



Climate data uncertainty for agricultural impact assessments in West Africa

Paula Romanovska¹ · Stephanie Gleixner¹ · Christoph Gornott^{1,2}

Received: 10 January 2022 / Accepted: 10 March 2023 / Published online: 30 March 2023
© The Author(s) 2023

Abstract

Reliable information on climate impacts can support planning processes to make the agricultural sector—which has cascading effects on food security, livelihoods and the security situation—more resilient. Subsequently, uncertainties in past and future climate data need to be decreased and better understood. In this study, we analysed the quality and limitations of different past and future climate data sets to be used for agricultural impact assessments in West Africa. The high differences between the three analysed past climate data sets underline the high observational uncertainty in West Africa and show the influence of selecting the observational data set for the bias-adjustment of climate model data. The ten CMIP6 (Coupled Model Inter-comparison Project Phase 6) models show regional and model-dependent biases with similar systematic biases as have been observed in earlier CMIP versions. Although the bias-adjusted version of this data (ISIMIP3b—Inter-Sectoral Impact Model Intercomparison Project) aligns overall well with observations, we could detect some regional strong deviations from observations for some agroclimatological indices. The use of the multi-model ensemble mean has resulted in an improved agreement of CMIP6 and the bias-adjusted ISIMIP3b data with observations. Choosing a sub-ensemble of bias-adjusted models could only improve the performance of the ensemble mean locally but not over the whole region. Therefore, our results suggest the use of the whole model ensemble for agricultural impact assessments in West Africa. While averaging the impact results over all climate models can serve as a best guess, the spread of the results over all models should be considered to give insights into the uncertainties. This study can support agricultural impact modelling in quantifying climate risk hotspots as well as suggesting suitable adaptation measures to increase the resilience of the agricultural sector in West Africa.

1 Introduction

To support the United Nations (UN) Integrated Strategy for the Sahel (UNISS), the UN wants to increase the preparedness to respond to growing multi-causal risk in the Sahel by improving the availability of high-quality data. The agricultural sector is the main source of livelihood in this region

and strongly influences development and peace-building. Hence, the data quality for climate impacts on the agriculture sector plays an important role in securing future livelihoods under climate risk (UN 2018, 2021).

West Africa is characterised by high variability in climate (Kothe et al. 2014), has a population that is expected to double from 2020 to 2050 (United Nations 2019) and is home to a population strongly reliant on rainfed agriculture (Schewe & Levermann 2017). Agricultural production is largely weather-dependent in West Africa and the risk of yield losses is increasing further under increasing temperatures (Sultan & Gaetani 2016). Even low future emissions scenarios were shown to push farming globally outside of its historical regimes under simulation of the latest climate and crop models (Jägermeyr et al. 2021). To adequately address climate risks and avoid further pressure on food security, evidence-based information on climate impacts and guidance on the suitability of adaptation measures is required. This is increasingly recognised by individual countries and the international community (e.g. in the UNISS strategy).

✉ Paula Romanovska
paula.romanovska@pik-potsdam.de

Stephanie Gleixner
gleixner@pik-potsdam.de

Christoph Gornott
gornott@pik-potsdam.de

¹ Potsdam Institute for Climate Impact Research (PIK),
Member of the Leibniz Association, 14473 Potsdam,
Germany

² Agroecosystem Analysis and Modelling, Faculty
of Organic Agricultural Sciences, University of Kassel,
37213 Witzenhausen, Germany

Simulations of regional impacts of climate change on crop production are strongly influenced by the climate data used as input. The selection of climate-forcing data is most influential in regions with high uncertainties in past climate data and where agricultural production varies greatly under climate variability (Ruane et al. 2021). Both are the case in West Africa, calling for an improved understanding of past and future climate data in the region.

The quality of past data over West Africa is limited by spatial and temporal gaps due to scarcely distributed long-term weather stations (Funk et al. 2019; Satgé et al. 2020). Validated data from satellites supports closing the existing gaps and reanalysis products developed with weather models are further providing solutions to missing data (Vy 2021). Nevertheless, uncertainties remain.

In recent decades, there has been substantial progress in climate modelling, but only limited improvements have been made over Sub-Saharan Africa (James et al. 2018). This can be seen in particular in West Africa, where the representation of the West African monsoon shows systematic biases in climate models and future climate projections carry high uncertainties. Especially for precipitation projections, climate models show strong disagreement, varying between much wetter and drier future conditions (Druryan 2011; James et al. 2018). Climate models of the Coupled Model Inter-comparison Project (CMIP) phase 3 and 5 project future climate that is highly model-dependent (Roehrig 2010; Sow et al. 2020). While CMIP6 models tend to simulate the West African monsoon better than CMIP5 models, common biases and large intermodal spreads persist (Wang et al. 2021), limiting the reliability of future projections in the region.

Simulations of general circulation models (GCMs) show too strong biases in mean climatology and extremes and the models are often on a too low resolution to produce meaningful results when applied in impact models to, e.g. simulate effects on the agricultural sector. Bias-adjusted data, meaning data that is statistically altered to match the current climate and often statistically downscaled, is preferred. While bias-adjustment is a necessary step, it cannot fix major problems of a climate model (Maraun 2016) but rather hides that physical processes might not be represented correctly by the model. Common bias-adjustment methods can easily adjust the mean climatology, but an understanding of how different bias-adjustment methods influence extremes or the trend is largely missing (Casanueva et al. 2020). Gampe et al. (2019) showed exemplary that the quality of the reference data set used for the bias-adjustment highly influences the bias-adjusted data, especially the representation of extremes. Famien et al. (2018) highlight the difference between bias-adjusted data sets depending on the bias-adjustment method and the reference data set that also leads to differing future yield projections. Thus,

the bias-adjusted data entails the model uncertainty as well as additional uncertainty introduced by the selection of the bias-adjustment method and observational uncertainty coming from the choice of the reference data set.

In this study, we analysed 10 CMIP6 models as well as their bias-adjusted versions developed in phase 3b of the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP3b) (Lange 2019; Lange & Büchner 2021). ISIMIP3b historical simulations and future projections are widely used by an international network of agricultural impact modelers and are one of few resources of bias-adjusted climate data designed for impact assessment and available globally. However, analyses of the quality and limitations of the data in West Africa are largely missing. The aim of our study is thus to contribute to a better understanding of uncertainties in climate data and offer guidance on the use of the CMIP6 and ISIMIP3b data for agricultural impact assessments in West Africa.

We analysed the uncertainties related to the data sets by comparing modelled data and observations in two steps:

- 1) To understand the general ability of the CMIP6 models to represent the main features of the West African climate, we compare monthly precipitation and temperature as well as past trends in the historical runs to three different observational data sets. Including observational data sets from different sources supports accounting for observational uncertainties.
- 2) Next, we assessed the bias-adjusted ISIMIP3b data. We tested its ability to simulate different agroclimatological indices in comparison to the observational data set used for the bias-adjustment. The indices were selected according to their reported influence on crop production and are ranging from rainy season onset and number of dry spells to temperature extremes.

The present study is structured as follows: in ‘Sect. 2’, we present the climate data used, the study region and its climate, and the methodology including all the definitions of indices. In ‘Sect. 3’, we present and discuss the results and finally, conclusions are given in ‘Sect. 4’.

2 Data and methods

2.1 Climate and agricultural data sets

To evaluate the performance of the climate models, the simulated data is compared against observational and reanalysis datasets. To account for uncertainties in the observational data sets and understand the influence of the selection of a reference data set used for the bias-adjustment, three observational data sets are included (Table 1). The data sets were

Table 1 Specification of observational and reanalysis data sets

Data set	Variable(s)	Covered period	Frequency	Resolution
ERA5	tas, pr	1979–2019	Hourly	0.25° × 0.25°
CHIRTS-daily	tas ¹	1983–2016	Daily	0.05° × 0.05°
CHIRPS-2.0	pr	1981–near-real time	Daily	0.05° × 0.05°
W5E5v2	tas, pr	1979–2019	Daily	0.5° × 0.5°

¹Obtained by calculating the mean between Tmax and Tmin.

selected based on their wide use, their availability over West Africa and daily resolution as well as their independence from one another. The data sets cover different available sources, namely observations, satellite data and/or data constructed through reanalysis. The three data sets cover precipitation and temperature data and are:

- (i) ERA5 (fifth generation of European Centre for Medium-Range Weather Forecast (ECMWF) atmospheric reanalyses), the latest high-resolution reanalysis produced by ECMWF. It combines vast amounts of historical observations into global estimates using advanced modelling and data assimilation systems (Hersbach et al. 2019, 2020).
- (ii) CHIRTS-daily, a high-resolution (0.05° × 0.05°) daily maximum and minimum temperature data set. CHIRTS-daily is based on the CHIRTSmax data set (Climate Hazards Center InfraRed Temperature with Stations). CHIRTSmax combines a global network of around 15,000 in situ station observations with remote sensing infrared land surface emission temperatures and thereby estimates the monthly mean maximum 2-m air temperature. ERA5 was used to disaggregate the monthly CHIRTSmax to daily maximum (Tmax) and minimum temperature (Tmin) values (Chris Funk et al. 2019; Verdin et al. 2020).

CHIRPS (Climate Hazards Center Infrared Precipitation with Station data) is the respective gridded precipitation data set incorporating satellite imagery with in situ station data (C. Funk et al. 2014).

- (iii) W5E5 v2.0, a dataset that was compiled to support the bias-adjustment of climate input data for the impact assessments carried out in ISIMIP3b (Cucchi et al. 2020; Lange et al. 2021). W5E5 is a combination of WFDE5 (WATCH Forcing Data methodology applied to ERA5 data and monthly values are bias-adjusted with CRU, Weedon et al. 2014; Cucchi et al. 2020) over land and ERA5 over the ocean. Additionally, W5E5 precipitation is adjusted to GPCC (Global Precipitation Climatology Centre, Schneider et al. 2018) precipitation over land and GPCP (Global Precipitation Climatology Project, Adler et al. 2003) over the ocean.

W5E5 and CHIRTS are partially dependent on ERA5 constraining an analysis of independent data sets. For both variables (temperature—tas and precipitation—pr), we analysed three data sets (compare Table 1). The observational and reanalysis data sets are for reasons of simplicity in the following called observational data sets.

