Climate Dynamics (2024) 62:1077–1106
https://doi.org/10.1007/s00382-023-06956-8

ORIGINAL ARTICLE

Robustness of climate indices relevant for agriculture in Africa deduced from GCMs and RCMs against reanalysis and gridded observations

Daniel Abel[1] · Katrin Ziegler[1] · Imoleayo Ezekiel Gbode[2] · Torsten Weber[3] · Vincent O. Ajayi[2] · Seydou B. Traoré[4] · Heiko Paeth[1]

Received: 9 January 2023 / Accepted: 7 September 2023 / Published online: 27 September 2023

Abstract

This study assesses the ability of climate models to represent rainy season (RS) dependent climate indices relevant for agriculture and crop-specific agricultural indices in eleven African subregions. For this, we analyze model ensembles build from Regional Climate Models (RCMs) from CORDEX-CORE (RCM_hist) and their respective driving General Circulation Models (GCMs) from CMIP5 (GCM_hist). Those are compared with gridded reference data including reanalyses at high spatio-temporal resolution ($\leq 0.25°$, daily) over the climatological period 1981–2010. Furthermore, the ensemble of RCM-evaluation runs forced by ERA-Interim (RCM_eval) is considered. Beside precipitation indices like the precipitation sum or number of rainy days annually and during the RS, we examine three agricultural indices (crop water need (CWN), irrigation requirement, water availability), depending on the RS' onset. The agricultural-relevant indices as simulated by climate models, including CORDEX-CORE, are assessed for the first time over several African subregions. All model ensembles simulate the general precipitation characteristics well. However, their performance strongly depends on the subregion. We show that the models can represent the RS in subregions with one RS adequately yet struggle in reproducing characteristics of two RSs. Precipitation indices based on the RS also show variable errors among the models and subregions. The representation of CWN is affected by the model family (GCM, RCM) and the forcing data (GCM, ERA-Interim). Nevertheless, the too coarse resolution of the GCMs hinders the representation of such specific indices as they are not able to consider land surface features and related processes of smaller scale. Additionally, the daily scale and the usage of complex variables (e.g., surface latent heat flux for CWN) and related preconditions (e.g., RS-onset and its spatial representation) add uncertainty to the index calculation. Mostly, the RCMs show a higher skill in representing the indices and add value to their forcing models.

Keywords Climate indices · Rainy season · Agriculture · Africa · CORDEX-CORE · CMIP5

1 Introduction

Agricultural water supply and use in Africa is mainly rain-fed (IPCC 2022). Thus, the continent's food security is highly vulnerable to events like droughts (e.g., Meza et al. 2020; Lottering et al. 2021) and heatwaves (e.g., Teixeira et al. 2013; Shew et al. 2020). This vulnerability is likely to increase due to climate change (IPCC 2022). This already manifests in more frequent extreme events over the last decades (Masih et al. 2014; Thomas and Nigam 2018). A general drying trend caused by higher precipitation uncertainty, higher evapotranspiration due to warmer temperatures, and increasing drought and heatwave risk are observed in recent decades and are likely to intensify in the future over Africa (e.g., Kotir 2011; Maidment et al. 2015; Dosio 2017; Weber

Daniel Abel and Katrin Ziegler have contributed equally to this work.

✉ Daniel Abel
daniel.abel@uni-wuerzburg.de

[1] Institute of Geography and Geology, University of Wuerzburg, Am Hubland, 97074 Wuerzburg, Germany

[2] Department of Meteorology and Climate Science, Federal University of Technology Akure (FUTA), Akure 340110, Ondo, Nigeria

[3] Climate Service Center Germany (GERICS), Helmholtz-Zentrum Hereon, Fischertwiete 1, 20095 Hamburg, Germany

[4] AGRHYMET Regional Centre, 425 Boulevard de l'Université, BP 11011, Niamey, Niger

et al. 2018; Ahmadalipour et al. 2019; Dosio et al. 2021a; IPCC 2021). These findings were also made for several African subregions: North (Elkouk et al. 2021; Zittis et al. 2021), East (Haile et al. 2020; Coppola et al. 2021), Southern (Abiodun et al. 2019; Mbokodo et al. 2020), Central (Fotso-Nguemo et al. 2022; Karam et al. 2022), and West Africa (Sylla et al. 2016; Sambou et al. 2021). Africa's vulnerability is further amplified by the rapidly growing population and, hence, increased food demand (Cleland 2013; Hall et al. 2017).

Climate change can have significant impacts on food security (Beltran-Peña and D'Odorico 2022), which declined over the past decades (Zhang et al. 2023) as it was also the case for crop yields by 5 to 20% recently (Sultan et al. 2019). Until 2050, the crop yields are expected to decrease further by 11% in West and 8% over entire Africa (Roudier et al. 2011; Knox et al. 2012). Not all crops and areas are affected to the same extent – some even may profit (Waha et al. 2013; Awoye et al. 2017; van Oort and Zwart 2018) – but e.g., for maize a general yield reduction has to be assumed over Africa and in several subregions (Waha et al. 2013; van Oort and Zwart 2018).

To investigate future climate changes under different relative concentration pathways (RCPs, van Vuuren et al. 2011), the above mentioned studies used general circulation (GCMs) and regional climate models (RCMs). Although these models are widely used to assess the impact of climate change on agriculture, their ability to reproduce the general circulation and precipitation characteristics over Africa is limited (e.g., Zebaze et al. 2019; Di Luca et al. 2020; Ayugi et al. 2020; Du et al. 2022) – as it is the case for the characteristics represented by different ETCCDI indices (Expert Team on Climate Change Detection and Indices, Zhang et al. 2011) (e.g., Sillmann et al. 2013; Ongoma et al. 2019; Sow et al. 2020; Dosio et al. 2021a; Ayugi et al. 2021). Additionally, model output evaluation is complicated due to the scarcity and/or unavailability of ground-based observation data, as well as the large difference among existing gridded precipitation products (e.g., Akinsanola and Ogunjobi 2017; Dembélé et al. 2020; Satgé et al. 2020; Dosio et al. 2021b). Nonetheless, it is important to know the model's uncertainties of the past to make valid statements on the future development of the climate.

To better understand the future risks and increasing vulnerability of Africa's rainfed agriculture, it is crucial to assess the quality of climate models in representing indices relevant for agriculture. Here, the onset of the rainy season plays an important role as it is the main factor in defining planting days for many crops (Akinseye et al. 2016; Dieng et al. 2018) and relevant for the definition of the agricultural indices used in this study. Furthermore, the amount and timing of rainfall during the rainy season as well as agricultural water needs are fundamental for the yield. Currently, crop-specific indices like crop water need (CWN), irrigation requirement (IR), or water availability (WA) (Allen et al. 1998) are used to estimate the water needs and are of major relevance for agriculture. These needs increased over the past decades (Rolle et al. 2022) and will enhance further under future climate change conditions (Fant et al. 2015; Jones et al. 2015; Dieng et al. 2018; Sylla et al. 2018). Most studies dealing with the crop-specific indices examined in this work used climate model data to force crop or hydrological models (e.g., Oettli et al. 2011; Konzmann et al. 2013; Waongo et al. 2015; Bonetti et al. 2022). Additionally, although the index names are the same, there are studies not using the FAO definition of the selected indices (Konzmann et al. 2013; Dieng et al. 2018; Rolle et al. 2021). Thus, few studies remain assessing the FAO-indices based directly on climate model data and focusing on a regional or larger scale (e.g., Gbode et al. 2022; Incoom et al. 2022) and not on small areas (e.g., Gurara et al. 2021). However, Dieng et al. (2018) focused on the performance of a single RCM in reproducing WA over West Africa, whereas WA followed a different definition. Gbode et al. (2022) considered all indices selected in our study for ensembles of GCMs and RCMs over West Africa but ignored the onset of the rainy season and different crop stages. Incoom et al. (2022) examined an ensemble of RCMs regarding a large area in Ghana regarding CWN and IR considering crop stages. Thus, information on the ability of climate models in reproducing agricultural indices over Africa is still rare. Moreover, current information on crop-specific indices is incomplete and fragmented over Africa (Rolle et al. 2022).

Consequently, this study aims to assess the ability of climate models simulating climate indices relevant to agriculture against gridded reference data for a historical period over Africa. The indices evaluated focus specifically on precipitation over the year and during the rainy seasons. In addition, crop-dependent indices which are determined by the onset of the rainy season are investigated. The novelty of our study consists of the assessment of CORDEX-CORE (Coordinated Regional Climate Downscaling Experiment – Coordinated Output for Regional Evaluation, Giorgi et al. 2022), the latest CORDEX generation, regarding the simulation of agricultural-relevant and crop-specific indices for several African subregions over a climatological period. The index selection underlying our study is demand-driven as the indices were requested by end-users (Weber et al. 2023b).

This study is organized as follows: Sect. 2 introduces the study area, the used data and related processing, calculated indices, and validation metrics. Section 3 starts with a comparison of gridded precipitation data and continues with an assessment of the ability of climate models to simulate precipitation. Afterwards, the rainy season and

related indices are compared and crop-specific indices are analyzed. Section 4 discusses the obtained results before drawing conclusions.

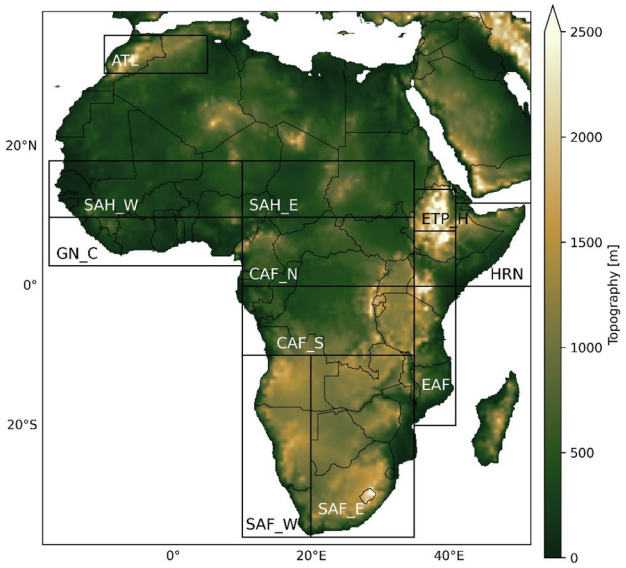

Fig. 1 Topography of Africa and overview of the investigated subregions based on Dosio et al. (2021a)

2 Data, indices, and methods

As the climate and precipitation patterns differ widely over the African continent, we examined several subregions shown in Fig. 1. Those are selected following Dosio et al. (2021a) as this definition is more differentiated than those used in IPCC's AR6 (Iturbide et al. 2020; IPCC 2021). The numerical longitudinal and latitudinal boundaries of the subregions are provided in Supplementary 1.

2.1 Climate data

2.1.1 Reference data

The high spatial resolution of CORDEX-CORE (0.22° × 0.22°) and some further reanalysis-driven RCM-simulations requires a restriction of the reference data to resolutions of 0.25° × 0.25° or higher. This is necessary as too coarse data are a significant source of uncertainty when it comes to the evaluation of models having a significantly finer resolution (Casanueva et al. 2020; Ciarlo et al. 2021). In addition, some of the investigated indices require daily data which is a further criterion. To create a recent climatology of the observed and modeled data, the common period 1981–2010 is considered. We have been less strict regarding the start year and included datasets beginning in 1983 as well (Table 1). A more detailed description of the individual

Table 1 Overview of datasets used in this study. A more detailed description of the datasets is given in Supplementary 2. A more comprehensive overview of precipitation and temperature datasets available for Africa showing the covered period, spatial resolution, and data source is provided in Supplementary 3

Dataset	Long name	Covered Period	Spatial resolution	Data source	Reference
AGERA5	Agrometeorological ECMWF ReAnalysis 5	1979-present	0.1°	Reanalysis	Boogaard et al. (2022)
ARC2	African Precipitation Climatology version 2	1983-present	0.1°	Merged product (station + satellite data)	Novella and Thiaw (2013)
CHIRPS	Climate Hazards Group InfraRed Precipitation with Stations	1981-present	0.05°	Merged product (station + satellite data)	Funk et al. (2015)
ERA5	ECMWF ReAnalysis 5	1950-present	0.25°	Reanalysis	Hersbach et al. (2020)
ERA5Land	ECMWF ReAnalysis 5 Land	1950-present	0.1°	Reanalysis	Muñoz-Sabater et al. (2021)
GPCC (monthly)	Global Precipitation Climatology Centre	1981-present	0.25°	Station	Schneider et al. (2020)
PERSIANN-CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Climate Data Record	1983-present	0.25°	Satellite	Ashouri et al. (2015)
TAMSAT	Tropical Applications of Meteorology using SATellite	1983-present	0.0375°	Merged product (station + satellite data)	Maidment et al. (2017)

datasets is given in Supplementary 2. Further, a comprehensive analysis of existing gridded precipitation datasets over Africa has been completed to obtain an overview of applicable observational datasets (Supplementary 3).

