of global climate models that captures the profile of regional changes for integrated climate impacts assessment, Earth Perspectives, 4, 1, https://doi.org/10.1186/s40322-017-0036-4, 2017.

Ruane, A. C., Goldberg, R., and Chryssanthacopoulos, J.: Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation, Agricultural and Forest Meteorology, 200, 233–248, https://doi.org/10.1016/j.agrformet.2014.09.016, 2015.

Schleiss, M., Jaffrain, J., and Berne, A.: Statistical analysis of rainfall intermittency at small spatial and temporal scales, Geophysical Research Letters, https://doi.org/10.1029/2011GL049000, 2011.

Shabalova, M. V., van Deursen, W. P. a., and Buishand, T. a.: Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model, Climate Research, 23, 233–246, https://doi.org/10.3354/cr023233, 2003.

Siegmund, J., Bliefernicht, J., Laux, P., and Kunstmann, H.: Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling, Journal of Geophysical Research Atmospheres, 120, 7316–7339, https://doi.org/10.1002/2014JD022692, 2015.

Sultan, B. and Gaetani, M.: Agriculture in West Africa in the twenty-first century: Climate change and impacts scenarios, and potential for adaptation, https://doi.org/10.3389/fpls.2016.01262, 2016.

Sutanto, S. J., van der Weert, M., Wanders, N., Blauhut, V., and Van Lanen, H. A.: Moving from drought hazard to impact forecasts, Nature Communications, 10, https://doi.org/10.1038/s41467-019-12840-z, 2019.

Tao, F., Yokozawa, M., and Zhang, Z.: Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis, Agricultural and Forest Meteorology, 149, 831–850, https://doi.org/10.1016/j.agrformet.2008.11.004, 2009a.

Tao, F., Zhang, Z., Liu, J., and Yokozawa, M.: Modelling the impacts of weather and climate variability on crop productivity over a large area: A new super-ensemble-based probabilistic projection, Agricultural and Forest Meteorology, 149, 1266–1278, https://doi.org/10.1016/j.agrformet.2009.02.015, 2009b.

Tao, F., Rötter, R. P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M. I., Semenov, M. A., Kersebaum, K. C., Nendel, C., Specka, X., Hoffmann, H., Ewert, F., Dambreville, A., Martre, P., Rodríguez, L., Ruiz-Ramos, M., Gaiser, T., Höhn, J. G., Salo, T., Ferrise, R., Bindi, M., Cammarano, D., and Schulman, A. H.: Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments, Global Change Biology, https://doi.org/10.1111/gcb.14019, 2018.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J. P.: How does bias correction of regional climate model precipitation affect modelled runoff?, Hydrology and Earth System Sciences, 19, 711–728, https://doi.org/10.5194/hess-19-711-2015, 2015.

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, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530–1544, https://doi.org/10.1002/joc.2168, 2011.

Van Bussel, L. G., Ewert, F., and Leffelaar, P. A.: Effects of data aggregation on simulations of crop phenology, Agriculture, Ecosystems and Environment, https://doi.org/10.1016/j.agee.2010.03.019, 2011.

von Trentini, F., Leduc, M., and Ludwig, R.: Assessing natural variability in RCM signals: comparison of a multi model EURO-CORDEX ensemble with a 50-member single model large ensemble, Climate Dynamics, 53, 1963–1979, https://doi.org/10.1007/s00382-019-04755-8, 2019.

Vrac, M.: Multivariate bias adjustment of high-dimensional climate simulations : The " Rank Resampling for Distributions and Dependences " (R 2 D 2) Bias Correction, Hydrology and Earth System Science, 22, 3175–3196, https://doi.org/10.5194/hess-22-3175-2018, 2018.

Waongo, M., Laux, P., Traoré, S. B., Sanon, M., and Kunstmann, H.: A crop model and fuzzy rule based approach for optimizing maize planting dates in Burkina Faso, West Africa, Journal of Applied Meteorology and Climatology, https://doi.org/10.1175/JAMC-D-13-0116.1, 2013.

Webber, H., Hoffmann, M., and Rezaei, E. E.: Crop Models as Tools for Agroclimatology, chap. 19, agronomy m edn., https://doi.org/10.2134/agronmonogr60.2016.0025, 2018.

Webber, H., Hoffmann, M., and Rezaei, E. E.: Crop Models as Tools for Agroclimatology, in: Agroclimatology: Linking Agriculture to Climate, Volume 60, edited by Hatfield, J. L., Sivakumar, M. V., and Prueger, J. H., chap. 19, pp. 519–546, Wiley, https://doi.org/10.2134/agronmonogr60.2016.0025, 2020.

White, J. W., Hoogenboom, G., Kimball, B. A., and Wall, G. W.: Methodologies for simulating impacts of climate change on crop production, Field Crops Research, https://doi.org/10.1016/j.fcr.2011.07.001, 2011.

Wilby, R. L., Wilby, R. L., Charles, S. P., Charles, S. P., Zorita, E., Zorita, E., Timbal, B., Timbal, B., Whetton, P., Whetton, P., Mearns, L. O., and Mearns, L. O.: Guidelines for Use of Climate Scenarios Developed From Statistical Downscaling Methods, Supporting material of the Intergovernmental Penel on Climate Change, pp. 1–27, 2004.

Xu, Z. and Yang, Z. L.: An improved dynamical downscaling method with GCM bias corrections and its validation with 30 years of climate simulations, Journal of Climate, 25, 6271–6286, https://doi.org/10.1175/JCLI-D-12-00005.1, 2012.

Zhao, G., Siebert, S., Enders, A., Rezaei, E. E., Yan, C., and Ewert, F.: Demand for multi-scale weather data for regional crop modeling, Agricultural and Forest Meteorology, 200, 156–171, https://doi.org/10.1016/j.agrformet.2014.09.026, 2015.