Table 2 Specification of the ten CMIP6 climate models

Model	Acronym	Institute	Resolution (atmosphere) ¹	Reference
CanESM5	Can	Canadian Earth System Model	T63 linear Gaussian grid (~2.8125° × 2.8125°); 49 levels	(Swart et al. 2019)
CNRM-CM6-1	CNCM	Centre National de Recherches Météorologiques	T127 (~100 km); 91 levels	(Voldoire 2018)
CNRM-ESM2-1	CNES	Centre National de Recherches Météorologiques	T127 (~100 km); 91 levels	(Voldoire 2019)
EC-Earth3	EC	EC-EARTH consortium	TL255 (~70 km); 91 levels	(EC-Earth 2019)
GFDL-ESM4	GFDL	Geophysical Fluid Dynamics Laboratory	1.25° × 1.00°; 49 levels	(Krasting et al. 2018)
IPSL-CM6A-LR	IPSL	Institute Pierre-Simon Laplace	N96 (2.5° × 1.259°); 79 levels	(Boucher et al. 2018)
MIROC6	MIROC	Model for Interdisciplinary Research on Climate	T85 (~1.4° × 1.4°); 81 levels	(Tatebe & Watanabe 2018)
MPI-ESM1-2-HR	MPI	Max Planck Institute Earth System Model	T127 (0.94° × 0.94°); 95 levels	(von Storch et al. 2017)
MRI-ESM2-0	MRI	Meteorological Research Institute	TL159 (~120 km); 80 levels	(Yukimoto et al. 2019)
UKESM1-0-LL	UKE	Met Office Hadley Centre	N96 (1.875° × 1.25°); 85 level	(Tang et al. 2019)

¹Information on resolution based on Richter and Tokinaga (2020)

Our analysis was carried out with temperature and precipitation data of ten GCMs from the CMIP6 historical simulations (specified in Table 2) as well as their bias-adjusted and downscaled counterparts of the ISIMIP3b data (Lange 2019; Lange & Büchner 2021). The ten GCMs are CanESM5 (short: Can), CNRM-ESM2-1 (short: CNES), CNRM-CM6-1 (short: CNCM), EC-Earth3 (short: EC), GFDL-ESM4 (short: GFDL), IPSL-CM6A-LR (short: IPSL), MIROC6 (short: MIROC), MPI-ESM1-2-HR (short: MPI), MRI-ESM2-0 (short: MRI) and UKESM1-0-LL (short: UKE). ISIMIP3b data was obtained by applying a trend preserving bias-adjustment at the spatial resolution of the CMIP6 data and a statistical downscaling method applied to the bias-adjusted simulations to increase and unify the resolution. The bias-adjustment method uses parametric quantile mapping aiming at a robust bias-adjustment of all percentiles of a distribution that preserves the trends in these percentiles (Lange 2019).

The subset of CMIP6 models selected within ISIMIP are representing a subset with relative structural independence and a fair representation of the whole ensemble in terms of climate sensitivity, with a tendency towards high climate sensitivity (Lange 2021; Nijse et al. 2020).

The different CMIP6 historical simulations are all provided on different grids. For the comparison of CMIP6 data to observational data, all observational datasets and CMIP6 model data have been regridded to a spatial resolution of $2.0^\circ \times 2.0^\circ$ using conservative remapping of CDO (Schulzweida 2019). Due to the inconsistent availability of data over the ocean, only data points over land are taken into consideration for all the analyses where different data sets are compared with one another by statistical means. The maps contain data over the ocean if available. ISIMIP3b data is provided on a $0.5^\circ \times 0.5^\circ$ grid. Since it is not the focus to compare the bias-adjusted and non-bias-adjusted versions with each other, the different grid scales for non-bias-adjusted CMIP6 data and ISIMIP3b data are unproblematic. To distinguish between the bias-adjusted and non-bias-adjusted versions of the GCMs in the following text, the bias-adjusted data is written in *italic*.

As a compromise between different time frames available by models and observations, the analysis focuses on a 30-year period between 1985 and 2014 available for all data sets. We consider this period sufficient to detect meaningful mean values of climate and agroclimatological indices not dominated by extreme values. At the same time, this is a period with higher observational data quality than earlier periods.

In addition to climate data, crop planting and harvesting dates were used in this study to define growing seasons. These were obtained from the Center for Sustainability and the Global Environment (Sacks et al. 2010) on a $0.5^\circ \times 0.5^\circ$ grid. The dataset is the result of digitising and georeferencing existing observations of crop planting

and harvesting dates from six sources and merging these crop calendar maps with monthly climatologies from CRU (Climate Research Unit gridded Time Series, Harris et al. 2020) with the help of derived climate statistics. We based our definition of the growing season on the maize crop calendar, as maize is widely grown in this region and assuming that the high data availability for maize leads to the most reliable results. We thereby defined the growing season as the time between the earliest planting date and the latest harvesting date of maize. A visualisation of the growing season dates can be found in the SI.

2.2 Study region

The model evaluation is done over the West African region from 19°W to 14°E and 3°S to 20°N entailing the main part of the Sahel region as defined in UNISS, a UN project aiming at increased availability of high-quality data to assess risks (UN 2018). The region is characterised by diverse climate conditions ranging from very dry conditions in the Sahara to a tropical climate near the Gulf of Guinea. The annual cycle of precipitation over the whole region is driven by the West African monsoon (WAM). The WAM is mainly driven by a temperature gradient between the ocean and land surface. The high temperatures over the Sahara in the northern hemisphere summer create a heat low which drives the moist air from the Atlantic inland towards the Sahel. The moisture condenses over land and thus releases latent heat, which again reinforces the temperature gradient and thereby the monsoon. This brings precipitation inland from April to October, with major differences between the wet Guinean region, where precipitation patterns show two peaks—one in June and one in September—and the arid Sahelian region showing a single peak in precipitation around August (Herzschuh et al. 2014; Kothe et al. 2014; Sow et al. 2020).

Due to substantially different climatological conditions as well as past trends within the West African region, we introduced three sub-regions within the whole region for parts of the analysis: the Sahel, Savanna and Guinea regions (Fig. 1). The north–south boundaries of the sub-regions are defined following (Abiodun et al. 2012; Dieng et al. 2018; Omotosho and Abiodun 2007) as 12°N – 16°N (Sahel), 8°N – 12°N (Savanna) and 4°N – 8°N (Guinea).

Due to predominantly rainfed agriculture, the agricultural season is largely limited to the rainy season. Most of the precipitation occurs in the main monsoon season between June and September. Rainy season length and precipitation amount highly differ between the regions. The Sahel and Savanna zones have a short rainy season of between 6 and 2 months with a peak in August and decreasing precipitation amounts towards the north. The Guinea zone receives more precipitation over a longer period, most of it falling between

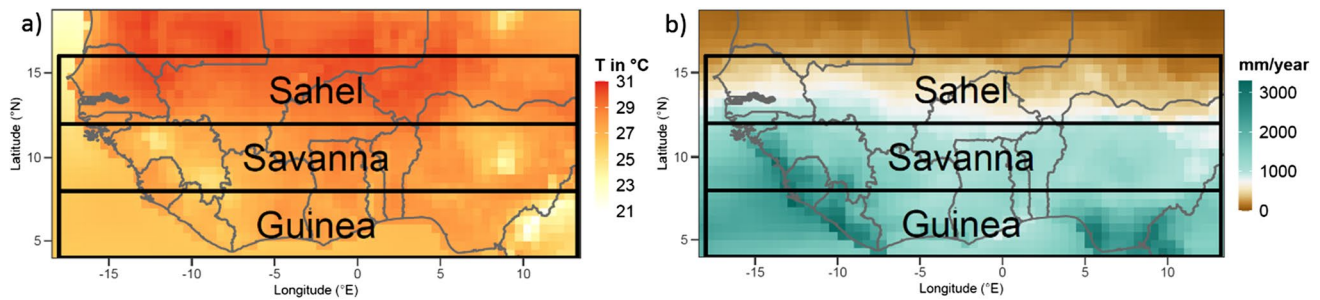


Fig. 1 **a** Mean annual temperature and **b** annual precipitation based on W5E5 data from 1985 to 2005. The three sub-regions are marked in black

February and October. Close to the equator, the rainy season is interrupted by a short dry season around July.

2.3 Methods

2.3.1 Representation of temperature and precipitation in the CMIP6 models

The bias-adjustment leading to the ISIMIP3b data hides the fact that the GCMs are only partially able to reproduce the main physical processes of the West African climate. Thus as a first step, we analyse the ability of the ten CMIP6 models to represent the main features of the West African climate by comparing the not bias-adjusted data from the CMIP6 models and their equally weighted multi-model ensemble mean (MMM) and the three different observational data sets CHIRTS, CHIRPS and ERA5 with the reference data set W5E5 ('Sect. 3.1'). The W5E5 data set is selected as a reference data set since it was used for the bias-adjustment of the ISIMIP3b data. Therefore, the comparison between the CMIP6 models and W5E5 shows the bias, which is addressed with the bias-adjustment. First, we compared the mean climatology from 1985 to 2014, including the spatial distribution of annual precipitation and mean annual temperature in the main monsoon season. We assessed the representation of the seasonality by comparing monthly temperature and precipitation along the latitudinal cross-section of the different data sets.

To compare the model simulations and different observational data sets against the W5E5 data set, we used different tools, which are further specified in the SI:

- Taylor diagrams for a quantitative comparison (Taylor 2001).
- Hovmöller diagrams to display the north–south distribution of temperature and precipitation across West Africa on a monthly scale (Hovmoeller 1949).
- Mann–Kendall test for trend detection from 1985 to 2014 (Kendall 1975; Mann 1945). The trend is in the

following named statistically significant if the significance level is below 5%. The trend analysis was done over the three regions Sahel, Savanna and Guinea.

2.3.2 Representation of agroclimatological indices in the ISIMIP3b data

To understand the representation of the agroclimate in the bias-adjusted ISIMIP3b data, we analysed the agreement between the bias-adjusted data of the ten models and the W5E5 data set for every grid point on land in the second part of the paper ('Sect. 3.2.1'). We validated the bias-adjusted data in how well it represents 14 temperature and precipitation-based agroclimatological indices which were selected due to their documented influence on the growth of main staple crops in West Africa or areas with similar climate and thus following an approach by Dieng et al. (2018). The comparison was done with the help of two statistics, the Pearson correlation coefficient (r) and the mean absolute error (MAE) as described in the supplemental information (SI), two complementary statistics that were also used in other model evaluation studies like Akinsanola et al. (2018). The agroclimatological indices were compared for each grid point.

The 14 agroclimatological indices are listed and defined in Table 3 and the following.

The onset of the rainy season, as an important agroclimatological index, was obtained using a definition adapted from Stern et al. (1981) and Laux et al. (2008). The rainy season onset is considered to be the first day of the year on which the following three conditions are simultaneously met:

- (1) At least 20 mm precipitation sum within 5 days.
- (2) The starting day and at least two other days in these 5 days are wet (defined as days with more than 0.5 mm precipitation).
- (3) No dry period of seven or more consecutive days within the next 30 days.

Table 3 Agroclimatological indices used for the analysis of the ISIMIP3b data

Agroclimatological indices	Definition	Unit
onset	First day of the year with at least 20 mm of precipitation within 5 days, of which the starting day and at least two other days are wet (<0.5 mm) and no dry period of 7 or more consecutive days within next 30 days	DoY
offset1	First day after the onset without a wet day (> 0.5 mm) within the next 20 days	DoY
offset2	Day after which 90% of precipitation has occurred	DoY
mm20	Number of days in the growing season with precipitation above 20 mm	mm
mm40	Number of days in the growing season with precipitation above 40 mm	mm
dry7	Number of times where 7 days are consecutively dry (<0.5 mm) within the growing season	Number of events
dry15	Number of times where 15 days are consecutively dry (<0.5 mm) within the growing season	Number of events
tmax95	95th percentile of maximum daily temperature in the growing season	°C
tmax99	99th percentile of maximum daily temperature in the growing season	°C
tmin05	5th percentile of minimum daily temperature in the growing season	°C
tmin01	1st percentile of minimum daily temperature in the growing season	°C
trange_sd	Standard deviation of diurnal temperature range in the growing season	°C
trange	Mean diurnal temperature range (maximum daily temperature–minimum daily temperature) in the growing season	°C
tmax_sd	Standard deviation of maximum daily temperature in the growing season	°C

We applied two definitions for the cessation date of the rainy season. In line with Laux et al. (2008), we defined cessation1 as the first day after the onset without a wet day (> 0.5 mm) within the next 20 days. The definition of cessation2 uses a threshold-based definition defining the end of the rainy season as the day after which 90% of the annual precipitation has fallen. This threshold is based on Odekunle (2006). Since this definition could not detect onset and cessation for all of West Africa (some missing points in the very dry north), all further agroclimatological indices are calculated for the growing season which is defined as dates between the earliest planting date and the latest harvesting date of maize (see ‘Sect. 2.1’).