In contrast to ARC-2, TAMSAT, GPCC, PERSIANN-CDR, and CHIRPS, the reanalyses data from ERA5 and its child products ERA5Land and AGERA5 also provide 2 m-temperature data (TMIN, TMAX, TMEAN) that are required for the computation of some indices. Other satellite or station-based temperature datasets fulfilling the necessary temporal and spatial criteria for our study are not available for the reference period (Supplementary 3). Additionally, reanalysis data are crucial when it comes to the consideration of further variables, which are more complex to measure, for instance the latent heat flux.

There are some studies that already compare some of the aforementioned and other observational datasets with a focus on subregions like West Africa (e.g., Akinsanola and Ogunjobi 2017; Dembélé et al. 2020; Satgé et al. 2020), Ethiopia (Degefu et al. 2022), or on entire Africa (e.g., Dosio et al. 2021b). However, these comparative studies demonstrate that the quality of the observational datasets differs with study area, temporal resolution, and underlying research question (e.g., flood monitoring, drought, extreme events, total amount, indices etc.). Therefore, we also undertake a thorough comparative analysis of the selected datasets for Africa and its subregions.

2.1.2 Model data

In Table 2, selected models from CMIP5 (Climate Model Intercomparison Project 5, Taylor et al. 2012) examined in this study are listed. The selection is based on the models used to force the available CORDEX-CORE simulations (Giorgi et al. 2022) shown also in Table 2. CMIP5 has been widely used to simulate Africa's climate over historical and future periods. The ensemble or parts of it have undergone a manifold evaluation on a global or subregional African scale for temperature and precipitation characteristics (e.g., Zebaze et al. 2019; Di Luca et al. 2020; Ayugi et al. 2020; Du et al. 2022) as well as ETCCDI indices (e.g., Sillmann et al. 2013; Ongoma et al. 2019; Sow et al. 2020; Ayugi et al. 2021).

In general, the GCMs' resolution is too coarse to account for important processes and simulate land surface heterogeneities adequately, which is also true for Africa (Dosio et al. 2019). Thus, we also consider RCMs from the recently published CORDEX-CORE ensemble (Table 2). This data is used instead of the previous and established CORDEX-AFR ensemble due to its higher spatial resolution (0.22°

Table 2 Overview of available RCMs from CORDEX-CORE for Africa in 0.22° resolution and the respective forcing data from ERA-Interim for evaluation runs (red) and from GCMs from CMIP5 for historical runs (purple). The historical ensemble of GCMs used is presented in green. *HadGEM2-ES and RCMs forced by this are not considered due to the assumption of 30 days each month (see Sect. 2.1.3)

	GCM/Forcing				
	Reference for GCM	Bentsen et al. (2012, 2013)	Martin et al. (2011), Jones et al. (2014)	Giorgetta et al. (2013)	
	spatial resolution for GCM	2.5° x 1.894737°	1.85° x 1.25°	1.875°	
	global pixel number for GCM	144x96	192x144	192x96	
Model name	ERA-Interim	NCC: NorESM1-M	MOHC: HadGEM2-ES*	MPI-M: MPI-ESM-LR	MPI-M: MPI-ESM-MR
GERICS-REMO2015_v1	1979-2017	1970-2099	1970-2099*	1970-2100	
CLMcom-KIT-CCLM5-0-15_v1	1979-2016	1950-2100	1950-2098*	1950-2100	
ICTP-RegCM4-7_v0	1979-2010	1970-2099	1970-2099*		1970-2099

instead of 0.44°) and the updated development stage of the RCMs. Dosio et al. (2021a) investigated some daily precipitation characteristics and showed that CORDEX-CORE represents these indices better than CORDEX-AFR in the majority of the subregions and seasons compared to an ensemble of gridded observational data. Samuel et al. (2023) showed that CORDEX-AFR represents extreme precipitation characteristics over four Southern African river basins better than CORDEX-CORE. Generally, RCMs need a forcing by either GCMs or reanalysis data to take the general large-scale circulation into account. The usage of reanalysis data, like ERA-Interim for CORDEX-CORE, enables a valid estimation of the RCMs as the reanalysis act as so called "perfect boundary conditions" (Wang et al. 2004). In summary, we consider three GCMs (GCM_hist), six RCM-simulations forced by them (RCM_hist), and three RCMs forced by ERA-Interim (RCM_eval).

2.1.3 Data processing

For the data processing, we use the Climate Data Operators (CDO) (Schulzweida 2019). All datasets are analyzed using the smallest common area covered by TAMSAT. Additionally, the data are processed such that the daily precipitation sum [mm] is the basis for further analyses. For spatial comparisons of the reference data, a remapping on the coarsest resolution of the respective datasets shown in the respective figures is done.

For the model data, the historical period of CMIP5 and CORDEX-CORE ends in 2005. Thus, we extend the time series until 2010 by using data from the RCP8.5 scenario as this is closest to recently observed greenhouse-gas emissions (Schwalm et al. 2020). With this, we have a recent climatology which is covered by reference data and upon which future climatologies and related climate change induced impacts can be based and compared to.

As the GCMs have different spatial resolutions, we calculate each index for each individual dataset first and interpolate them afterwards on the coarsest resolution to produce the ensemble mean. The RCM RegCM does not have the standard CORDEX-CORE resolution of 0.22° but approximately 0.225°, which is caused by the underlying Mercator projection. Thus, the RegCM-runs are interpolated to the standard resolution. The regridding for precipitation data is done using an inverse distance interpolation. For temperature data and for integer values, we perform a nearest neighbor interpolation which is more appropriate for the underlying variable's distribution (Casanueva et al. 2020).

Another aspect is the different handling of single models regarding leap years. Thus, we decided to dismiss the 29th of February in all datasets and for all analyses. In HadGEM, a month has 30 days by default. Therefore, the model and forced RCMs have been neglected for the ensembles as annual daily indices and the calculation of the rainy season depend on 365 days. Furthermore, the available CORDEX-CORE data of RegCM's evaluation run do not contain the December 2010. We took this inaccuracy as given and did not perform any processing regarding this aspect.

The quality of single models and their individual ranking within the ensemble strongly depend on the investigated variable, region, time period, and season. Consequently, we focus on the equal-weighted ensemble mean of the respective model simulations. This comes along with the advantage that the mentioned inconsistencies between the different model outputs have a reduced effect on the results compared to the consideration of single models.

The seasonal cycles are calculated by averaging the monthly precipitation sum of individual months over the overlapping period of the datasets. Further, we show linear trends (Wilks 2011) of the reference data. These are calculated over the overlapping period (1983–2019) of the considered datasets and multiplied by ten to get information on the trend per decade. Regarding individual subregions, the respective area is selected and averaged spatially to create a single time series for each subregion.

2.2 Rainy season definition and indices

2.2.1 Definition of the rainy season

To define the rainy season and its climatology, its onset and cessation have to be identified at a daily scale. There are several approaches defining the onset and cessation dates on grid-point scale (Bombardi et al. 2020) as required in this study. We use the method of Dunning et al. (2016) – with some modifications following Weber et al. (2018) – which is a more specialized form of Liebmann et al. (2012) since it can detect more than one rainy season per year. Further, it is a cumulative instead of a threshold-based approach. Thus, it avoids the detection of a so-called "false onset" caused by a single heavy precipitation event (Dunning et al. 2016; Bombardi et al. 2020) – although threshold approaches exist to overcome such limitations as well (e.g., Laux et al. 2008). Recently, this method has been frequently used to detect rainy seasons (Dunning et al. 2018; Weber et al. 2018; Chapman et al. 2020; Ferijal et al. 2021).

In a first step, the climatological cumulative sum of the daily rainfall anomaly is determined for each grid box and afterwards smoothed using a 30-day-running mean. The minimum (maximum) of the climatological cumulative daily $C(d)$ rainfall anomaly ($Q_i - Q$) (Eq. 1) is considered as the onset (cessation) day of the climatological rainy season if the onset (cessation) day is lower (higher) than the four preceding and the four following days. If neither a minimum nor a maximum is found, the smoothing period is extended by 15 days until an equal number of minima

and maxima is detected. Otherwise, a 120-day-running mean is achieved. Thereby, we assume that the first maximum after a preceding minimum defines a rainy season (Weber et al. 2018). In the case that more than two rainy seasons are detected, we consider only the two longest rainy seasons. Furthermore, if the number of days between two rainy seasons is less than 40 or if two rainy seasons overlap, one rainy season is assumed.

Equation 1: Cumulative daily precipitation anomaly.

$$C(d) = \sum_{i=1^{st}Jan}^{d} Q_i - \overline{Q}$$

In a second step, the onset and cessation of the rainy seasons are determined for each individual year. This is done by calculating the cumulative rainfall anomaly (daily rainfall minus climatological daily mean rainfall over the period) and searching for the absolute minimum/maximum 20 days prior to the climatological onset date to 20 days past the climatological cessation date for each year.

A limitation is that the algorithm detects a rainy season independent of the absolute precipitation amount. To avoid a misleading detection in arid climates, we solely consider grid points with an annual precipitation ≥ 100 mm for the final rainy season masks. The masks display the binary behavior of the presence or absence of the rainy season on each day of the year averaged over a climatological period.

2.2.2 Climate and agricultural relevant indices

Table 3 gives an overview of the indices used in this study. Most indices focus on precipitation characteristics, which are defined by the ETCCDI (Zhang et al. 2011), on annual scale or during the rainy season. This enables a comparison of these characteristics over the year as well as a separation between the rainy and the dry seasons. Additionally, characteristics defining the rainy season are defined as rainy season-related indices. Precipitation also is the basis for the agricultural indices. However, information on either the actual or potential evapotranspiration is required as well. Thus, these are dealt with in detail in Sect. 2.2.3.

2.2.3 Agricultural indices

The Crop Water Need (CWN) is the amount of water needed for the optimal growth of individual crops. This index depends on the temperature variables required by the applied potential evapotranspiration scheme, precipitation, and time dependent plant properties. There are different approaches available to derive CWN. We use the potential evapotranspiration ($ET0$) based on the Hargreaves scheme (Hargreaves and Samani 1985) requiring the mean, minimum, and maximum temperature. We build the mean of $ET0$ over the days of the corresponding plant phase weighted by a specific crop factor (Kc) (Eq. 2). The length of each phase, also called stages, and the Kcs are plant specific. A common characteristic of the plant

Table 3 Climate indices used in this study. Indices marked by * can be calculated based on other indices from the table. The calculation of precipitation-based indices depend on the considered time period (annual, rainy season (appended "_rs")), the rainy season-related indices define the rainy season. Per definition, the agricultural indices are calculated over the rainy season as well (see Sect. 2.2.3)

Category of climate indices	Abbreviation	Definition	Input variables
Precipitation-based indices	RTOT	Total precipitation from wet days over period (a wet day with precipitation ≥ 1 mm)	Daily precipitation data in mm
	Rd	Number of rainy days over period (days with precipitation ≥ 1 mm)	Daily precipitation data [mm]
	Dd*	Number of dry days over period (days with precipitation < 1 mm)	Daily precipitation data [mm]
	CWD	Maximum number of consecutive wet days over period	Daily precipitation data [mm]
	CDD	Maximum number of consecutive dry days over period	Daily precipitation data [mm]
	nCDD	Number of at least five consecutive dry days over period	Daily precipitation data [mm]
Rainy season-related indices	Ons	Starting day of rainy season	Daily precipitation data [mm]
	Ces	Ending day of rainy season	Daily precipitation data [mm]
	dur*	Duration of rainy season	Daily precipitation data [mm]
Agricultural indices	CWN	Water needed during the crop stages	Daily potential evapotranspiration [mm]
	IWR	Water balance of the root zone (difference between the reference evapotranspiration (ET0) and effective precipitation)	Daily CWN and precipitation data [mm]
	WA	Water available for crop growth (difference between actual precipitation and actual evapotranspiration)	Daily precipitation [mm] and actual evaporation [mm] or latent heat flux [Wm^{-2}] data

stages is that the initial stage (IS) starts with the onset of the first rainy season. We used the coefficients published by the FAO (Allen et al. 1998) and chose 12 prominent African crops for our analysis. However, as this would be beyond the scope of this paper, we solely focus on Maize (grain) of the long growing season due to its widespread use in entire sub-Saharan Africa (Cairns et al. 2013). The corresponding *Kc*s for each growing stage are given in Table 4.