To better understand the diverging performance of the models in different regions, we assessed the agreement with W5E5 additionally over the three regions the Sahel, Savanna and Guinea separately (‘Sect. 3.2.2’). Next to the multi-model ensemble mean (MMM), we assessed the multi-model mean of a sub-ensemble of the best-performing models (MMSM). This can give us a better idea of if a selection of best-performing models or exclusion of low-performing models can improve the general result (‘Sect. 3.2.3’).

3 Results and discussion

3.1 Comparing CMIP6 models with observational data sets

This section describes the agreement between different observational data sets as well as the agreement of the historical simulations of non-bias-adjusted CMIP6 models with the observations. This includes an analysis of annual precipitation and mean annual temperature, their monthly distribution, as well as past climate trends.

3.1.1 Representation of main climatological patterns in the observational data sets

To assess uncertainties in observational data and thereby set the base for evaluating the performance of the climate models, we first discussed the agreement between the three gridded observational data sets. Annual (Fig. 2) and monthly (Fig. 3) temperature values are higher in CHIRTS and lower in ERA5 than in W5E5 over most of West Africa. Cold temperature biases in ERA5 over Western Equatorial Africa were already found by Gleixner et al. (2020).

Annual (Fig. 4) and monthly (Fig. 5) precipitation is lower, most visibly over the Sahel, in ERA5 than in W5E5 and CHIRPS. CHIRPS agrees well with W5E5, showing only small regionally and monthly varying differences. W5E5 seems to have a strong wet bias over Gabun as both other observational data sets and all GCMs depict a much

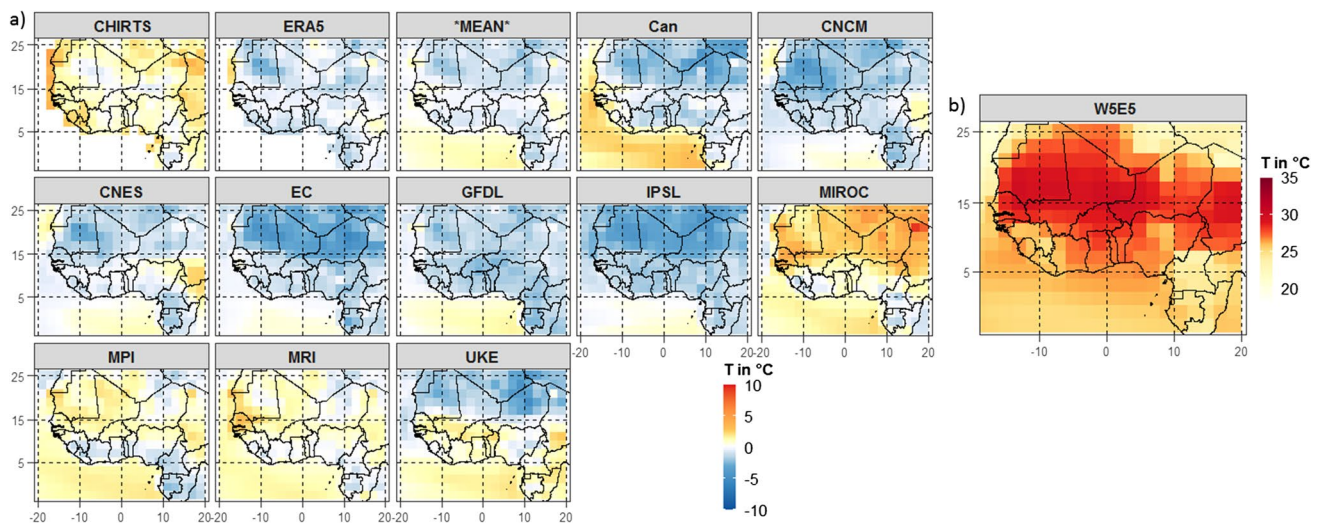


Fig. 2 **a** Difference in mean annual temperature between GCMs, CHIRTS and ERA5 compared to W5E5 for the period 1985–2014. **b** Mean annual temperature in °T averaged over 1985–2014 based on W5E5

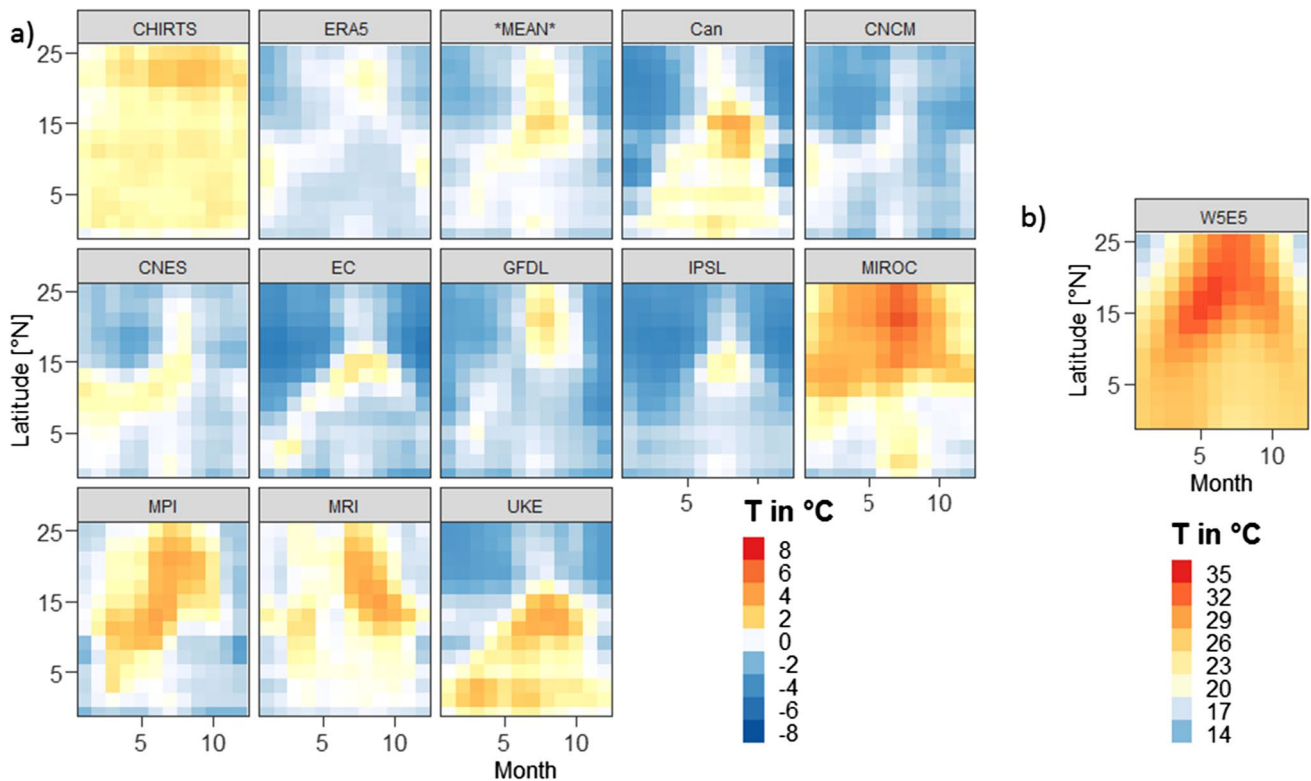


Fig. 3 Hovmöller diagrams of mean monthly temperature along the latitudinal cross-section between longitude -20°E and 20°W over land for the period 1985–2014. **a** Anomalies compared to W5E5. **b** Hovmöller diagram based on W5E5 as a reference

drier climate. Remarkable is that the cold temperature bias of ERA5 compared to CHIRTS and W5E5 shows patterns similar to the annual precipitation distribution (compare Figs. 5a and 3b) indicating that the reanalysis product overestimates the cooling effect of precipitation.

Additionally, the double passing of the rain belt and the movement of the WAM to its most northern peak are shown less clearly by ERA5. This is in line with a recent study that found biases in the position of the rain band in reanalysis products (Quagraine et al. 2020). Satagé

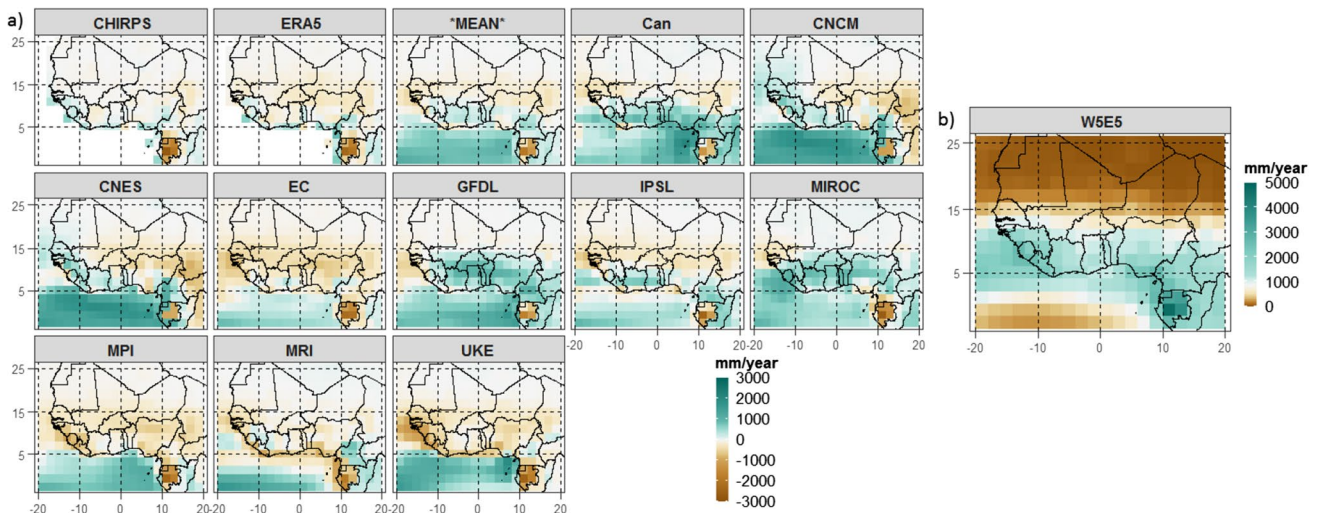


Fig. 4 a Difference in mean annual precipitation of GCMs, CHIRPS and ERA5 compared to W5E5 for the period 1985–2014. b Mean annual precipitation in mm averaged over 1985–2014 based on W5E5