Equation 2: CWN per plant phase.

$$CWN_p = \frac{1}{e_p - s_p}\sum_{i=s_p}^{e_p} CWN_i = \frac{1}{e_p - s_p}\sum_{i=s_p}^{e_p}\left(ET0_i * Kc_p\right) = Kc_p\left(\frac{1}{e_p - s_p}\sum_{i=s_p}^{e_p} ET0_i\right)$$

$i = day\ of\ year$

$p = phase\ (IS,\ CDS,\ MSS,\ or\ LSS)$

$CWN_p = crop\ water\ need\ per\ plant\ phase\ [mm]$

$CWN_i = daily\ crop\ water\ need\ [mm]$

$Kc_p = crop\ factor\ per\ plant\ phase$

s_p = *day of year when the* p^{th} *plant phase starts*
= *for IS it is the start day of the rainy season and* $e_{(p-1)+1}$ *for other phases*

$e_p = day\ of\ year\ when\ the\ p^{th}\ plant\ phase\ ends$

$ET0_i = daily\ potential\ evapotranspiration\ [mm]$

The second index is the irrigation requirement (IR, Eq. 3). It is defined as the amount of water that is required in addition to precipitation in order to satisfy the CWN. We calculate IR as the difference between CWN and the effective precipitation, which is a value derived from precipitation as follows (Ali and Mubarak 2017):

Equation 3: IR per plant phase.

$$IR_p = \frac{1}{e_p - s_p}\sum_{1=s_p}^{e_p}\left(CWN_i - efftp_i\right)$$

$$efftp_i = \begin{cases} 0 & for\ tp_i < 6.5mm \\ 75 & for\ tp_i \geq 75mm \\ else & tp_i \end{cases}$$

$tp_i = daily\ precipitation\ [mm]$

$efftp_i = daily\ effective\ precipitation\ [mm]$

This equation indicates that for daily precipitation amounts below 6.5 mm, the daily values for IR_i equal the daily CWN_i. However, we present IR as the mean of the daily difference values per plant phase because this leads to differences compared to *CWN*.

As a third index we consider WA (water availability). It is calculated as the mean daily values per plant phase which are derived from the difference between daily precipitation and the actual evapotranspiration *ET*. *ET* is based on the daily surface latent heat flux [in Wm^{-2}] (Eq. 4, Allen et al. 1998). If the soil water storage is neglected this index is analogous to the surface runoff (Sylla et al. 2018).

Equation 4: WA per plant phase.

$$WA_p = \frac{1}{e_p - s_p}\sum_{1=s_p}^{e_p}\left(tp_i - ET_i\right)$$

$$ET_i = \frac{\frac{hfls_i}{1000000}}{2.45\frac{MJ}{kg}}$$

$ET_i = daily\ actual\ evapotranspiration\ [mm]$

$hfls_i = surface\ latent\ heat\ flux\ [\frac{W}{m^2}]$

Table 4 Plant-specific properties per plant phase of Maize (grain) (Allen et al. 1998)

Plant: Maize (grain)	Initial Stage (IS)	Crop Development Stage (CDS)	Mid-Season Stage (MSS)	Late-Season Stage (LSS)
Crop Factor (Kc) [-]	0.4	0.8	1.15	0.7
Phase Length [days]				
Long	30	50	60	40
Short	20	35	40	30

2.3 Validation metrics

To validate the performance of the agricultural indices based on the different model ensembles compared to the reference data, we use three different metrics: (1) the mean absolute error (MAE), (2) the Kling-Gupta-Efficiency (KGE), and (3) the Taylor Skill Score (TSS). The MAE describes the average of the absolute differences between the model and the reference data with lower values referring to a higher model quality (Wilks 2011).

The KGE (Gupta et al. 2009) considers the correlation, the bias, and the variability of the model and the validation:

Equation 5: Kling-Gupta-Efficiency.

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_m}{\sigma_v} - 1\right)^2 + \left(\frac{\mu_m}{\mu_v} - 1\right)^2}$$

$r = Pearson\ correlation\ coefficient$

$\sigma = standard\ deviation\ of\ model\ (m)\ and\ validation\ (v)$

$\mu = arithmetic\ mean\ of\ model\ (m)\ and\ validation\ (v)$

With this, the three components are weighted equally. The KGE can represent positive and negative values $[-\infty;1]$ with higher values showing a better representation.

The TSS (Taylor 2001) is based on the correlation coefficient and the variability and covers values between 0 and 1. The higher a value the better.

Equation 6: Taylor Skill Score.

$$TSS = \frac{4(1+r)^4}{\left(\frac{\sigma_m}{\sigma_v} + \frac{1}{\frac{\sigma_m}{\sigma_v}}\right)^2 (1+r_0)^4}$$

$r_0 = maximum\ correlation\ attainable,\ here\ 0.999999$

We use these scores as MAE prevails the unit and value range of the indices while KGE and TSS combine several characteristics of the model and reference data. The skill scores are applied on the agricultural indices presented in Sect. 3.4. First, we assess the mean temporal evolution over the period 1981–2010 for precipitation and CWN. For this, we use the absolute instead of the accumulated values shown in the respective figure and remove the seasonal cycle before calculating the skill scores. Second, we validate the spatial representation of the agricultural indices by remapping the climate models to the resolution of ERA5Land acting as reference data. This has the advantage that the number of grid points is the same for all model ensembles. Further, it preserves the added value of high spatial resolution and the corresponding spatial variance – which becomes more smoothed with coarser model resolutions – of the reference data. A limitation of this procedure is the creation of information on the fine grid without the related fine-scale spatial information, e.g., orography, which is considered in dynamical and statistical downscaling approaches.

3 Results

3.1 Comparison of reference data

The first assessment of the eight precipitation datasets listed in Table 1 is based on the respective time series generated by the spatial mean of annual precipitation sums (Fig. 2). For entire Africa (AFR, bottom left), the previously mentioned large spread among the data becomes clear. This spread is not only related to the absolute amount of annual precipitation, but also to the interannual variations and the trends in the time series. Exemplarily, ARC2 consistently shows lower values than the other considered data yet has an increasing trend. This behavior is caused by the number of missing values which decreases over time and a remarkable high precipitation amount in 2020. Due to these limitations, ARC2 is not suitable for detailed analyses of subregions and further aspects, regardless its long time series and high spatial resolution. Generally, it can be observed that the spread decreases over time and the data are more in line with each other in recent years.

Based on this, only ERA5Land, CHIRPS, and TAMSAT are considered for the time series of the subregions and subsequent analyses. ERA5Land provides all variables required for the agricultural indices. CHIRPS and TAMSAT have a high spatial resolution and proofed reliability (Dembélé et al. 2020; Satgé et al. 2020).

These three time series have similar annual peaks within most subregions, leading to high correlation coefficients among each other (not shown). Regarding the linear trends (colored numbers in the subfigures of Fig. 2 and 1983–2019), the picture is more complicated. CHIRPS and TAMSAT have the same magnitude and direction in most subregions. However, more complex areas like Central Africa or ETP_H show different magnitudes or even signs (CAF_S). ERA5Land consistently shows the highest precipitation sum of the selected datasets in ETP_H, SAF_W, and SAF_E, although some peaks are outbid by the spread of the other datasets. For GN_C, CAF_N, CAF_S, and SAH_E, the first half of the period of ERA5Land (and ERA5, not shown) is marked by the highest precipitation amount as well. This behavior changes around the year 2000 when the sum decreases and better corresponds with the other two datasets. This "abrupt transition" (Hersbach et al. 2020)

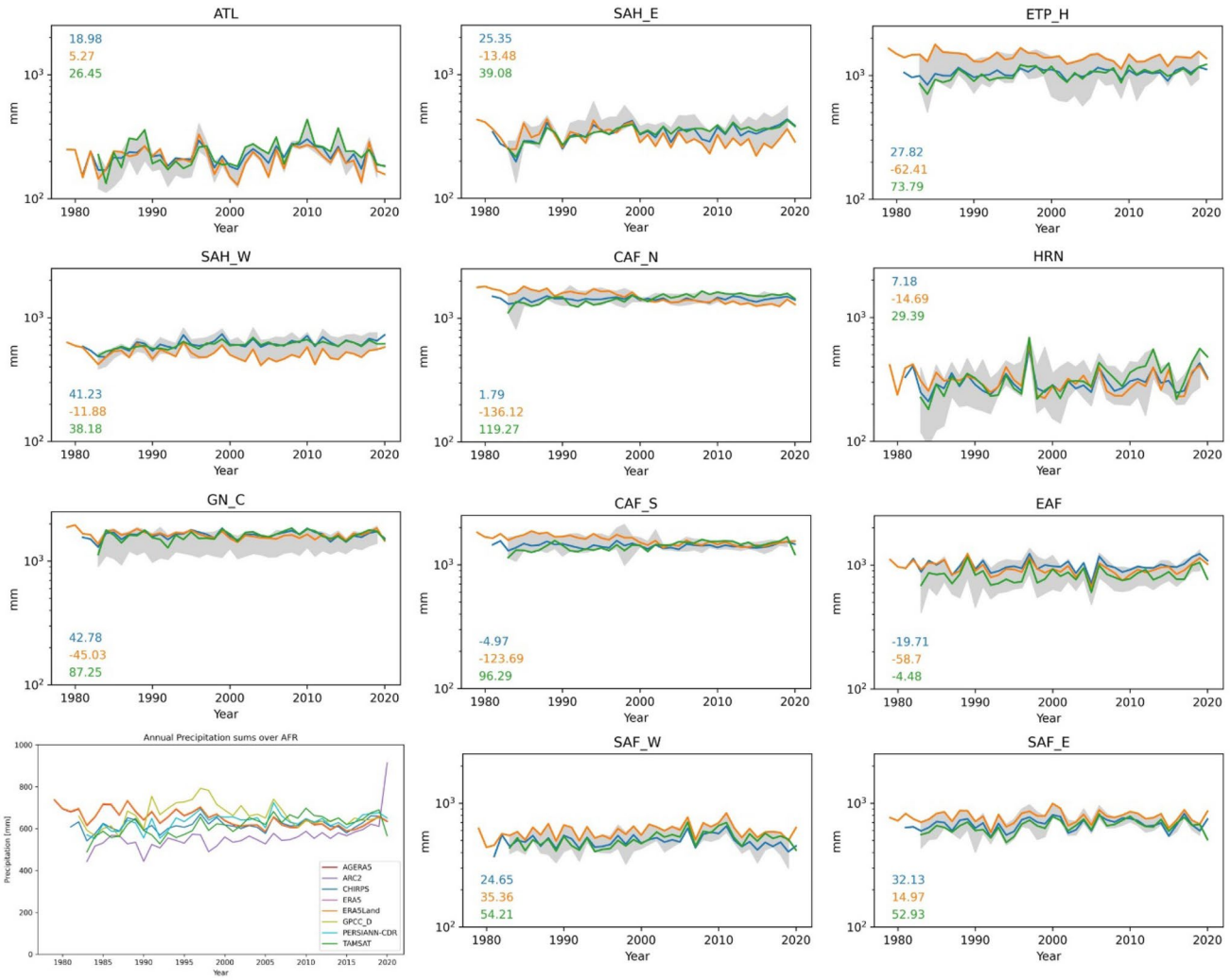

Fig. 2 Time series of yearly precipitation sums for Africa and the respective subregions for different gridded precipitation datasets. While AFR (bottom left) contains all selected datasets (Sect. 2.1.1), the other plots are limited to CHIRPS, ERA5Land, and TAMSAT. Numbers show the linear trend [mm/decade] over the period 1983–2019. The gray-shaded area shows the spread (minimum and maximum) of all eight datasets in the subregions and during their overlapping period (1983–2019). Furthermore, the subregional plots use a logarithmic scale

is caused by the assimilation of different satellite data and underlines data inhomogeneity as a general issue of reanalyses. Consequently, the linear trends are affected substantially resulting in negative, and thus diverging, signs over most subregions compared to CHIRPS and TAMSAT. Solely the trends in ATL and Southern Africa show the same sign in all three datasets. The inconsistency of precipitation trends has been highlighted, e.g., for West (Paeth et al. 2011; Dosio et al. 2020) and entire Africa (Zebaze et al. 2019).