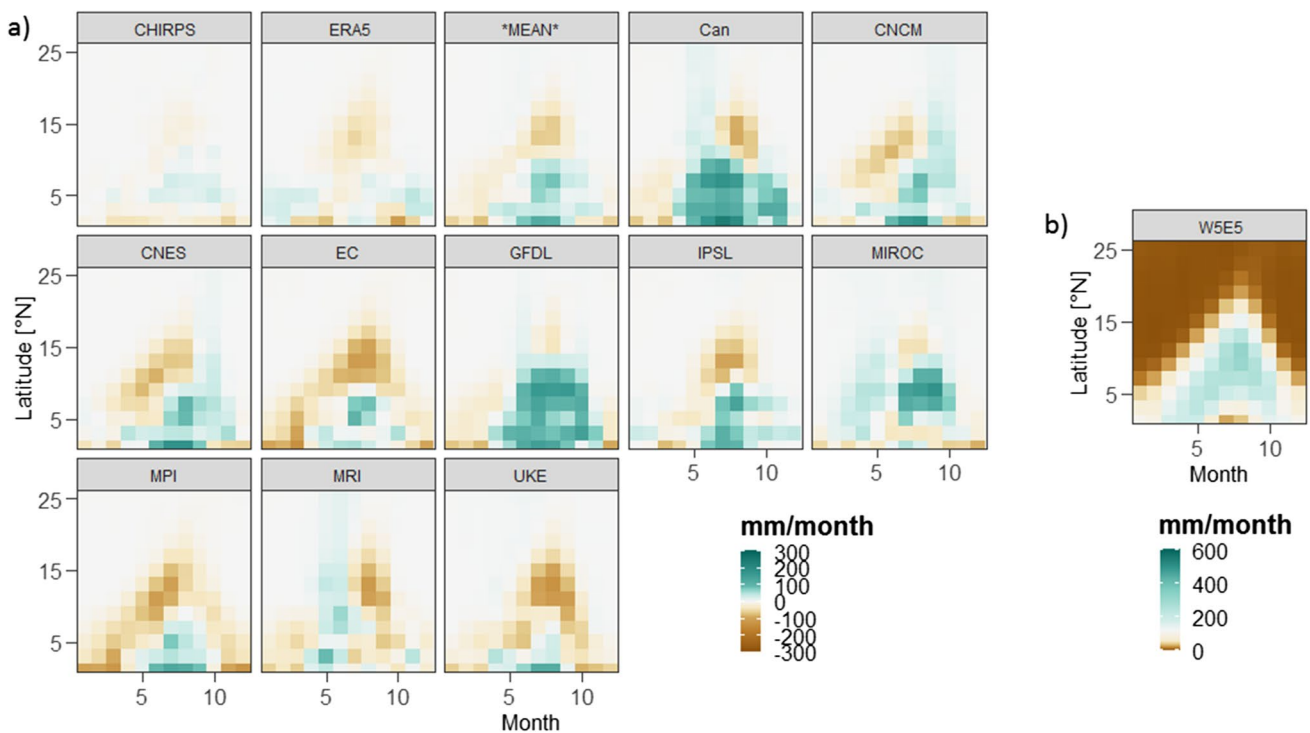


Fig. 5 Hovmöller diagrams of mean monthly precipitation along the latitudinal cross-section between longitude -20°E and 20°W over land for the period 1985–2014. a Anomalies compared to W5E5. b Hovmöller diagram based on W5E5 as a reference

et al. (2020) analysed 23 precipitation data sets over West Africa and found the data sets reliability to vary in space and time. They found CHIRPS to be the best-performing data set for monthly averages.

All in all, we found substantial differences in observational data sets. Since W5E5 and CHIRPS are partially dependent on ERA5, the observational uncertainties shown here might be even bigger than the displayed results. Sylla

et al. (2013) state that the large discrepancies in observational and reanalysis data sets over West Africa put constraints on a reliable model evaluation and Gómez-Navarro et al. (2012) point to the importance of a model evaluation that takes uncertainties in observations into account. The found differences in observations over West Africa show that the following validation of the performance of the models by a comparison to observations is largely dependent on and limited by the quality of the observational data set.

3.1.2 Representation of main climatological patterns in the CMIP6 models

Comparing the ten CMIP6 models and their ensemble mean to the observational data sets, the models exhibit some systematic and individual biases that are higher than the spread between the different observational data sets. Looking at the biases of each model in reproducing the mean annual temperature (Fig. 2), a systematic warm bias over the Gulf of Guinea is visible for all ten models and the MMM. Since the temperature gradient between the ocean and the Sahara is influencing the northward shift of the monsoon (Roehrig et al. 2013), questions arise about how far the models are reproducing the West African monsoon correctly. Over land, MIROC stands out with a strong warm bias over the Sahel and Sahara throughout the year, while Can, CNCM, CNES, EC, GFDL, IPSL and UKE show a clear cold bias over this region in the northern hemisphere winter (Fig. 3b). An analysis of the monthly distribution of temperature along the latitudinal cross-section shows that all ten GCMs can reproduce the temporal distribution of temperature with low inter-annual temperature differences in Guinea and higher differences in the Sahel and Sahara (Fig. 3). The models tend to show warm biases in the monsoon season in the Sahel and Sahara and

cold biases in the dry season. Similar to ERA5, some models show temperature biases that are similar to the annual precipitation distribution (compare Figs. 5a and 3b) indicating that also some models are not capable of simulating the cooling effect of precipitation well. This is especially visible for UKE and Can.

The simulations of precipitation in the CMIP6 models show a systematic southward shift of the rain belt, indicated by a wet bias over the Guinean coast and a dry bias over the Sahel in JJAS in most models and the MMM (Fig. 4). Only GFDL and MIROC overestimate the precipitation in the southern Sahel. This systematic underestimation of the northwards movement of the WAM could be linked with the systematic warm bias over the Gulf of Guinea in JJAS (Fig. 2) as Roehrig et al. (2013) have found a relation between the north–south temperature gradient between the Sahara and the ocean in the monsoon season and the position of the monsoon maximum in CMIP3 and CMIP5 historical simulations.

All models capture the annual movement of the WAM including the northward movement of the rain belt, reaching a maximum around July and migrating southwards afterwards (Fig. 5a) with some temporal and spatial biases (Fig. 5b). None of the models can detect the position and time of the most northern position of the WAM correctly. EC, MPI and MRI simulate a late onset of the rainy season. Due to the double passage of the tropical rain belt around the equator, two rainy seasons are shown by the observational data sets. Only the MRI model can simulate this pattern reasonably well, showing a distinct dry period in July near the Gulf of Guinea.

For temperature and precipitation, the MMM shows smaller absolute differences to W5E5 than most individual models. Some systematic biases persist in the MMM. Biases in precipitation are similar to biases found in earlier

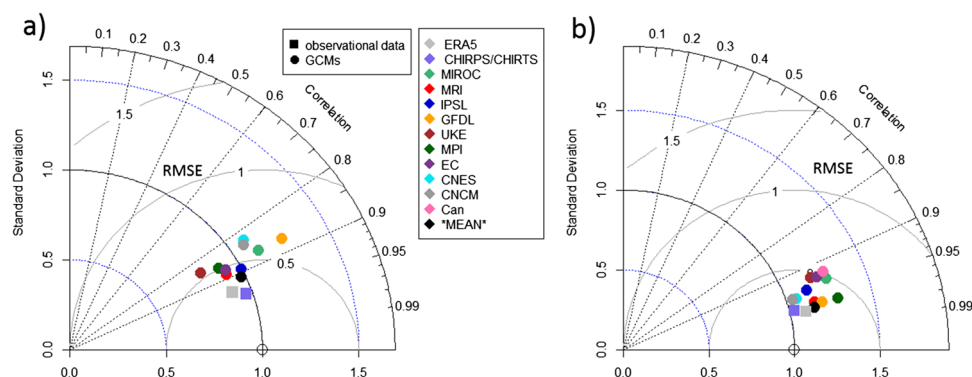


Fig. 6 Taylor diagram showing the agreement of **a** monthly precipitation and **b** monthly temperature over West Africa. Displayed are the ten CMIP6 data sets and the three observational data sets in reference to the W5E5 data set. The Taylor diagram displays the spatial (each grid point) combined with the temporal (12 months) agreement based

on three statistics: (1) the Pearson correlation coefficient (azimuthal angle), the root-mean-square error (distance to the centre point) and the standard deviation (radial distance from the origin, normalised to 1)

(downscaled) CMIP versions by the IPCC 4th Assessment Report (IPCC 2007) and Akinsanola and Zhou (2018).

When comparing the biases of the multi-model mean with the ones of ERA5, it becomes visible that the stronger systematic biases in the models (warm bias over the Sahel and Sahara and underestimation of the northwards movement of the WAM) are to a weaker degree also visible in ERA5.

Additionally, we analysed the representation of mean monthly temperature and precipitation quantitatively with Taylor diagrams (Fig. 6). The diagrams confirm that the three observational data sets agree slightly better with W5E5 than the GCMs. For mean temperature and precipitation over West Africa, the three observational data sets are correlated (> 0.97 for tas and > 0.93 for pr) and have low values of root-mean-square errors. The normalised standard deviations are around 1 (> 1 for tas and < 1 for pr). ERA5 and CHIRPS had an initially higher resolution than W5E5 and were regridded with conservative remapping. The remapping could partially explain the different standard deviations of ERA5 and CHIRPS compared to W5E5 (Diaconescu et al. 2015). Comparing the CMIP6 models to W5E5, temperature shows lower values for the root-mean-square error and a better correlation than precipitation with a correlation coefficient between 0.92 and 0.975 compared to a range between 0.82 and 0.91. The overall high correlation values are partially emerging from the heterogenic climate over space and time. The models tend towards high standard deviations for temperature, indicating a stronger spatial and/or temporal divergence in comparison to W5E5.

For precipitation, the ten models depict an equal spread with some models showing higher and some lower standard deviations. Only the MMM shows a high correlation and a normalised standard deviation close to one. The MMM is closer to W5E5 than all other models for mean annual temperature and precipitation being placed very close to the observational data sets in the Taylor diagrams.

3.1.3 Past trends in the observational data sets and CMIP6 models

We compared the temporal evolution by applying a Mann–Kendall trend test. A significantly increasing past temperature trend of similar magnitude is indicated by all observational data sets in all the regions (Fig. 7b). The weakest trend is observed in the Sahel. Also, the climate models represent the temperature trend well. Models with weak trends tend to display weak trends in all sub-regions and vice versa. Only MIROC does not detect a significant past trend in the Sahel and Savanna.