Figure 3 shows the spatial climatologies and differences of CHIRPS, ERA5Land, and TAMSAT. Although the time series in Fig. 2 and the overall patterns of the yearly precipitation sum show a good match among each other, there are large differences in some areas. Considering the different origins (station, satellite, and/or reanalysis) of the datasets and their processing methods, it is reasonable that there are some interpolation fragments (e.g., near Marrakech, Morocco) caused by individual stations. Especially the discrepancy between subtropical and tropical regions, where the climatological difference between the datasets reaches more than 300 mm, is remarkable and has already been noted. While CHIRPS and TAMSAT are, apart from EAF, in good accordance, ERA5Land shows large differences in Central Africa (CAF_N and CAF_S) and ETP_H. This is partially caused by the mentioned inhomogeneity that is pronounced in these regions. Further, the representation of the Intertropical Convergence Zone (ITCZ) (Quagraine et al. 2020) and parameterizations of subgrid processes (Sun et al. 2018) can cause these differences. Consequently, ERA5 as

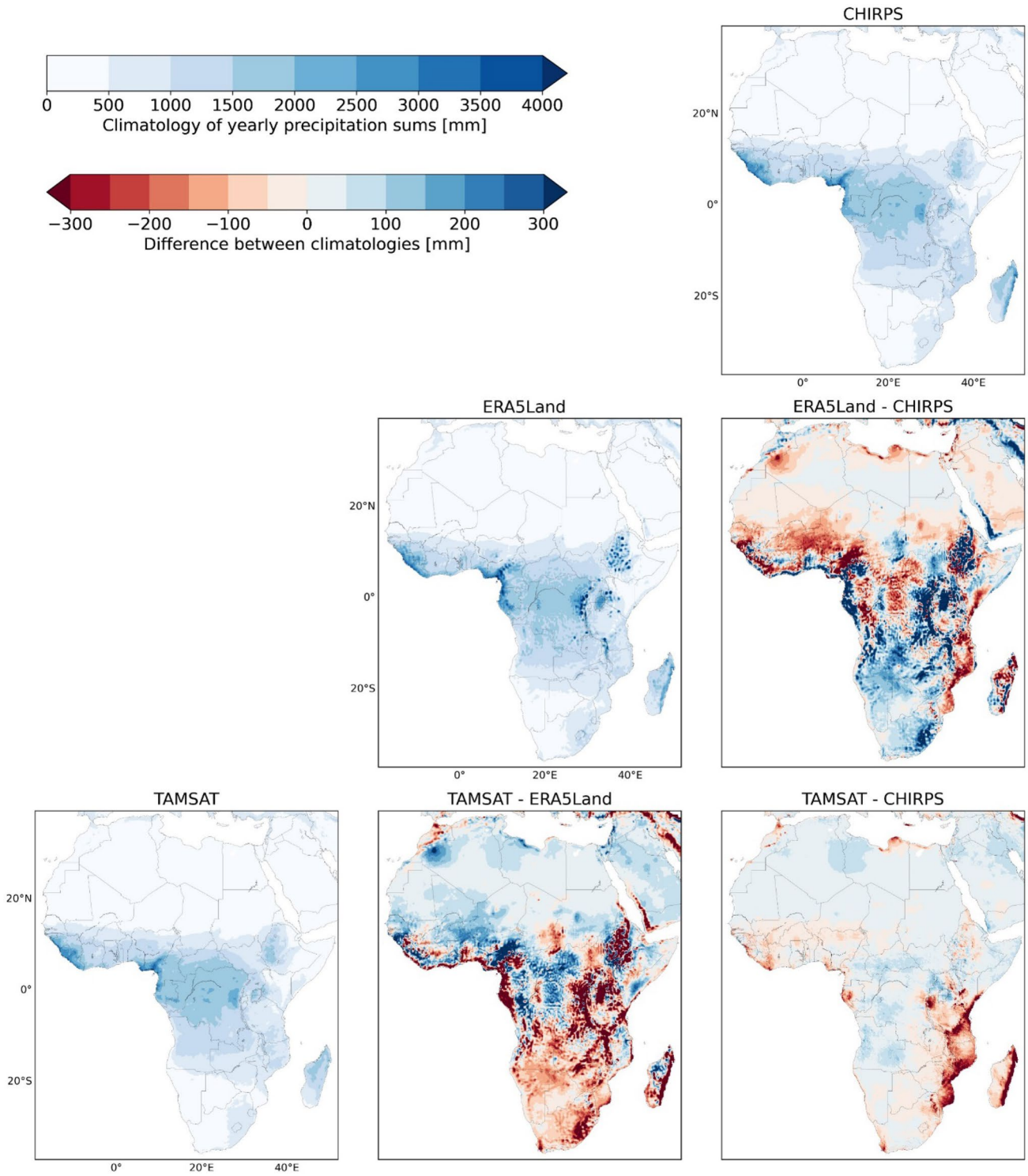

Fig. 3 Climatologies of annual precipitation sums over Africa of CHIRPS, ERA5Land, and TAMSAT for their overlapping period (1983–2010) and their respective differences. The data is remapped to ERA5Land (0.1°) as this dataset has the coarsest resolution

the parent product of ERA5Land has shown to be less skillful in the tropics than in the extratropics (Lavers et al. 2022).

Focusing on the mean seasonal cycle of rainfall averaged over the subregions (Fig. 4), the three observational datasets exhibit differences among each other that strictly depend on the individual subregions. These differences are also underlined by the spread of all considered datasets. However, all three highlighted datasets are in line

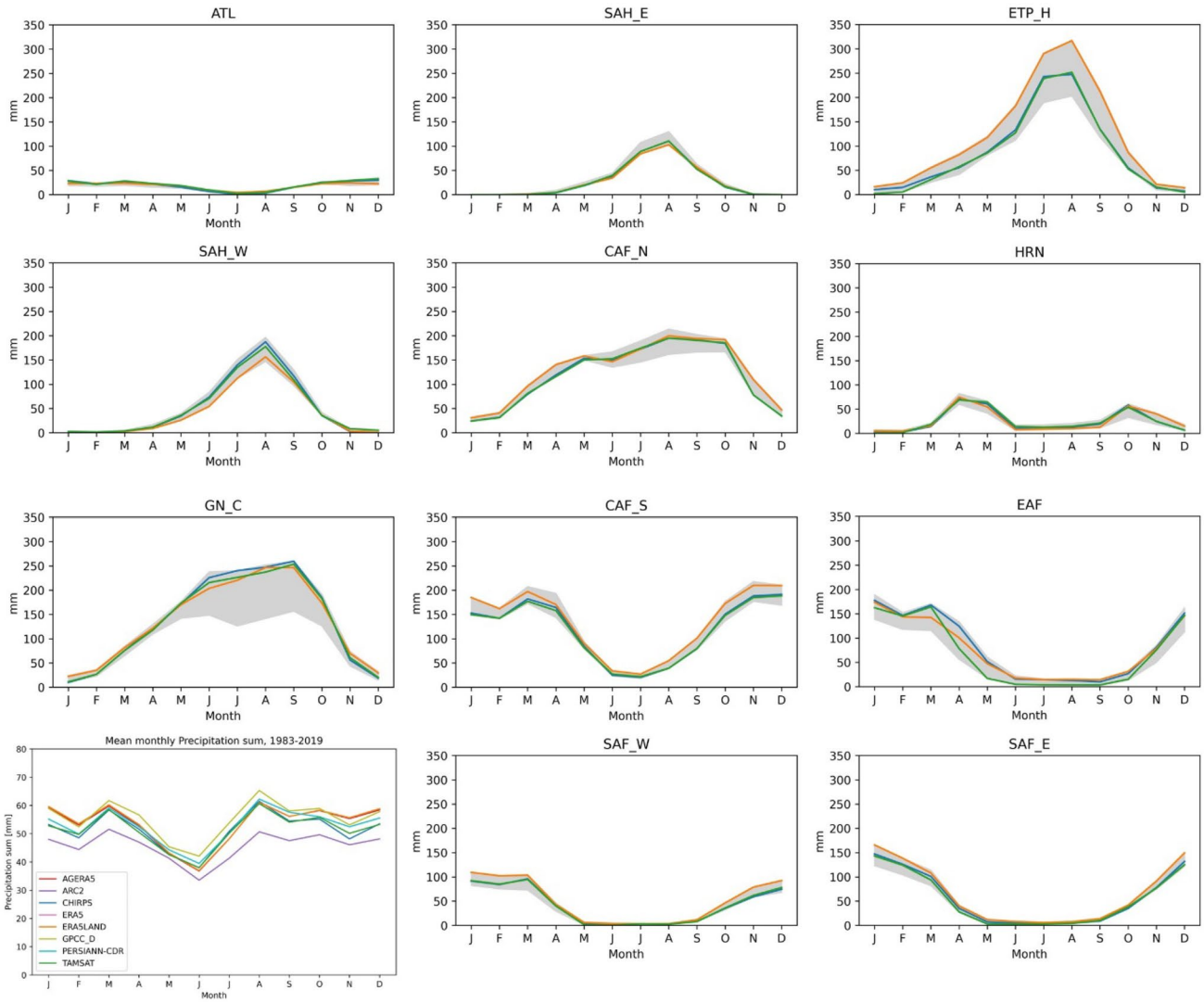

Fig. 4 Mean seasonal cycles of monthly precipitation sums from CHIRPS, ERA5Land, and TAMSAT for the period 1983–2010 for the different subregions. The gray-shaded area shows the spread (minimum and maximum) of all eight datasets plotted for entire Africa (bottom left)

with each other with CHIRPS and TAMSAT being more similar. Solely ETP_H with its complex topography shows larger discrepancies. Having the spatial differences in mind, it is shown that there are some differences among the datasets that balance out each other by calculating the spatial mean – especially within the tropical subregions. Additionally, it can be stated that for most subregions the inconsistency between the datasets as well as their standard deviation increases with wetter conditions during the seasonal cycle. The former is not the case for CAF_N and EAF where differences are larger before and after the precipitation peak than during the peak.

From Fig. 4 it also becomes clear that there are three classes of subregions: Regions with one rainy season (e.g., ATL, SAH_E/W, ETP_H, and SAF_W/E); regions with two separate rainy seasons over the year (e.g., HRN); and regions representing a mixture of these classes (especially CAF_N/S, and GN_C). In the latter, the spatial mean is built over regions with one and two rainy seasons. This is shown as an example in Fig. 8 for CHIRPS and is treated in more detail in Sect. 3.3.1.

In summary, the interannual behavior of the precipitation time series and their annual cycle is in good accordance with CHIRPS, TAMSAT, and ERA5Land. Considering the spatial pattern, CHIRPS and TAMSAT also match well. Due to the combination of station and satellite data used to generate CHIRPS, its longer time period, and a couple of validation studies within Africa approving its high quality in different subregions (e.g., Dinku et al. 2018; Harrison et al. 2019; Dembélé et al. 2020; Satgé et al. 2020; Tarek et al. 2021), we define CHIRPS as our

baseline for precipitation and precipitation-based indices in the subsequent analyses. As the reanalysis of ERA5Land contains the variables that is necessary to calculate the agricultural indices and has the advantage of physical consistency between these, its precipitation characteristics are shown in the results as well to have an idea of its deviations from CHIRPS' characteristics.

3.2 Evaluation of CMIP5 and CORDEX-CORE

Subsequently, an evaluation of the precipitation characteristics of the climate models' ensemble means is done. In Fig. 5, their annual cycles are compared with CHIRPS and ERA5Land for the African subregions. To account for the spread of CORDEX-CORE, the shaded areas represent the minimum and maximum of the respective RCM ensembles.

With exception of a strong underestimation of RCM_eval at GN_C between July and September, all model ensembles are able to represent the general monthly precipitation characteristics well. For the historical ensembles, this is also shown by Dosio et al. (2021a). However, the authors did not consider the evaluation ensemble. Further, the selection of reference data differs between our study and the work by Dosio et al. (2021a). Therefore, we justified that showing the seasonal cycle of CORDEX-CORE together with the reference data used in our study is important. As RCM_hist is not showing this behavior at GN_C and not even the ensemble minima and maxima are overlapping in July and August, the drop might be induced by the forcing data of ERA-Interim as indicated by Nikulin et al. (2012). Additionally, Quagraine et al. (2020) found a too narrow northward propagation of

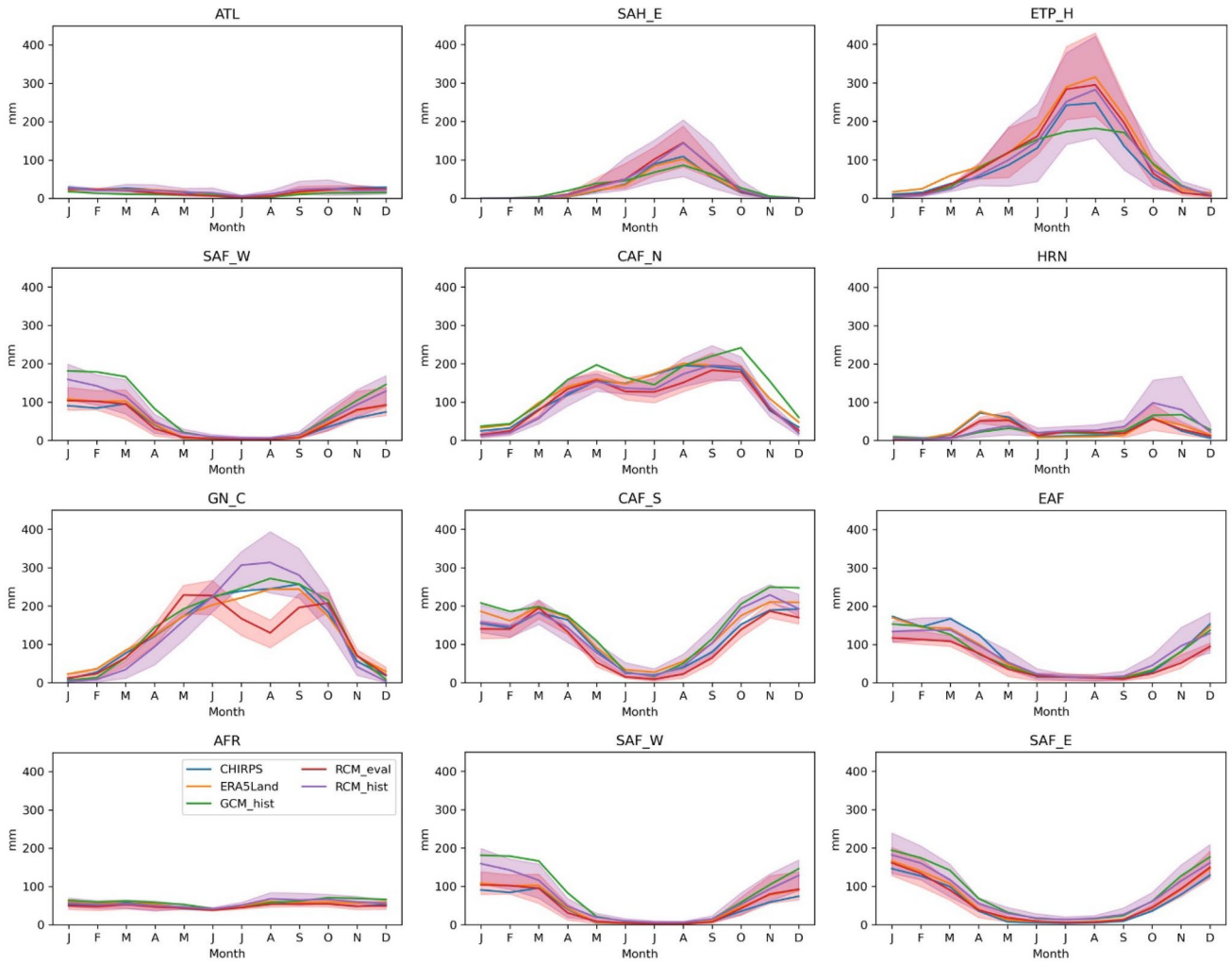

Fig. 5 Mean seasonal cycle of rainfall for the ensemble means from the GCMs (GCM_hist), RCMs forced by ERA-Interim (RCM_eval) and the GCMs (RCM_hist), respectively, and from CHIRPS and ERA5Land for AFR and its subregions over the period 1981–2010. For RCM_eval and RCM_hist, the ensemble minimum and maximum is shown by the shaded areas

the monsoonal precipitation in ERA-Interim. This results in overestimated precipitation amounts along the Guinean coast and underestimated ones in the Sahel zone. However, this cannot be observed in the comparison of the RCM ensembles conducted here.

The GCMs show a strong underestimation of precipitation in ETP_H which is caused by the complex topography of the region that cannot be represented by the coarse resolution of the GCMs. As RCM_hist's annual cycle is much closer to the reference data, an added value of the dynamical downscaling is detected for that region. The overestimation of GCM_hist in SAF_W, SAF_E, CAF_S, and – to a smaller extent – EAF is in accordance with Zebaze et al. (2019). Both RCM ensembles perform better than GCM_hist. However, as the historical run is closer to the GCMs than the evaluation run, the effect of the forcing data on the RCM quality can be seen. Interestingly, this overprinting effect of the forcing GCMs on the RCMs is not present in the Sahel. On the one hand, the GCMs show an underestimation in SAH_W. However, both RCM ensembles are in good accordance with the reference data. On the other hand, the GCMs match the reference data well in SAH_E while the RCM ensembles show a systematic overestimation. As a consequence, the RCMs' behavior has to be a result of the rainfall-related model physics, which improves the GCM-forcing over the Western part but fail in the Eastern part.

To compare the spatial differences in Fig. 6, all datasets have been interpolated to the resolution of GCM_hist. The GCMs show an overestimation in Southern Africa and CAF_S and an underestimation in EAF. This behavior is in line with Fig. 5 and the findings of Zebaze et al. (2019). As indicated in Sect. 3.1, the model biases with regard to CHIRPS are stronger than to ERA5Land. Interestingly, the differences in SAH_W show a different sign compared to ERA5Land and CHIRPS. This is a notable example of the non-uniform representation of precipitation in available datasets and the resulting challenges for model evaluation.

Mostly, RCM_eval is in better accordance with CHIRPS than with ERA5Land in most parts of Africa. An exception of this is Southern Africa. However, the sign of the bias is the same at most grid points. For RCM_hist, the same difference patterns as for GCM_hist can be detected. This is highlighted by the general overestimation in Southern Africa. Compared to CHIRPS, this behavior is present along the western coast and even north of the equator. The overestimation in Southern Africa might be caused by a warm sea surface temperature bias which occurs in MPI-ESM-LR (Weber et al. 2023a). However, EAF and Central Africa are simulated closer to the reference data by RCM_hist than by GCM_hist and RCM_eval. This is also true in West Africa compared to RCM_eval.

We can conclude two major issues regarding the model performance. First, the quality of the model ensemble strongly depends on the subregion considered. Second, the forcing data can have a significant effect on the RCM simulations. This is a known issue (Wang et al. 2004; Di Luca et al. 2016; Sørland et al. 2021) and has been investigated for Southern Africa by Karypidou et al. (2022) in more detail. The authors state that the RCMs are acting more independently at the beginning of the rainy season where precipitation is a small-scale process and mainly coupled to land surface-atmosphere interactions. During the rainy season, when precipitation is governed by large-scale effects, the GCM forcing plays a stronger role. Finally, the authors conclude that RCMs are able to counteract GCM-induced biases and add value to the simulations in Southern Africa.

3.3 Representation of the rainy season and related indices

3.3.1 Rainy season

We now focus on the occurrence of one and two rainy seasons and their respective onset and cessation. Figure 7 shows the number of rainy seasons based on CHIRPS. The Sahara and Namib deserts as well as some parts of HRN are not characterized by a rainy season as these regions show arid conditions (see Sect. 2.2.1). HRN is the only subregion which is clearly dominated by two rainy seasons while GN_C, Central Africa, EAF, and ETP_H show larger areas where one as well as two rainy seasons occur. The other subregions show only small areas or single grid points with a second rainy season (e.g., SAH_E and SAF_W). Some of these local occurrences are caused by the high resolution of CHIRPS (0.05°). This can be concluded from a comparison with the results of Dunning et al. (2016) and Chapman et al. (2020), who also consider CHIRPS but in coarser resolutions (0.25°) and over other periods. However, the areas of one and two rainy seasons agree well with the results of these two studies. As the precipitation between various datasets differs (Sect. 3.1) the rainy seasons mostly differ as well, as found in Chapman et al. (2020).

To be consistent when comparing the datasets, we use the rainy season mask of CHIRPS for all subsequent rainy season-related analyses. For this purpose, the CHIRPS mask is remapped to the respective dataset's resolution using a nearest neighbor interpolation. Figure 8 shows the spatial mean of the onset and cessation dates in four selected subregions. The selection has been made to cover all described rainy season types across Africa. For SAH_W with its one rainy season, the datasets are very similar in representing the onset (rs1_ons) and cessation (rs1_ces). They show a higher standard deviation for the onset than for the cessation days. However, all model ensembles indicate an earlier

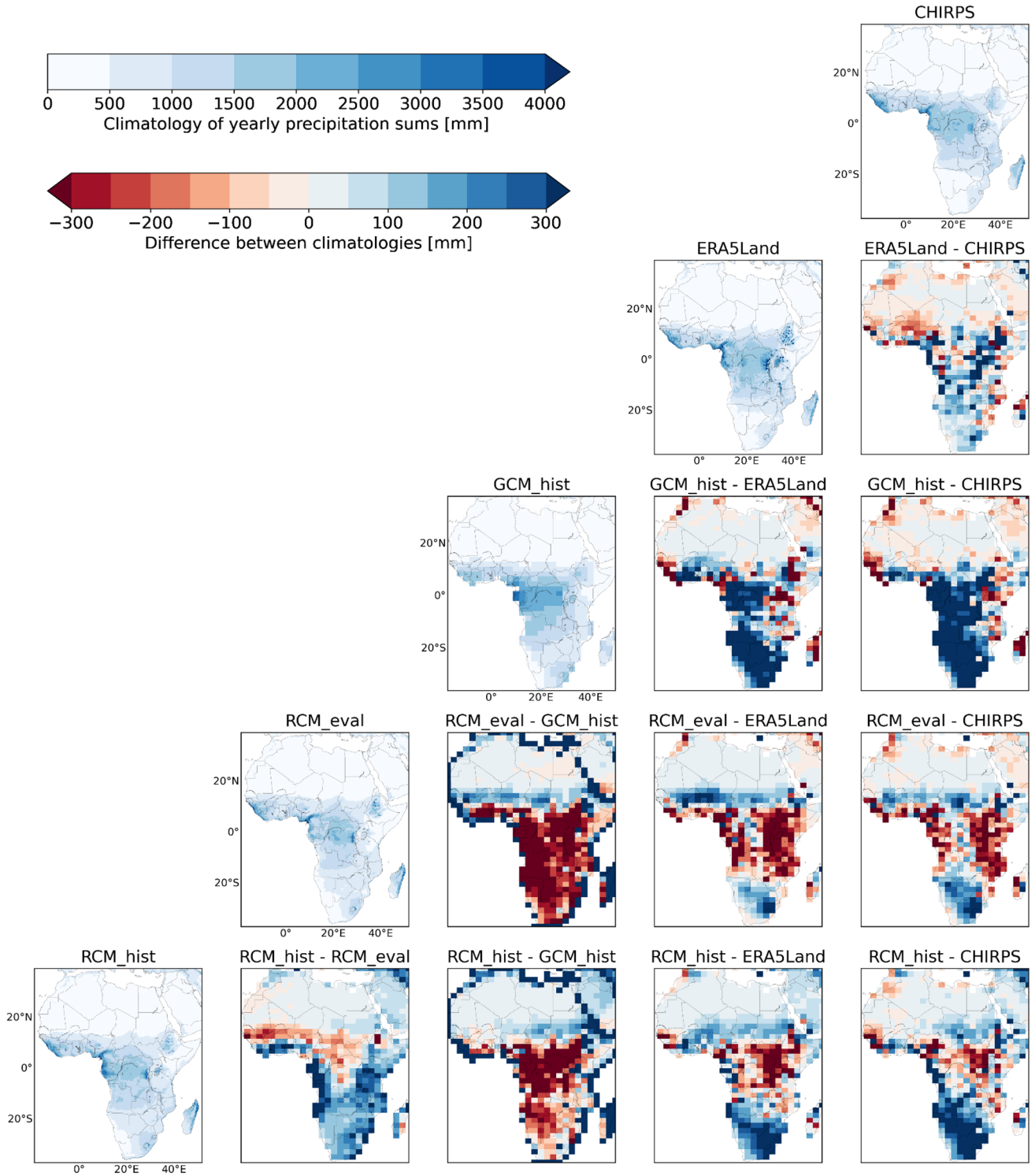

Fig. 6 Spatial differences of annual precipitation sums between the reference data and the model ensembles (1981–2010). For the absolute values, the dataset's individual resolution is conserved, for the differences, all datasets are interpolated to the coarsest resolution (NorESM1-M)

onset which results in a longer rainy season compared to the reference data.