While for temperature the trends are almost uniform amongst the regions and agree well between the different data sets, the indicated trends for precipitation differ by region and data set (Fig. 7a). The observational data sets do not agree on a past trend in precipitation, with ERA5 not showing an increasing trend in the Sahel. According to other literature (Biasutti 2019; Kothe et al. 2014; Roehrig et al. 2013), seasonal precipitation amounts have partially

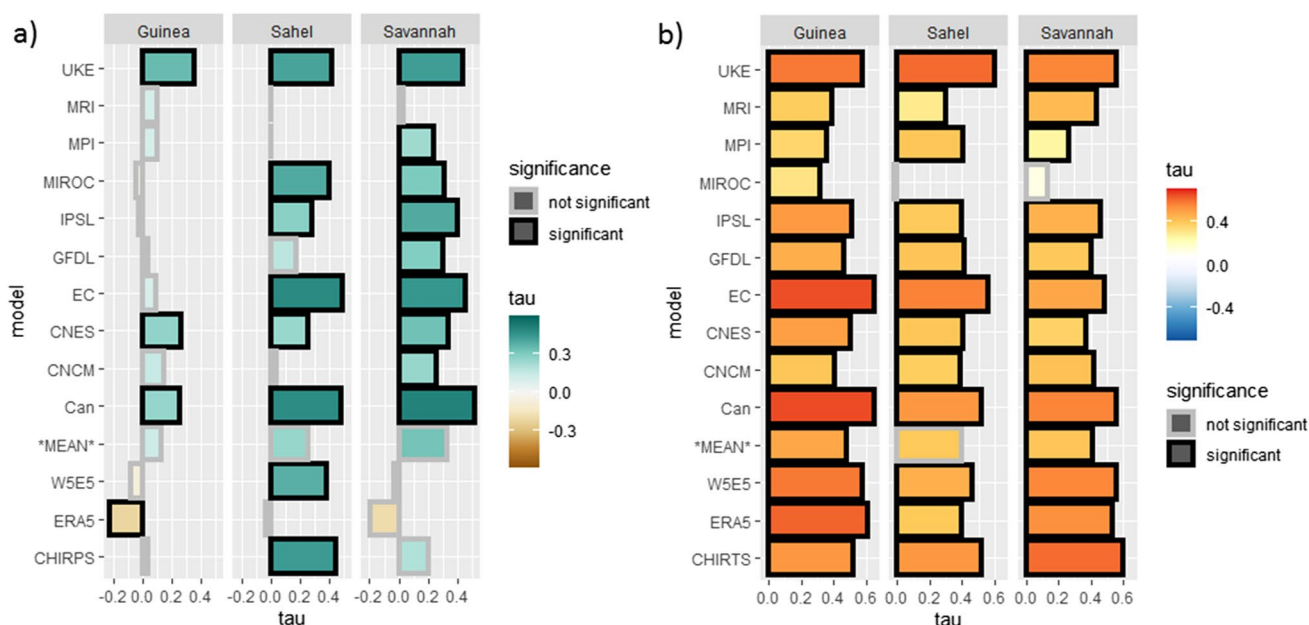


Fig. 7 Trend in **a** annual precipitation and **b** mean annual temperature between 1985 and 2014 for the ten GCMs and the three observational data sets. A Mann–Kendall test was applied. The magnitude of the trend and its direction is given by tau. The significance level is 5%

recovered since the droughts of the 1970s and 1980s, supporting the increasing trend shown by CHIRPS and W5E5 in the Sahel. Reanalysis data was found to be unsuitable for the investigation of trends (Bengtsson et al. 2004; Thorne & Vose 2010; Trenberth et al. 2008) and large differences in local trend patterns in West Africa have been detected amongst reanalysis products and in comparison with observations (Quagraine et al. 2020). Also, Gleixner et al. (2020) have found opposing trends in ERA5 temperature and precipitation trends on the African continent compared to other observational datasets. When excluding the ERA5 data set, the other two observational data sets agree on an increasing trend in the Sahel and detect partially opposing trends that are not significant in the other two regions. Despite excluding ERA5 for its known limitations in reproducing past trends, the differences in the other data sets point to uncertainties in past trends.

Six out of the ten models agree on the increasing trend in the Sahel. In disagreement with W5E5, nine out of ten models simulate a past increasing precipitation trend in the Savanna. The inability to correctly represent the precipitation trend in the Savanna might be due to the rain belt being placed too far south in most models.

3.2 Representation of agroclimatological indices in the bias-adjusted GCMs

3.2.1 Performance over West Africa

We have clearly shown that the CMIP6 data has strong biases and can simulate the climatology in West Africa, not to the extent that it is suitable for agricultural impact modelling. Thus, a bias-adjustment of the data is necessary before it can be adequately used for crop modelling as has been done in the ISIMIP project. The bias-adjustment adjusts the distribution of temperature and precipitation

and therefore also its long-term averages and variability of the climate simulations in CMIP6 to the W5E5 dataset. Taylor diagrams (Fig. 8) compare the agreement between CMIP6 data, bias-adjusted ISIMIP3b data and observational data sets with one another. Since the bias-adjusted data sets are on a different grid, the data displayed in one single Taylor diagram is not completely comparable. Nevertheless, the Taylor diagrams clearly show that the bias-adjustment done as part of the ISIMIP project leads to a spatial distribution of mean historical climate data that is very close to the W5E5 data set which was used for the bias-adjustment. This does not allow a clear conclusion on the agreement of ISIMIP3b data with the real climate as the observational uncertainty was shown to be high over the region. For example, as W5E5 might have a strong wet bias over Gabun, the bias-adjusted climate models inherit this bias. As the bias-adjusted data sets agree much better with one another than the observational data sets, we can see that the choice of the observational data set has a high influence on the bias-adjusted data as has also been highlighted by Famien et al. (2018) and Gampe et al. (2019).

While mean monthly values are almost perfectly aligned with the observations, it is not guaranteed that the bias-adjusted dataset captures other agriculturally relevant indicators and extreme events. Therefore, the focus of this section is on indicators that are important for ensuring reliable agricultural impact assessments (specification of the agroclimatological indices in ‘Sect. 2.3.2’).

Evaluating the bias-adjusted ISIMIP3b data shows that overall the agreement in agroclimatological indices between the ISIMIP3b data and the observational data set W5E5 is high (Fig. 9). There is not one bias-adjusted ISIMIP3b model that can simulate all aspects of the agroclimatology better than others. While *MIROC* (bias-adjusted data in *italic*) cessation dates agree best with W5E5, *MRI* can best simulate heavy precipitation events and *CNES* performs

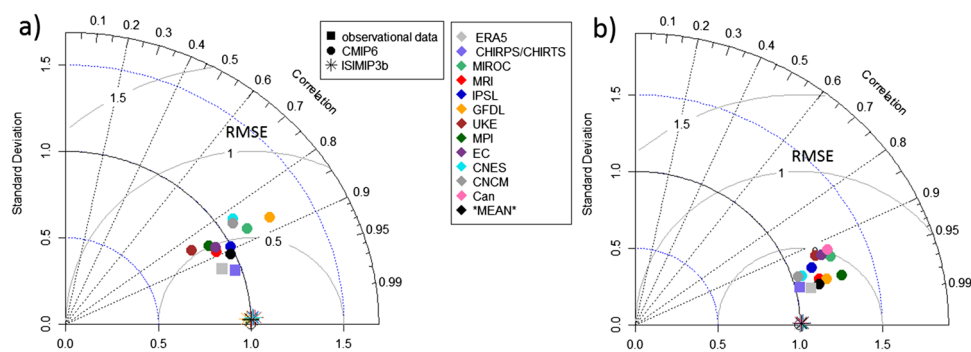


Fig. 8 Taylor diagram showing the agreement of **a**) monthly precipitation and **b**) monthly temperature over West Africa. Displayed are the CMIP6 data sets (circle), the 10 bias-adjusted ISIMIP3b data sets (star) and the three observational data sets (square) in reference to the W5E5 data set. The Taylor diagram displays the spatial (each grid point)

combined with the temporal (12 months) agreement based on three statistics: (1) the Pearson correlation coefficient (azimuthal angle), the root-mean-square error (distance to the centre point) and the standard deviation (radial distance from the origin, normalised to 1)

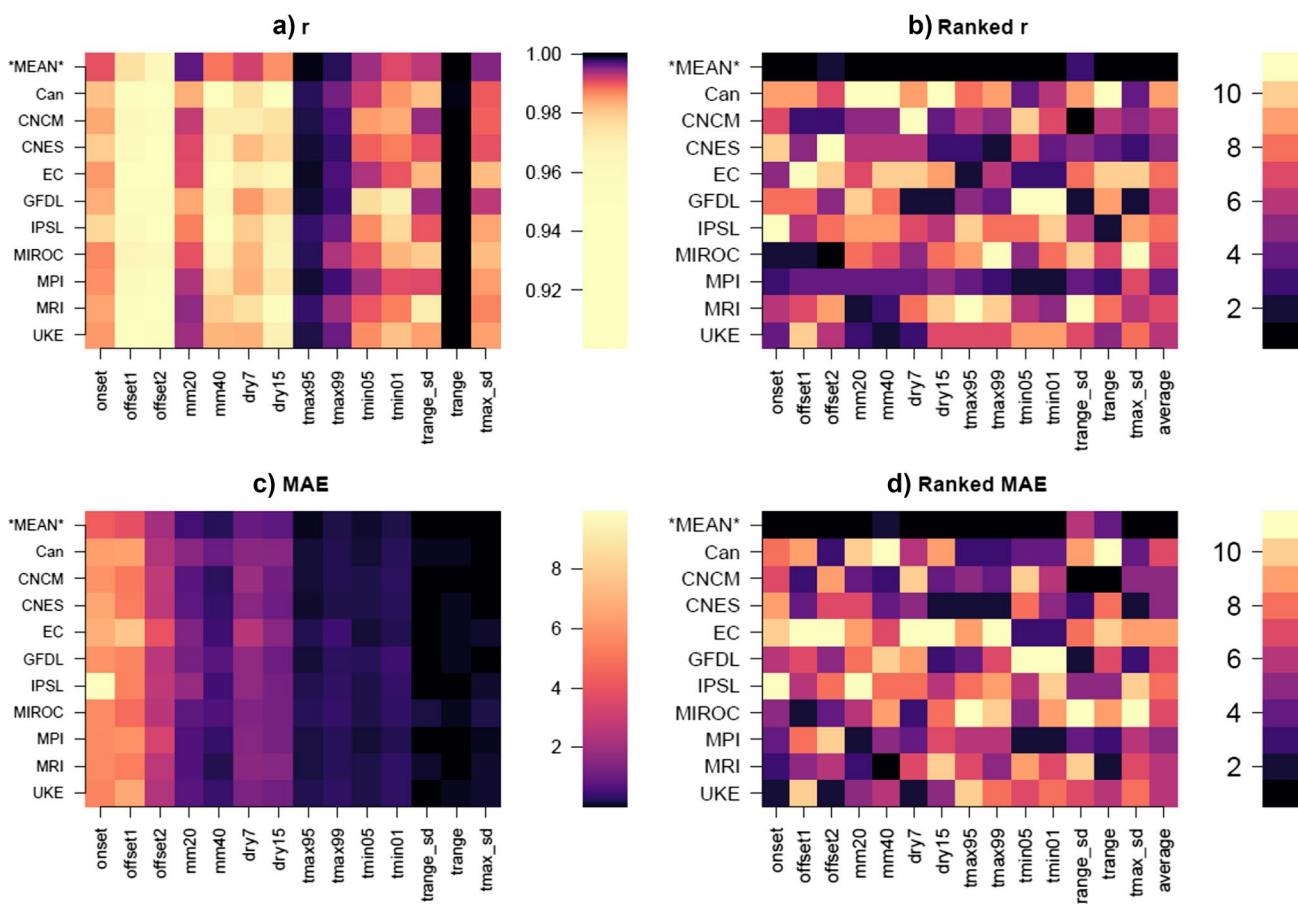


Fig. 9 Diagrams showing the Pearson correlation coefficient (a, b) and the mean absolute error (c, d) relating all models and the MMM to W5E5 for 14 different agricultural indicators. The right diagrams

display the ranking of the models in representing the individual indicators. The most right column displays the average rank of an individual model

best for high temperature extremes. The spatial distribution and direction of the bias for the different agroclimatological indices is shown in the SI Fig. 2–Fig. 15 and exemplary maps for 15 days dry spells are shown in Fig. 10. Most models show a bias towards more 15 days dry spells in the growing season in Guinea and less in the Sahel and Savanna compared to W5E5. Thus, this bias can also be seen to a weak extent in the ensemble mean. Other biases cancel each other out, making the MMM the closest to W5E5. Abrupt changes of values, especially at country borders, can be seen in all indicators that are calculated within the growing season. This is not due to the climate data sets but due to the definition of the growing season which is in this data set often defined for a whole country (Sacks et al. 2010) (SI Fig. 1). Since the performance analysis of the models is based on a comparison to W5E5, we assume that a more fine-grained definition of the growing season would not have a strong influence on the results here.