In the area of GN_C having one rainy season, its duration is longer than in SAH_W. This is reasonable due to the northward monsoonal propagation. While the reference data are quite similar as well, the models show a larger spread of the onset and cessation dates than in SAH_W. Compared to the reference data, this is expressed in an earlier onset in

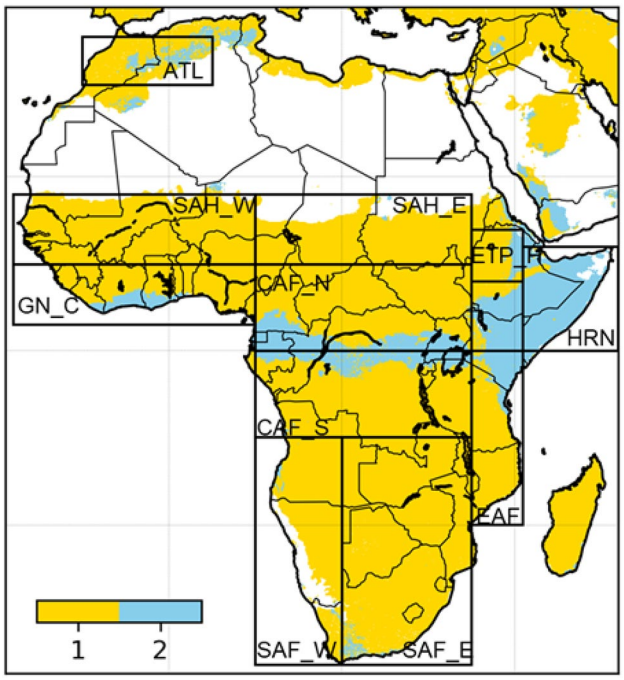

Fig. 7 Mask of the occurrence of one or two rainy seasons in Africa based on CHIRPS' climatology (1981–2010). Ocean and arid areas are excluded

GCM_hist and RCM_eval compared to the reference data, but the latest onset in RCM_hist. As a consequence, the latter has the shortest rainy season while its driving models simulate the longest. Thus, the forcing is not overruling the dynamical downscaling since this represents small-scale processes more adequately as it has also been stated by Karypidou et al. (2022) for Southern Africa. A too short rainy season in GN_C in CORDEX-CORE, being the successor of CORDEX-AFR, is in line with the results of Chapman et al. (2020). Most datasets show a higher standard deviation of the onset compared to the cessation. A notably different behavior is present in RCM_eval where the cessation's uncertainty is much higher than in the other datasets. This could be related to the already mentioned inadequate representation of precipitation in ERA-Interim over West Africa which has been improved with ERA5 (Quagraine et al. 2020). However, considering solely grid points with two rainy seasons in GN_C (GN_C rs2), it seems that ERA5Land (like ERA5 and AGERA5, not shown) is not able to depict the rainy season dates in an adequate way as the differentiation between the first and second rainy season is not possible. This behavior is caused by the usage of CHIRPS's rainy season mask which differs noticeably from the rainy season mask of the ERA5-products. In fact, ERA5Land shows a smaller area with two rainy seasons (map not shown). This leads to an overweight of the first rainy season as the grid points with one rainy season in ERA5Land are attributed to the CHIRPS-area with two rainy seasons. However, the areas with a common second rainy season agree well in both reference datasets regarding the onset (rs2_ons) and cessation (rs2_ces) dates. The onset of the first rainy season is also shown in an adequate way. GCM_hist and RCM_hist also show one long rainy season but are not able to simulate a second rainy season. One could argue that this is due to the fact that the rainy season masks of the ensembles do not show an overlap with CHIRPS. This also shows the inability of these models to represent the second rainy season. Actually, the rainy season masks of GCM_hist and RCM_hist do not represent an area with two rainy seasons in GN_C. This is in line with the results of Chapman et al. (2020), who examined the rainy seasons' representation in CMIP5 and CORDEX-AFR. On the other hand, RCM_eval is able to represent the second rainy season at GN_C. Hence, it demonstrates that the RCMs depend on the ability of the forcing data when it comes to simulating two rainy seasons. The problem that both rainy seasons are too short is also occurring in the area of GN_C.

Regarding the rainy season dates, CHIRPS and ERA5Land are in line with each other for the first rainy season in HRN. The second rainy season begins later in ERA5Land but shows a comparable duration. RCM_eval agrees well with ERA5Land apart from a later onset of the first rainy season. This onset is delayed in GCM_hist which also has a higher standard deviation regarding the first rainy season. The onset of the second rainy season is represented well but the cessation is delayed. Interestingly, the standard deviation of the second rainy season is very small compared to the high uncertainty of the first rainy season. RCM_hist is not able to adequately simulate the first rainy season since its onset date lies around the cessation date of the reference data and RCM_eval. As the second rainy season is simulated slightly better than in GCM_hist, this results in a very short break between the two rainy seasons. From these three subregions, we can conclude that the comparison of reference data in regions with a bimodal seasonal cycle is challenging. However, as the onset dates of the first rainy season are quite similar, this shortcoming is not affecting the agricultural indices considered later-on since the first onset is the relevant factor for these. The performance of the model ensembles strongly depends on the considered subregion, rainy season and – for the RCMs – forcing data.

3.3.2 Precipitation-related indices

Subsequently, we examine the precipitation-related indices introduced in Sect. 2.2.1 on a climatological scale. As it is noted previously, the rainy season shows a high spatial variability which was balanced by using the spatial mask of CHIRPS to calculate onset and cessation of the rainy season of the individual datasets. To calculate the indices

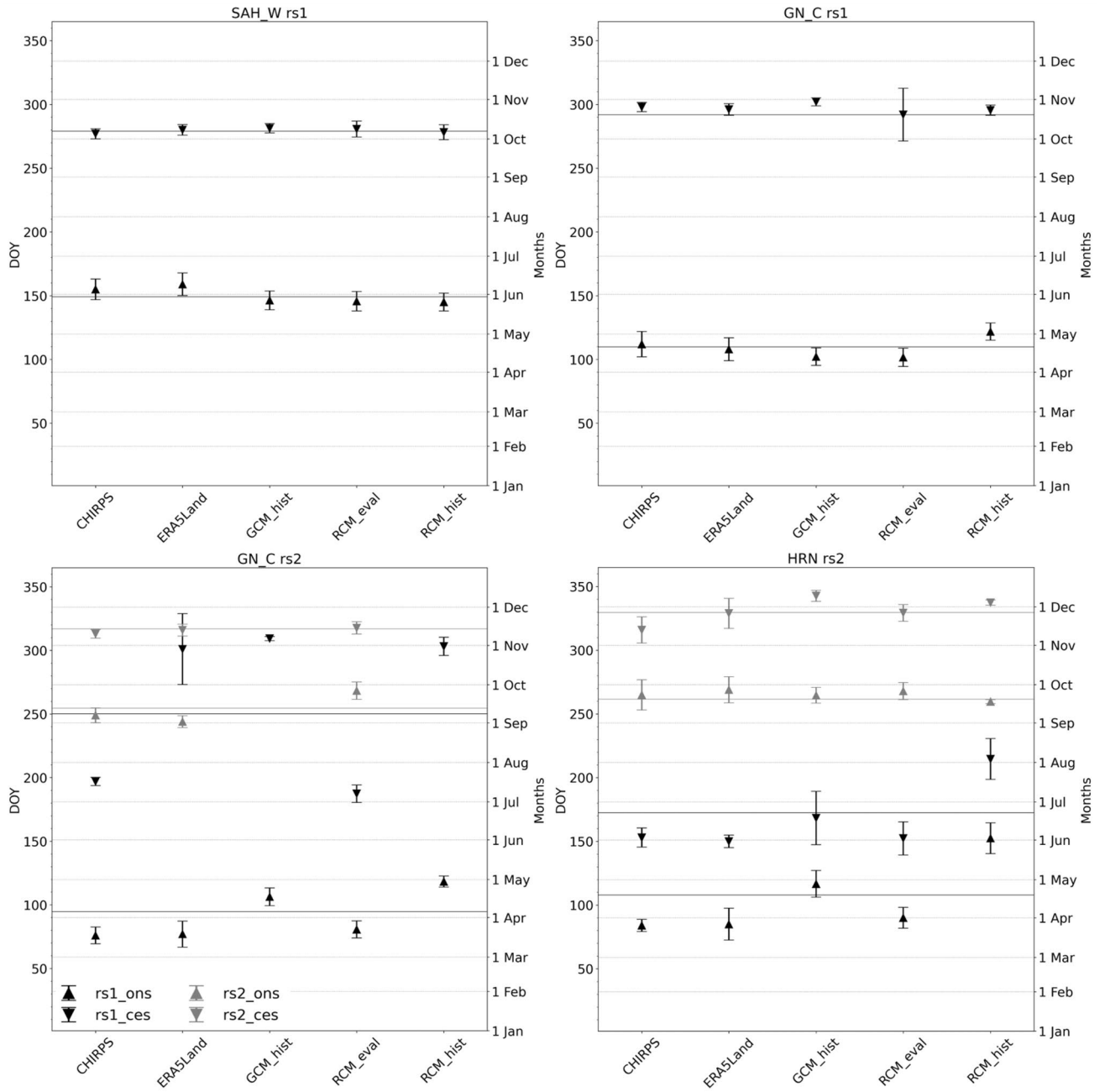

Fig. 8 Spatial mean of onset (ons) and cessation (ces) dates of the rainy seasons in the subregions SAH_W, GN_C, and HRN for the reference data and the model ensembles. The center of an arrow represents the median of all grid points belonging to either one or two rainy seasons. The bars represent the spatial standard deviation over the area. The horizontal line represents the mean of the five datasets. If a subregion contains two rainy seasons, the first is marked by black and the second by gray symbols

over their "own" rainy season, the temporal occurrence of the rainy season also adds variability, depending on the examined region. Hence, we decided to apply the temporal rainy season mask of CHIRPS as well. This leads to an increased comparability of the models with CHIRPS and, thus, allowing a more stringent assessment of their ability to reproduce the indices. Nevertheless, the spatio-temporal differentiation between one or two rainy seasons is not done subsequently as we calculate the spatial mean of each subregion.

Figure 9 displays the total precipitation sum over the year (RTOT) and during the rainy season (RTOT_rs) over the three selected subregions. The respective relative contribution of RTOT_rs to RTOT is noted in the bars. Focusing on SAH_W, the relative contribution of the rainy season is similar in all five datasets. This is

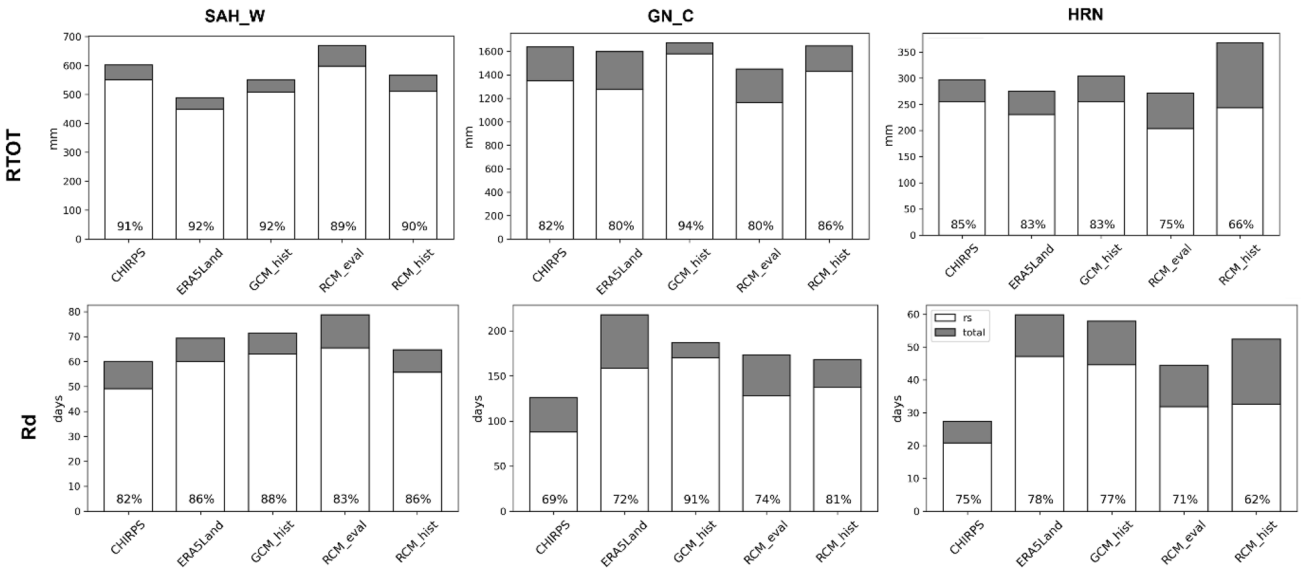

Fig. 9 Spatial climatology (1981–2010) of the annual precipitation (RTOT) and the amount of precipitation that fell during the rainy season (RTOT_rs) in the upper row. The number of rainy days (Rd and Rd_rs) over the two periods is displayed in the bottom row. The three subregions SAH_W, GN_C, and HRN are considered for reference and model data. The temporal rainy season mask of CHIRPS is applied to all datasets

true despite the different absolute amounts of RTOT and RTOT_rs. Here, RCM_hist is closest to CHIRPS while RCM_eval shows an over- and ERA5Land an underestimation. However, the case of SAH_W is comparably simple as the entire area is marked by one rainy season and the onset and cessation dates of the data are quite similar (see Fig. 8).

GN_C with its two rainy seasons has a better agreement among the datasets regarding the absolute values while the relative contribution of the rainy season to annual rainfall totals differs more strongly. This is particularly the case for GCM_hist which simulates only one rainy season. Consequently, GCM_hist has nearly the entire annual precipitation in this period although it starts significantly later than CHIRPS. Having this in mind, the behavior of RCM_hist must be pronounced as it is much closer to CHIRPS in all aspects despite the overestimation of the driving models. This reveals an added value of the RCMs compared to their driving data. The underestimation of RCM_eval might be caused by the shorter second rainy season.

The applied temporal rainy season mask of CHIRPS has a strong effect on the relative contributions in HRN where the first rainy season of CHIRPS and RCM_hist are not overlapping. As a consequence, RCM_hist has a significantly lower contribution of RTOT_rs to RTOT. However, this shows that the overestimation of RTOT is caused by too much precipitation during the second rainy season leading to RTOT_rs showing the same amount as CHIRPS. The relative contribution's underestimation of RCM_eval – despite its ability to simulate the temporal characteristics of both rainy seasons – reveals that there is a lack of water during this important period. In contrast, GCM_hist shows a high ability to reproduce CHIRPS.

Subsequently, we focus on the number of rainy days per year (Rd), the rainy days during the rainy season (Rd_rs), and their relative contributions to the total annual number of days with rainfall. A common observation is that CHIRPS is consistently showing the lowest number of Rd that peak in less than 50% of the rainy days in ERA5Land. As RTOT in CHIRPS is in a comparable range with the other data, this means that the precipitation intensity of CHIRPS is generally higher – at least compared to ERA-Interim but well within the range of observations over Africa (Dosio et al. 2021a). Nevertheless, the relative contribution is comparable with ERA5Land leading to an agreement between the two datasets in this regard. Considering SAH_W, all datasets have comparable contributions with RCM_eval and RCM_hist being in between CHIRPS and ERA5Land but showing higher and lower absolute values, respectively. In GN_C, the absolute amounts of Rd lie between CHIRPS and ERA5Land while the relative contributions are overestimated compared with the reference data. Here, RCM_eval is closest to the reference data. While GCM_hist shows a strong overestimation of Rd_rs compared to Rd, RCM_hist performs somewhat better. As this is also the case in SAH_W we argue that RCM_hist is adding value compared to its forcing data in West Africa.

The most complex situation prevails in HRN as the reference as well as the model data differ the most in that subregion. Regarding the absolute amount, the model ensembles

are closer to ERA5Land with the lowest values for RCM_eval. GCM_hist is closest to ERA5Land regarding the absolute amount and to both reference datasets regarding the proportions. Both RCM ensembles are underestimating the relative contributions. However, RCM_eval is closer to the reference data while RCM_hist shows a strong underestimation which originates from applying the temporal rainy season mask of CHIRPS. This behavior is similar to RTOT. Having this in mind, both indices are strongly overestimated during the original RCM_hist rainy season. Consequently, RCM_hist is not able to represent the considered precipitation characteristics in HRN which is caused by the interaction of RCM behavior and the forcing. The RCM behavior is visible in the general underestimation of the relative contributions, the forcing effect can be deduced from the delayed onset of the rainy season in RCM_hist where the behavior of GCM_hist is enhanced.

We further examine the occurrence of wet (CWD_rs) and dry (CDD_rs) spells during the rainy season. The behavior of the datasets generally follows the already described patterns – CWD_rs is related to Rd_rs and CDD_rs behaves vice versa. In Fig. 10, the focus lies on the maximum length and the number of such dry spells, defined as at least five days without precipitation, in the subregions SAH_W, GN_C, and HRN. As it is also the case for Fig. 9, the general quality of the datasets depends on their ability to represent the rainy season mask of CHIRPS adequately. Further, it should be mentioned that the y-axes in Fig. 11 are not the same for the subregions to prevail readability as the values differ strongly between the subregions.

In SAH_W, ERA5Land and the RCM ensembles represent CDD_rs and nCDD_rs from CHIRPS well. ERA5Land and RCM_eval tend to have slightly more and longer dry spells while RCM_hist behaves vice versa. In contrast, GCM_hist shows much drier conditions as represented by these two indices with CDD_rs being twice as high compared to CHIRPS. Thus, the largest difference between the examined datasets is represented by GCM_hist and RCM_hist which is not the case for the earlier studied rainy season indices where the two ensembles showed a similar behavior. Thus, RCM_hist is able to reduce the bias of CDD_rs and nCDD_rs present in its forcing data in SAH_W.

Focusing on GN_C, CHIRPS represents the dataset with the longest dry conditions during the rainy season. This is in line with Rd in Fig. 9 where rainfall is more seldom but does not result in a reduction of the total precipitation amount in CHIRPS meaning that precipitation events are more intense. ERA5Land shows the strongest deviation from CHIRPS with nearly no dry spell and the shortest CDD_rs. This also is in line with the results of Fig. 9. The model ensembles are between these reference datasets with GCM_hist being closest to CDD_rs and RCM_eval being closest to nCDD_rs of CHIRPS, respectively. In total, one can argue that the models generally are good in representing dry spells during the rainy season in GN_C. In comparison to SAH_W, the nCDD_rs is solely slightly lower in GN_C. CDD_rs is of a similar amount in CHIRPS and RCM_hist but doubles in GCM_hist and RCM_eval and is even the threefold in ERA5Land. This behavior underlines the differences in the rainy season masks of the second rainy season as represented by the other datasets compared to CHIRPS and, thus, highlights the complexity of precipitation processes in the region as well as.

When looking at the y-axes of HRN, CDD_rs is way longer than in the other subregions. While CHIRPS and ERA5Land show approximately 25 days as maximum duration, RCM_eval is around 30 days. The strongest difference is represented in GCM_hist with CDD_rs of 60

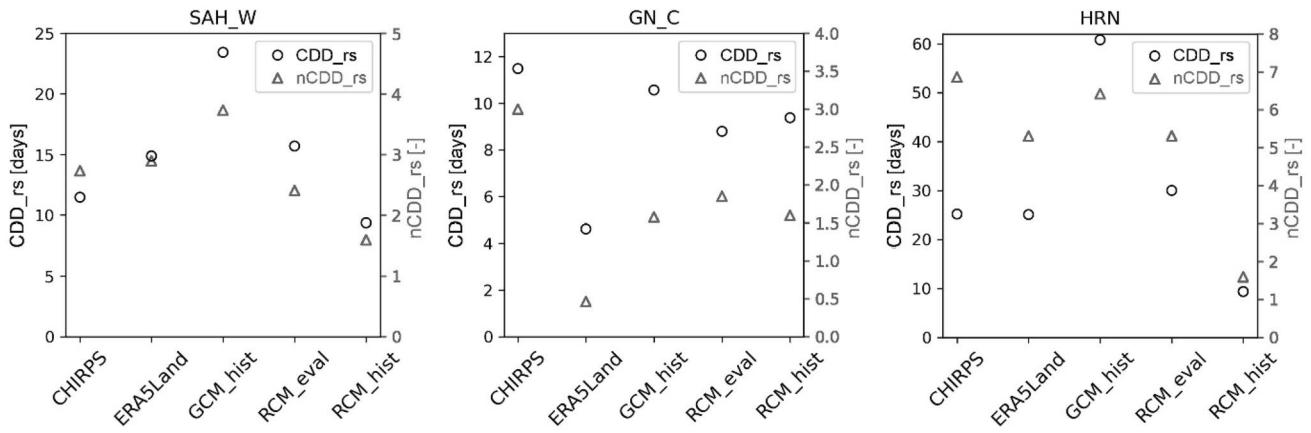

Fig. 10 Spatial climatology (1981–2010) of the maximum number of consecutive dry days during the rainy season (CDD_rs, circle, left y-axis) and the number of periods of consecutive dry days during the rainy season (nCDD_rs, triangle, right y-axis) for the three subregions SAH_W, GN_C, and HRN. The value range of the y-axes differ between the subregions. The temporal rainy season mask of CHIRPS is applied to all datasets

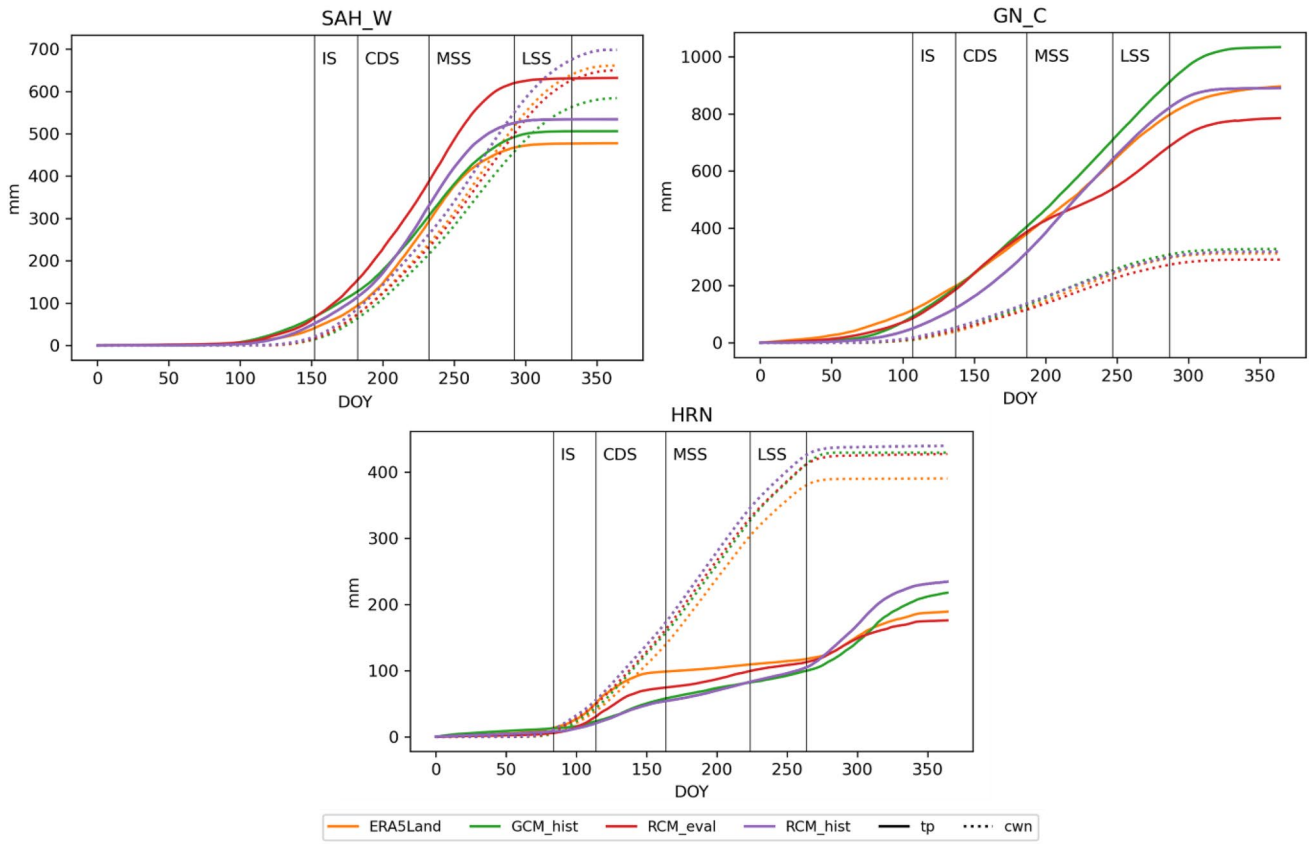

Fig. 11 Cumulative sum of the climatology of daily precipitation (tp) and crop water need (CWN) over the crop stages of maize (grain, long) in the three subregions SAH_W, GN_C, and HRN for ERA5Land and the model ensembles. The vertical lines represent the regional mean of the four crop-specific stages. The skill scores of the variables' temporal evolution are shown in Supplementary 4

days. In contrast, RCM_hist shows the shortest duration of 10 days. However, this is caused by the generally bad representation of the first rainy season in HRN (Fig. 8). Further, the differences between the forcing GCMs and the RCMs are highlighted again. Regarding nCDD_rs, CHIRPS has the highest number of dry spells followed by GCM_hist. Having Figs. 8 and 9 in mind, this is not intuitive as the rainy season overlap of GCM_hist with CHIRPS is worse than of ERA5Land and RCM_eval and Rd in GCM_hist is way higher than in CHIRPS. Additionally, the reduction of nCDD_rs by RCM_hist is notable and is mainly caused by the short overlap of the first rainy season.