Overall, the agreement between ISIMIP3b data and W5E5 is dependent on the index, the model, the area and the chosen statistics. On average *MPI*, *CNCM* and *CNES*

have a better ranking for all three statistics. The MMM outperforms any single model in most agroclimatological indices under all applied statistics. While the agreement is generally good, outliers exist on a regional scale. Information on these outliers (as can be seen in the maps in Fig. 10 and in SI Fig. 2–SI Fig. 15) can help impact models understand the behaviour of the model when applying the different bias-adjusted GCMs. For example, *Can*, *IPSL* and *GFDL* underestimate the number of days with heavy precipitation of at least 20 mm and overestimate the number of days with very heavy precipitation of at least 40 mm. *EC* shows the opposite biases (SI Fig. 5 and SI Fig. 6). These biases in climate input data can influence crop model outputs, especially for crops and models that are sensitive to heavy precipitation. Some indices, namely onset and cessation of the rainy season and 15 days dry spells, show in individual regions strong negative or positive deviations from W5E5. Regional outliers in crop model outputs driven by individual climate models could be better understood and eliminated by taking these results into account.

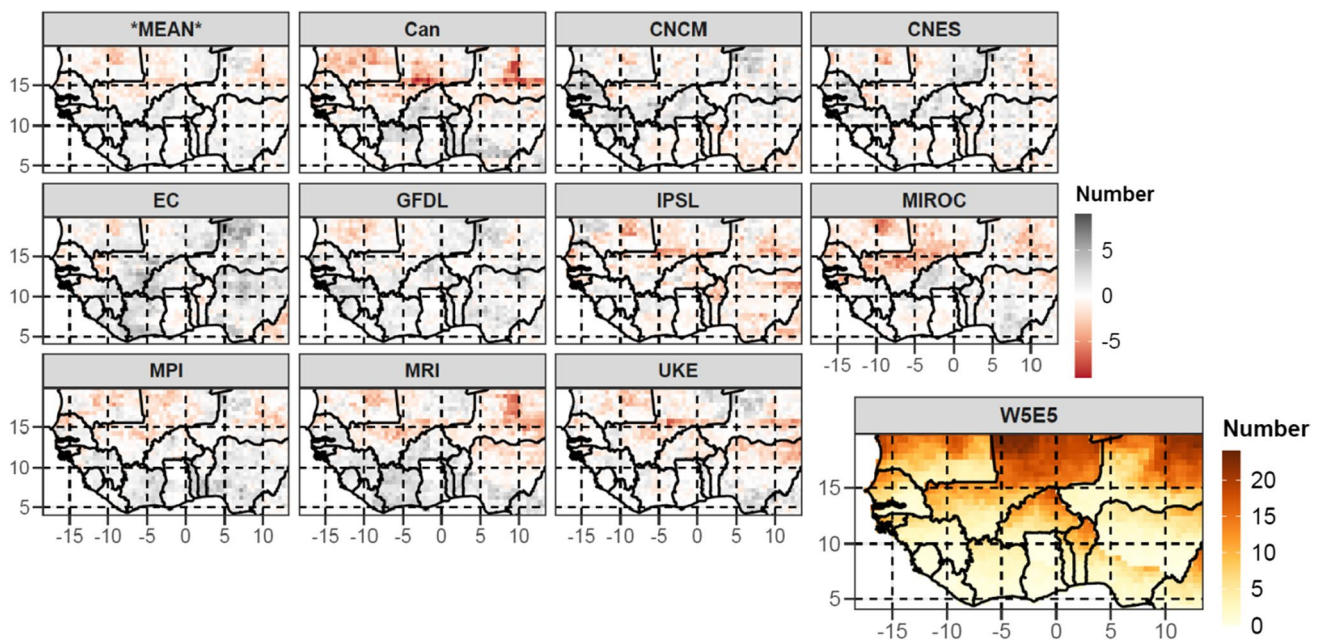
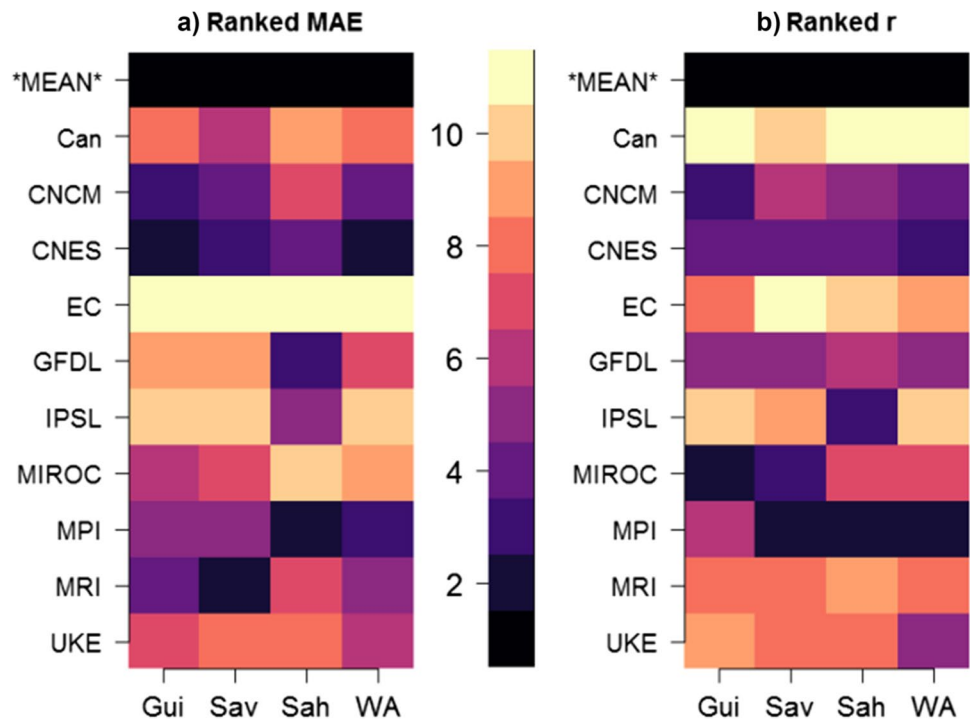


Fig. 10 Difference in number of 15 days dry spells within the growing season detected by the 10 bias-adjusted GCMs and the ensemble mean in reference to the W5E5. Right bottom: number of 15 days dry spells in the W5E5 data set

Fig. 11 Ranking of agreement between ISIMIP data and W5E5 for each sub-region Guinea, Savanna, Sahel and whole of West Africa based on the mean absolute error (a) and the Pearson correlation coefficient (b). Rankings are averages over the 14 agroclimatological indices



Although the bias-adjusted models agree to a satisfying degree with the W5E5 data, our ability to detect their representation of the true current climate is limited by the quality of the W5E5 data set. Our results point to uncertainties here since the three different observational data sets show clear differences ('Sect. 3.1.1'). We want to underline the high

influence of the selection of the reference data set applied in the bias-adjustment on the bias-adjusted data.

3.2.2 Model ranking over sub-regions

The models show a slightly different performance in reproducing the agricultural indices in the individual sub-regions as could also be seen on a fine scale already in Fig. 10 and SI Fig. 2–SI Fig. 15. To quantify regional differences in the performance of the individual models, Fig. 11 shows the ranked mean absolute error and Pearson correlation coefficient for each model and sub-region separately, whereby the ranking is averaged over all 14 agroclimatological indices. While the ranked correspondence to W5E5 is also for the regions dependent on the statistic, *CNCM* and *CNES* tend to perform better over Guinea while *MPI* performs better over the Sahel for both statistics. Averaged over all indices, the MMM still clearly outperforms any of the individual models in all regions.

3.2.3 Performance of sub-ensemble of models

Since *MPI*, *CNCM* and *CNES* show higher agreement to W5E5 for all analysed statistics and many agroclimatological indices, we analysed the performance of the mean of the three model sub-ensemble (here called multi-model sub-ensemble mean: MMSM). Averaged over the whole region, the MMM does still represent all but one agroclimatological index better than the MMSM (SI Table 1 shows the correlation coefficients for all agroclimatological indices between the MMSM/MMM and the W5E5 data set). Figure 12 shows the number of indices where the MMSM outperforms the MMM for each grid point. Overall, the regions where the MMM is closer to W5E5 than the MMSM dominate (displayed in pink in Fig. 12) indicating that the selection of the best-performing models for a multi-model mean is not meaningful over the whole of West Africa. Nevertheless, on a regional scale, we could show that the selection of models could improve the performance of the mean for a few

regions, for example northern Burkina Faso. Thus, in line with Schaller et al. (2011), we find that the selection of a sub-ensemble of climate models can be considered on a regional scale after a further assessment.

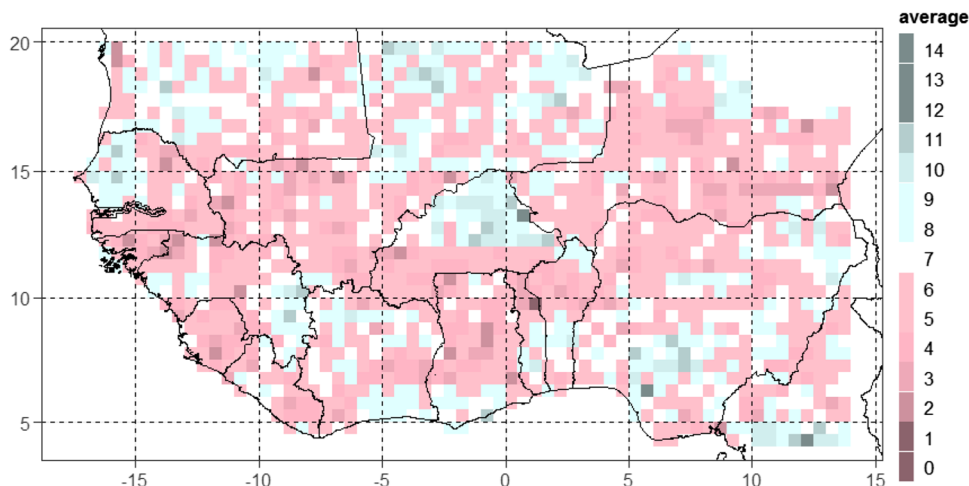
This simple selection of a sub-ensemble of models does not allow a conclusion for (not) applying any weighting or exclusion of models to reach a higher performance of the ensemble mean. More complex approaches that apply a weighting according to the model's uniqueness as well as an agreement with observations (e.g. Sanderson et al. 2017) and maintain the model spread (e.g. Herger et al. 2018) could lead to different results. A more detailed process-based analysis could also allow model exclusion on a regional scale, but the method requires substantial resources and in depth expertise on the regional climate dynamics (Siderius et al. 2021).

4 Summary and conclusions

To answer the need for high-quality climate data under increasing climate risks, uncertainties in past and future climate data need to be reduced and better understood. This study provides a better understanding of the quality and limitations of past and future climate data sets to be used for agricultural impact assessments in West Africa. For this, we compared ten CMIP6 GCMs and their respective bias-adjusted ISIMIP3b versions against the three different observational and reanalysis data sets CHIRPS, CHIRTS, ERA5 and W5E5 in terms of representing mean temperature and precipitation, past trends as well as crucial agroclimatological indices.

The observational and reanalysis data sets show regionally dependent differences in their representation of past trends and mean climatology over West Africa pointing at large observational uncertainties. The reanalysis data set ERA5 displays an overestimated cooling effect of precipitation and shows an opposing past precipitation trend in

Fig. 12 Number of indices for which the MMSM values are closer to the W5E5 data set than the MMM values. The blueish colour indicates that the mean of the best three models is closer to W5E5 for more than half of the indices, while pink shades indicate the opposite



comparison to the other data sets. W5E5 and thus also the bias-adjusted ISIMIP3b data show a wet bias over Gabon compared to the other data sets.

The differences between the GCMs exceed the differences between the observational data sets. While the models successfully simulate the large climatological patterns of annual and monthly temperature and precipitation, systematic as well as individual biases exist in all models. With similar systematic biases found in earlier CMIP versions by IPCC 4th Assessment Report (IPCC 2007) and Akinsanola and Zhou (2018), we conclude that CMIP6 could only partially improve the performance over West Africa. Overall, the multi-model ensemble mean presents the highest correlation with observations.

The bias-adjustment leading to ISIMIP3b data almost perfectly aligns mean precipitation and temperature to the reference data set W5E5. Thus, the differences between the observational data sets are much higher than between the bias-adjusted models. This highlights the importance of ensuring the quality of the reference data set used for the bias-adjustment and underpins the need to account for observational uncertainty when developing and working with bias-adjusted data. The W5E5 data set and the bias-adjusted ISIMIP3b data are overall also in good agreement for 14 selected agroclimatological indices like rainy season onset and cessation as well as extreme precipitation and temperature values. This lends confidence in the use of ISIMIP3b data for crop modelling over West Africa but this confidence is limited by the high observational uncertainty we found earlier and thus the limited trust we can put in the reference data set W5E5. Some regionally higher systematic and individual differences to W5E5 could be shown in the individual bias-adjusted models which can help to explain outliers in agricultural modelling results. Not a single model can simulate all agroclimatological indices better than the other models, but the performance depends on the indicator, the location and to some extent on the chosen statistics. The ensemble mean consistently outperforms the individual bias-adjusted models. A sub-ensemble of the slightly better performing models (*MPI*, *CNCM* and *CNES*) did not perform better than the full ensemble over the whole region but could improve the correlation to W5E5 in a few regions.

Therefore, the results of this study suggest using the complete ensemble mean of the ISIMIP3b data over West Africa or carefully selecting a subset of models based on an analysis on a regional scale. While the multi-model mean of the climate models shows the highest performance, it cannot directly be used as an input for impact models since it is not a physically consistent data set and suppresses part of the variability of the climate system. Instead, we recommend to run impact models with each individual model of the whole model ensemble and average at the end over the impact results. Nevertheless, averaging can only serve

as a best guess. Considering results from all individual models is essential to understand the range of uncertainties in climate impact studies. The detailed analysis found in this study can support agricultural modellers to understand their impact results under different climate input data. These recommendations are limited by the selection of indicators used for this assessment and the simple selection of best-performing models.

As the UN aims to increase readiness in the light of multi-causal risks in the Sahel region, the strategies need to root in solid analyses of the crop response to climate change and thus, high-quality data is key. Our findings point to further efforts needed in two main research areas to increase the reliability of climate data used for agricultural impact assessments and other cascading climate change impacts. First, observational uncertainties highlight the necessity to increase the quality of observed weather data by, e.g. expanding the station network over the study domain complemented with further improvements of reanalysis and satellite products. Second, model uncertainties and systematic model biases in the CMIP and ISIMIP3b data underpin the need to improve the representation of the complex West African climate in future CMIP versions.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00704-023-04430-3>.

Acknowledgements We highly appreciate the creation and provision of datasets by all institutions and spatial product creators. Namely, we would like to thank the modelling groups listed in Table 2 as well as the Center for Sustainability and the Global Environment, the Climatic Research Unit, the Copernicus Climate Change Service and the Climate Hazards Center. We are especially grateful to the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) team in producing, coordinating and making available the ISIMIP input data. For their valuable feedback and input, we would like to thank Bernhard Schaubberger, Rahel Laudien and Stefan Lange.

Author contribution All authors contributed to the study conception and design. The data analysis was performed by P.R.. C.G. initiated the research. The first draft of the manuscript was written by P.R. and all authors commented on and added to the following versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work has been funded by the Agrica, the CC&H PI3 and the Predictive Analytics project. Agrica is financed by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH on behalf of the German Federal Foreign Office and Federal Ministry for Economic Cooperation and Development (BMZ). CC&H PI3 is financed by the German Research Foundation (DFG). Predictive Analytics is financed by UNHCR to support the United Nations Integrated Strategy for the Sahel (UNISS) and the work of the UN Special Coordinator for Development in the Sahel.

Data availability The datasets analysed during the current study are available in the ISIMIP repository (<https://data.isimip.org/search/>), the Climate Hazards Center repository (<https://data.chc.ucsb.edu/products/>), the Copernicus repository (<https://cds.climate.copernicus.eu#!/home>) and the Center for Sustainability and the Global Environment

repository (<https://sage.nelson.wisc.edu/data-and-models/datasets/crop-calendar-dataset/>).

Code availability Code can be made available upon request.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abiodun BJ, Adeyewa ZD, Oguntunde PG, Salami AT, Ajayi VO (2012) Modeling the impacts of reforestation on future climate in West Africa. *Theoret Appl Climatol* 110(1–2):77–96. <https://doi.org/10.1007/s00704-012-0614-1>
- Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P-P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E (2003) The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J Hydrometeorol* 4(6):1147–1167. [https://doi.org/10.1175/1525-7541\(2003\)004%3c1147:TVGPCP%3e2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004%3c1147:TVGPCP%3e2.0.CO;2)
- Akinsanola AA, Ajayi VO, Adejare AT, Adeyeri OE, Gbode IE, Ogunjobi KO, Nikulin G, Abolude AT (2018) Evaluation of rainfall simulations over West Africa in dynamically downscaled CMIP5 global circulation models. *Theoret Appl Climatol* 132(1–2):437–450. <https://doi.org/10.1007/S00704-017-2087-8>
- Akinsanola AA, & Zhou W (2018) Projection of West African summer monsoon rainfall in dynamically downscaled CMIP5 models. *Climate Dynamics* <https://doi.org/10.1007/s00382-018-4568-6>
- Bengtsson L, Hagemann S, Hodges KI (2004) Can climate trends be calculated from reanalysis data? *J Geophys Res* 109:11111. <https://doi.org/10.1029/2004JD004536>
- Biasutti M (2019) Rainfall trends in the African Sahel: characteristics, processes, and causes. *Wiley Interdiscip Rev: Clim Chang* 10(4):e591. <https://doi.org/10.1002/WCC.591>
- Boucher O, Denvil S, Levvasseur G, Cozic A, Caubel A, Foujols MA, Meurdesoif Y, Cadule P, Devilliers M, Ghattas J, Lebas N, Lurton T, Mellul L, Musat I, Mignot J, & Cheruy F (2018) *IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.1534>
- Casanueva A, Herrera S, Iturbide M, Lange S, Jury M, Dosio A, Maraun D, & Gutiérrez JM (2020) Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmospheric Science Letters*, 21(7). <https://doi.org/10.1002/ASL.978/FORMAT/PDF>
- Cucchi M, Weedon GP, Amici A, Bellouin N, Lange S, Muller Schmied H, Hersbach H, & Buontempo C (2020) *WFDE5: bias adjusted ERA5 reanalysis data for impact studies*.
- Diaconescu EP, Gachon P, Laprise R (2015) On the remapping procedure of daily precipitation statistics and indices used in regional climate model evaluation. *J Hydrometeorol* 16(6):2301–2310. <https://doi.org/10.1175/JHM-D-15-0025.1>
- Dieng D, Laux P, Smiatek G, Heinzeller D, Bliedernicht J, Sarr A, Gaye AT, Kunstmann H (2018) Performance analysis and projected changes of agroclimatological indices across West Africa based on high-resolution regional climate model simulations. *J Geophys Res: Atmos* 123(15):7950–7973. <https://doi.org/10.1029/2018JD028536>
- Druyan, L. M. (2011). Studies of 21st-century precipitation trends over West Africa. In *Int J Climatol* <https://doi.org/10.1002/joc.2180>
- EC-Earth EEC (2019) *EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4700>
- Famien AM, Janicot S, Delfin Ochoa A, Vrac M, Defrance D, Sultan B, Noël T (2018) A bias-corrected CMIP5 dataset for Africa using the CDF-t method - a contribution to agricultural impact studies. *Earth System Dynamics* 9(1):313–338. <https://doi.org/10.5194/ESD-9-313-2018>
- Funk C, Peterson P, Peterson S, Shukla S, Davenport F, Michaelsen J, Knapp KR, Landsfeld M, Husak G, Harrison L, Rowland J, Budde M, Meiburg A, Dinku T, Pedreros D, Mata N (2019) A high-resolution 1983–2016 Tmax climate data record based on infrared temperatures and stations by the Climate Hazard Center. *J Clim* 32(17):5639–5658. <https://doi.org/10.1175/JCLI-D-18-0698.1>
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Rowland J, Romero B, Husak G, Michaelsen J, & Verdin P (2014) A quasi-global precipitation time series for drought monitoring. *U.S. Geological Survey Data Series*. <https://doi.org/10.3133/ds832>
- Gampe D, Schmid J, Ludwig R (2019) Impact of reference dataset selection on RCM evaluation, bias correction, and resulting climate change signals of precipitation. *J Hydrometeorol* 20(9):1813–1828. <https://doi.org/10.1175/JHM-D-18-0108.1>
- Gleixner S, Demissie T, Diro GT (2020) Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere* 11(9):996. <https://doi.org/10.3390/ATMOS11090996>
- Gómez-Navarro JJ, Montávez JP, Jerez S, Jiménez-Guerrero P, & Zorita E (2012) What is the role of the observational dataset in the evaluation and scoring of climate models? *Geophysical Research Letters*, 39(24). <https://doi.org/10.1029/2012GL054206>
- Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* 7(1):1–18. <https://doi.org/10.1038/s41597-020-0453-3>
- Herger N, Abramowitz G, Knutti R, Angéilil O, Lehmann K, Sander BM (2018) Selecting a climate model subset to optimise key ensemble properties. *Earth Syst Dyn* 9(1):135–151. <https://doi.org/10.5194/ESD-9-135-2018>
- Hersbach H, Bell W, Berrisford P, Horányi AJMS, Nicolas J., Radu R, Schepers D, Simmons A, Soci C, & Dee D (2019) *Global reanalysis: goodbye ERA-Interim, hello ERA5*. 159, 17–24. <https://doi.org/10.21957/vf291hehd7>
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M., ... Thépaut JN (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal*

- Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Herzschuh, U., Borkowski, J., Schewe, J., Mischke, S., & Tian, F. (2014). Moisture-advection feedback supports strong early-to-mid Holocene monsoon climate on the eastern Tibetan Plateau as inferred from a pollen-based reconstruction. *Palaeogeography, Palaeoclimatology, Palaeoecology* <https://doi.org/10.1016/j.palaeo.2014.02.022>
- Hovmoeller, E. (1949). The trough-and-ridge diagram. *Tellus* <https://doi.org/10.3402/tellusa.v1i2.8498>
- IPCC (2007) Climate change 2007: The physical science basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.) Contribution of working group I to the fourth assessment. Report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, and New York, p 996
- Jägermeyr J, Müller C, Ruane AC, Elliott J, Balkovic J, Castillo O, Faye B, Foster I, Folberth C, Franke JA, Fuchs K, Guarin JR, Heinke J, Hoogenboom G, Iizumi T, Jain AK, Kelly D, Khabarov N, Lange S, ... Rosenzweig C (2021) Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, 2(11), 873–885. <https://doi.org/10.1038/s43016-021-00400-y>
- James R, Washington R, Abiodun B, Kay G, Mutemi J, Pokam W, Hart N, Artan G, & Senior C (2018) Evaluating climate models with an African lens. *Bull Am Meteorol Soc.* <https://doi.org/10.1175/BAMS-D-16-0090.1>
- Kendall MG (1975) Rank correlation methods, 4th edn. Charles Griffin, London
- Kothe S, Lüthi D, & Ahrens B (2014) Analysis of the West African monsoon system in the regional climate model COSMO-CLM. *Int J Climatol* <https://doi.org/10.1002/joc.3702>
- Krasting JP, John JG, Blanton C, McHugh C, Nikonov S, Radhakrishnan A, Rand K, Zadeh NT, Balaji V, Durachta J, Dupuis C, Menzel R, Robinson T, Underwood S, Vahlenkamp H, Dunne KA, Gauthier PPG, Ginoux P, Griffies SM., ... Zhao M (2018) *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.1407>
- Lange S (2019) Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development Discussions*. <https://doi.org/10.5194/gmd-2019-36>
- Lange S (2021) *ISIMIP3b bias adjustment fact sheet*.
- Lange S, & Büchner M (2021) ISIMIP3b bias-adjusted atmospheric climate input data. *ISIMIP Repository*. <https://doi.org/10.48364/ISIMIP.842396.1>
- Lange S, Menz C, Gleixner S, Cucchi M, Weedon GP, Amici A, Bellouin N, Müller Schmied H, Hersbach H, Buontempo C, & Cagnazzo C (2021) *WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0)*. ISIMIP Repository. <https://doi.org/10.48364/ISIMIP.342217>
- Laux P, Kunstmann H, Bárdossy A (2008) Predicting the regional onset of the rainy season in West Africa. *Int J Climatol*. <https://doi.org/10.1002/joc.1542>
- Mann HB (1945) Nonparametric tests against trend. *Econometrica* 13(3):245. <https://doi.org/10.2307/1907187>
- Maraun D (2016) Bias correcting climate change simulations - a critical review. *Curr Clim Chang Rep* 2(4):211–220. <https://doi.org/10.1007/S40641-016-0050-X>
- Nijssen FJMM, Cox PM, Williamson MS (2020) Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models. *Earth Syst Dyn* 11(3):737–750. <https://doi.org/10.5194/esd-11-737-2020>
- Odekunle TO (2006) Determining rainy season onset and retreat over Nigeria from precipitation amount and number of rainy days. *Theor Appl Climatol* <https://doi.org/10.1007/s00704-005-0166-8>
- Omotsho JB, Abiodun BJ (2007) A numerical study of moisture build-up and rainfall over West Africa. *Meteorol Appl* 14(3):209–225. <https://doi.org/10.1002/met.11>
- Quagraine KA, Nkrumah F, Klein C, Klutse NAB, Quagraine KT (2020) West African summer monsoon precipitation variability as represented by reanalysis datasets. *Clim* 8(10):111. <https://doi.org/10.3390/CLI8100111>
- Richter I, Tokinaga H (2020) An overview of the performance of CMIP6 models in the tropical Atlantic: mean state, variability, and remote impacts. *Clim Dyn* 55(9–10):2579–2601. <https://doi.org/10.1007/S00382-020-05409-W/FIGURES/16>
- Roehrig R (2010) Variabilité intrasaisonnière de la mousson africaine : caractérisation et modélisation. Retrieved March 21, 2023, from <http://www.theses.fr/2010PEST1089>
- Roehrig R, Bouniol D, Guichard F, Hourdin F déric & Redelsperger JL (2013) The present and future of the West African monsoon: a process-oriented assessment of CMIP5 simulations along the AMMA transect. *J Clim* <https://doi.org/10.1175/JCLI-D-12-00505.1>
- Ruane AC, Phillips M, Müller C, Elliott J, Jägermeyr J, Arneth A, Balkovic J, Deryng D, Folberth C, Iizumi T, Izaurralde RC, Khabarov N, Lawrence, P, Liu W, Olin S, Pugh TAM, Rosenzweig C, Sakurai G, Schmid E., ... Yang H (2021) Strong regional influence of climatic forcing datasets on global crop model ensembles. *Agricultural and Forest Meteorology*, 300, 108313. <https://doi.org/10.1016/J.AGRFORMET.2020.108313>
- Sacks WJ, Deryng D, Foley JA, Ramankutty N (2010) Crop planting dates: an analysis of global patterns. *Glob Ecol Biogeogr* 19(5):607–620. <https://doi.org/10.1111/j.1466-8238.2010.00551.x>
- Sanderson BM, Wehner M, Knutti R (2017) Skill and independence weighting for multi-model assessments. *Geosci Model Dev* 10(6):2379–2395. <https://doi.org/10.5194/GMD-10-2379-2017>
- Satgé F, Defrance D, Sultan B, Bonnet MP, Seyler F, Rouché N, Pierron F, Patrel JE (2020) Evaluation of 23 gridded precipitation datasets across West Africa. *J Hydrol* 581:124412. <https://doi.org/10.1016/J.JHYDROL.2019.124412>
- Schaller N, Mahlstein I, Cermak J, Knutti R (2011) Analyzing precipitation projections: a comparison of different approaches to climate model evaluation. *J Geophys Res: Atmos* 116(D10):10118. <https://doi.org/10.1029/2010JD014963>
- Schewe J, & Levermann A (2017) Non-linear intensification of Sahel rainfall as a possible dynamic response to future warming. *Earth Syst Dyn* <https://doi.org/10.5194/esd-8-495-2017>
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, & Ziese M (2018) *GPCC full data monthly product version 2018 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historical data, Global Precipitation Climatology Centre*. https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_050
- Schulzweida U (2019) *CDO user guide* <https://doi.org/10.5281/zenodo.3539275>
- Siderius C, Geressu R, Todd MC, Kolusu SR, Harou JJ, Kashaigili JJ, & Conway D (2021) High stakes decisions under uncertainty: dams, development and climate change in the Rufiji River Basin. *Clim Risk in Africa*, 93–113. https://doi.org/10.1007/978-3-030-61160-6_6
- Sow M, Diakhaté M, Dixon RD, Guichard F, Dieng D, Gaye AT (2020) Uncertainties in the annual cycle of rainfall characteristics over West Africa in CMIP5 models. *Atmos* 11(2):216. <https://doi.org/10.3390/atmos11020216>
- Stern, R. D., Dennett, M. D., & Garbutt, D. J. (1981). The start of the rains in West Africa. *J Climatol* <https://doi.org/10.1002/joc.3370010107>

- von Storch JS, Putrasahan D, Lohmann K, Gutjahr O, Jungclaus J, Bittner M, Haak H, Wieners KH, Giorgetta M, Reick C, Esch M, Gayler V, de Vrese P, Raddatz T, Mauritsen T, Behrens J, Brovkin V, Claussen M, Crueger, T., ... Roeckner E (2017) *MPI-M MPIESM1.2-HR model output prepared for CMIP6 HighResMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.762>
- Sultan B, Gaetani M (2016) Agriculture in West Africa in the twenty-first century: climate change and impacts scenarios, and potential for adaptation. *Front Plant Sci* 7(AUG2016):1262. <https://doi.org/10.3389/FPLS.2016.01262/BIBTEX>
- Swart NC, Cole JNS, Kharin VV, Lazare M, Scinocca JF, Gillett NP, Anstey J, Arora V, Christian JR, Hanna S, Jiao Y, Lee WG, Majaess F, Saenko OA, Seiler C, Seinen C, Shao A, Sigmund M, Solheim L, ... Winter B (2019) The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, 12(11), 4823–4873. <https://doi.org/10.5194/gmd-12-4823-2019>
- Sylla MB, Giorgi F, Coppola E, & Mariotti L (2013) Uncertainties in daily rainfall over Africa: assessment of gridded observation products and evaluation of a regional climate model simulation. *Int J Climatol* <https://doi.org/10.1002/joc.3551>
- Tang Y, Rumbold S, Ellis R, Kelley D, Mulcahy J, Sellar A, Walton J, & Jones C (2019) *MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.6113>
- Tatebe H, & Watanabe M (2018) *MIROC MIROC6 model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.881>
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res Atmos* <https://doi.org/10.1029/2000JD900719>
- Thorne PW, Vose RS (2010) Reanalyses suitable for characterizing long-term trends. *Bull Am Meteor Soc* 91(3):353–362. <https://doi.org/10.1175/2009BAMS2858.1>
- Trenberth KE, Koike T, Onogi K (2008) Progress and prospects for reanalysis for weather and climate. *Eos* 89(26):234–235. <https://doi.org/10.1029/2008EO260002>
- UN (2018) *UN support plan for the Sahel working together for a prosperous and peaceful Sahel 2*. https://www.un.org/africarenewal/sites/www.un.org.africarenewal/files/EnglishSummary_Report_0.pdf
- UN (2021) UNISS United Nations Integrated Strategy for the Sahel - progress report 2019/20. *Office of the United Nations Special Coordinator for Development in the Sahel*. <https://unis-sahel.org/wp-content/uploads/2021/06/UNISS-Progress-Report-English-.pdf>
- United Nations (2019) *World population prospects 2019*. <https://population.un.org/wpp/DataQuery/>
- Verdin A, Funk C, Peterson P, Landsfeld M, Tuholske C, Grace K (2020) Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set. *Sci Data* 7(1):1–14. <https://doi.org/10.1038/s41597-020-00643-7>
- Voldoire A (2018) *CMIP6 simulations of the CNRM-CERFACS based on CNRM-CM6-1 model for CMIP experiment abrupt-4xCO2*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.3916>
- Voldoire A (2019) *CMIP6 ScenarioMIP ssp245*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4191>
- Vy DNT (2021) *Evaluation and comparison of satellite-based rainfall product CHIRPS and reanalysis product ERA5 in West Africa*. <https://doi.org/10.13140/RG.2.2.22619.49445>
- Wang B, Biasutti M, Byrne MP, Castro C, Chang CP, Cook K, Fu R, Grimm AM, Ha KJ, Hendon H, Kitoh A, Krishnan R, Lee JY, Li J, Liu J, Moise A, Pascale S, Roxy MK, Seth A., ... Zhou T (2021) Monsoons climate change assessment. *Bulletin of the American Meteorological Society*, 102(1), E1–E19. <https://doi.org/10.1175/BAMS-D-19-0335.1>
- Weedon GP, Balsamo G, Bellouin N, Gomes S, Best MJ, & Viterbo P (2014) The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-Interim reanalysis data. *Water Resour Res* <https://doi.org/10.1002/2014WR015638>
- Yukimoto S, Koshiro T, Kawai H, Oshima N, Yoshida K, Urakawa S, Tsujino H, Deushi M, Tanaka T, Hosaka M, Yoshimura H, Shindo E, Mizuta R, Ishii M, Obata A, & Adachi Y (2019) *MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.6842>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.