From the indices we can conclude that the overlap of the rainy season masks is the dominant factor determining whether a dataset is representing the indices calculated from CHIRPS well or not. A further aspect is the temporal distribution of precipitation during the rainy season as it is highlighted by Rd and the consideration of dry spells. Here, it is shown that RCM_hist is able to change the boundary conditions from its forcing GCMs making these two ensembles the most diverging ones in SAH_W and HRN.

The investigation of the change of these indices in the future is an important and interesting aspect that is planned for the follow-up study.

3.4 Representation of agricultural indices

To assess the representation of the three agricultural indices CWN, IR, and WA by the model ensembles, we solely consider ERA5Land as reference because temperature and radiation are necessary to calculate CWN and WA, respectively (see Sect. 2.2.1). All three indices are based on four stages (IS – Initial Stage, CDS – Crop Development Stage, MSS – Mid Season Stage, and LSS – Late Season Stage) whose individual length and crop factor depends on the respective crop considered (Table 4). The initial stage begins with the onset of the first rainy season for which we consider the CHIRPS-mask for all four examined datasets to be temporally consistent. However, the higher uncertainty of the first rainy season's onset among the datasets might have a strong effect on the planting days when considering the datasets individually. The cessation date of the rainy season or the existence of a second rainy season is irrelevant for the three

agricultural indices as the plant processes are solely considered from the onset onwards.

Figure 11 shows the daily cumulative sum of the climatology (1981–2010) of precipitation (tp) and CWN for the field mean of the three subregions studied previously. Additionally, the four crop-specific stages affecting the crop factor Kc (see Sect. 2.2.3) are included as vertical lines. In SAH_W, tp is higher than CWN during the first two stages and until the second half of MSS meaning that there is enough water to match the crop's need. ERA5Land is the first dataset where the need exceeds the precipitation followed by RCM_hist and GCM_hist in the LSS. Solely RCM_eval simulates higher tp than CWN for all four crop stages. Hence, the historical simulations are representing the relation between the two variables better than the evaluation runs. The strong difference between RCM_eval and ERA5Land originates from the overestimation of precipitation as CWN is on a comparable level. While both historical ensembles overestimate precipitation, GCM_hist underestimates and RCM_hist overestimates CWN. Additionally, the temporal occurrence of tp over the year is different in the historical simulations. Here, GCM_hist is generating more precipitation in the beginning due to the earlier onset of the rainy season while RCM_hist has higher values later-on.

At the GN_C, all datasets simulate significantly higher tp than CWN. Nevertheless, the temporal occurrence of the precipitation differs between the models. During the first two stages, GCM_hist and RCM_eval are much closer to ERA5Land than RCM_hist which underestimates precipitation. With the beginning of MSS, RCM_eval shows a decrease of CWN resulting in a consistent underestimation while RCM_hist is quite close to ERA5Land. On the other hand, GCM_hist exhibits consistently higher values than ERA5Land. The different qualities of the models have already been seen in terms of the onset of the rainy seasons and complicates the assessment of climate models.

In HRN, the CWN is solely fulfilled by ERA5Land during IS. The later onset in the historical simulations and the high precipitation amounts during the second rainy season are not able to compensate the early precipitation deficit. RCM_eval is closest to ERA5Land in this regard.

The skill scores validating the temporal evolution over the year (Supplementary 4) do not give a clear idea of which model ensemble has the best representation. Considering CWN, it is clearer that RCMs have an added value compared to GCMs in GN_C and HRN. RCM_eval show a lower KGE and TSS than GCM_hist in SAH_W. Generally, the quality of the models firstly depends on the subregion and only afterwards on advantages of individual ensembles.

Figure 12 displays the spatial patterns of CWN, IR, and WA per day during the four crop stages for ERA5Land. This shows that CWN has the highest values during MSS and in subtropical regions as the higher temperatures compared to the tropics increase the input factor of potential evapotranspiration there. IR is creating a relation between CWN and precipitation by subtracting the latter from CWN. Thus, negative values show no irrigation need while required irrigation is marked by positive values. It becomes clear that irrigation is necessary in large parts of the Sahel as well as in HRN and Southern Africa for all crop stages of maize (grain). Especially in the MSS – when CWN is highest (see Fig. 12), but the ITCZ is already propagating southward and, thus, leading to the cessation of the rainy season in the northern parts of sub-Saharan Africa – a strong irrigation need of the same amount of the CWN is present. The cessation during MSS is also visible in the maps of WA which considers the actual evapotranspiration based on the latent heat flux. The maps of IR reveal the spatial differentiation that was missing in the spatial means. With this information it becomes clear that the higher tp in SAH_W during the early stages originates from the West Coast but does not occur further inland, like in Burkina Faso. Additionally, GN_C has regions where IR is positive. This is true for the early stages as well as for the break between the two rainy seasons in coastal areas.

Figure 13 displays the absolute differences of CWN between the model ensembles and ERA5Land during the four stages in the respective model resolution. It becomes clear that the coarse resolution of the GCMs is not appropriate for such a specific indicator as strong differences with different signs occur in heterogeneous regions like ETP_H or around Lake Victoria. South of the equator, a general underestimation is present while the patterns north of the equator are changing in West Africa and HRN with the growing stages. In contrast to the GCMs, RCM_eval displays a strong overestimation of CWN during all stages in EAF. Southern Africa is in good agreement with ERA5Land in most areas while its northern parts are marked by overestimations In all stages, yet are less prominent during IS and CDS. Too high values in HRN in the beginning are reduced and show a good agreement with ERA5Land towards later growing stages. In West Africa, CWN is generally underestimated during all stages with regional exceptions. For example, GN_C is showing too high values in IS and around Senegal overestimations are present during all stages. The patterns of RCM_hist display a composite of GCM_hist and RCM_eval. This is demonstrated by the underestimation in SAF_W and Southern Africa as well as the patterns in West Africa during CDS and MSS rising from the GCM-forcing. The overestimation in EAF originates from the RCMs' own model physics. Generally, one can observe that there is a different behavior south- and northward of the equator between the RCM

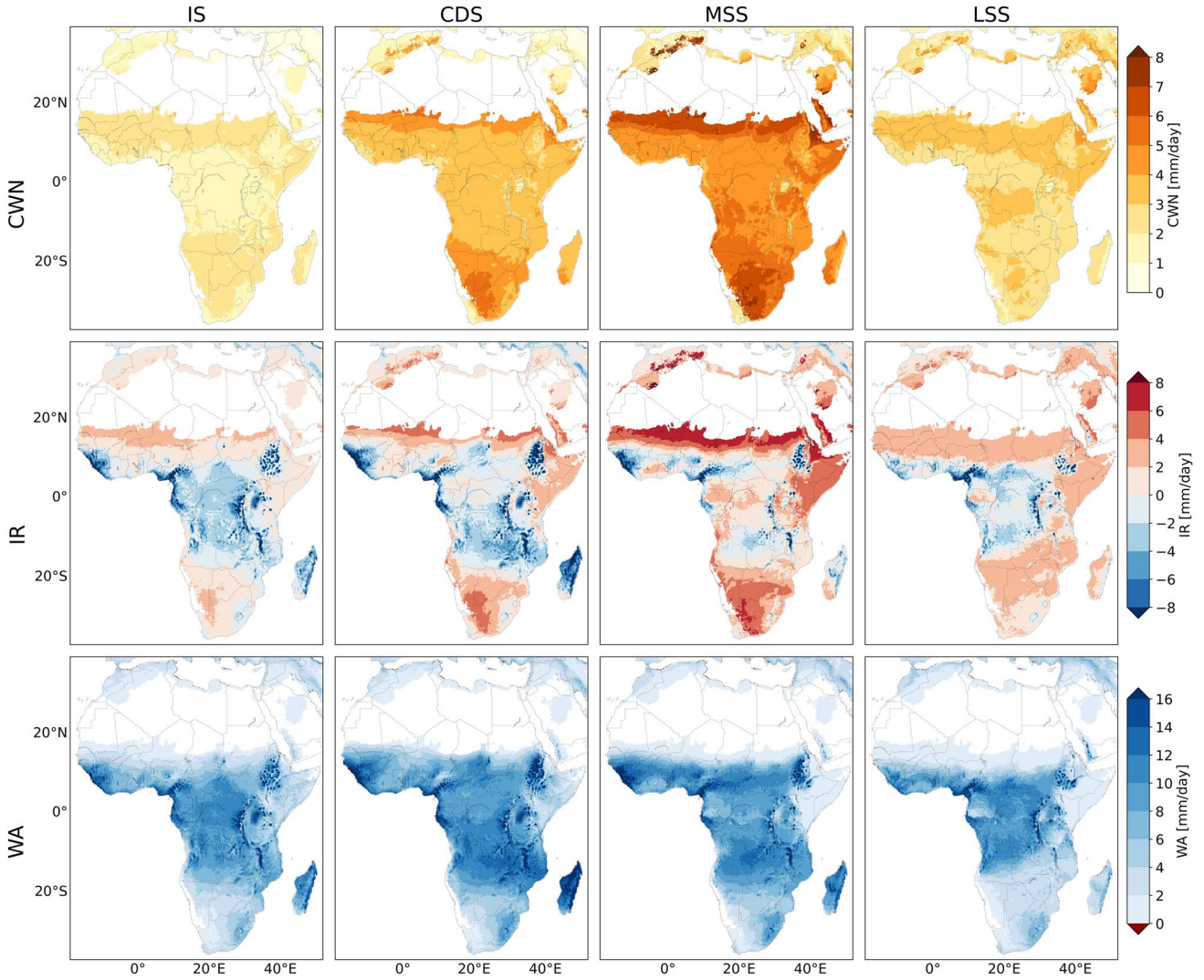

Fig. 12 Absolute values of the climatological (1981–2010) CWN (top row), IR (middle row), and WA (bottom row) for the four crop stages of maize (grain) based on ERA5Land. The indices are represented in mm/day

ensembles: While the values of RCM_hist are lower southwards, they are higher northwards. As a consequence, a reduced overestimation in EAF and SAF is simulated in RCM_hist compared to RCM_eval and the overestimation in Western Africa is more pronounced. These behaviors are another example of the complex and interactive effects of forcing and model physics on the model output in different subregions. However, for HRN e.g., the temporal development of the sign is consistent in all models.

In Fig. 14, the three validation metrics applied to the absolute values of CWN and the four stages are shown for AFR and its subregions. They represent the quality of CWN's spatial distribution as simulated by the model ensembles based on the high resolution of the reference data. AFR is simulated in a comparable quality throughout the stages by all ensembles with GCM_hist being slightly the best due to the lowest MAE. Considering subregions, both RCM-ensembles show lower MAEs and higher KGEs and TSSs in most of them. Additionally, the RCM-ensembles are much closer to each other than RCM_hist to its forcing of GCM_hist in most subregions and throughout most stages. This highlights the general ability in reducing errors introduced by the forcing. The clearest difference between RCMs and GCM_hist prevails in ETP_H with its complex terrain and demonstrates the ability of RCMs to take small-scale land surface characteristics into account. However, not all subregions and stages are consistently better simulated by the RCMs. This especially is the case in EAF but also in some combinations of ensemble, subregion, stage, and applied metric (e.g., SAH_E or GN_C).

Figure 15 shows the absolute differences of IR of the model ensembles to ERA5Land. Generally, the difference

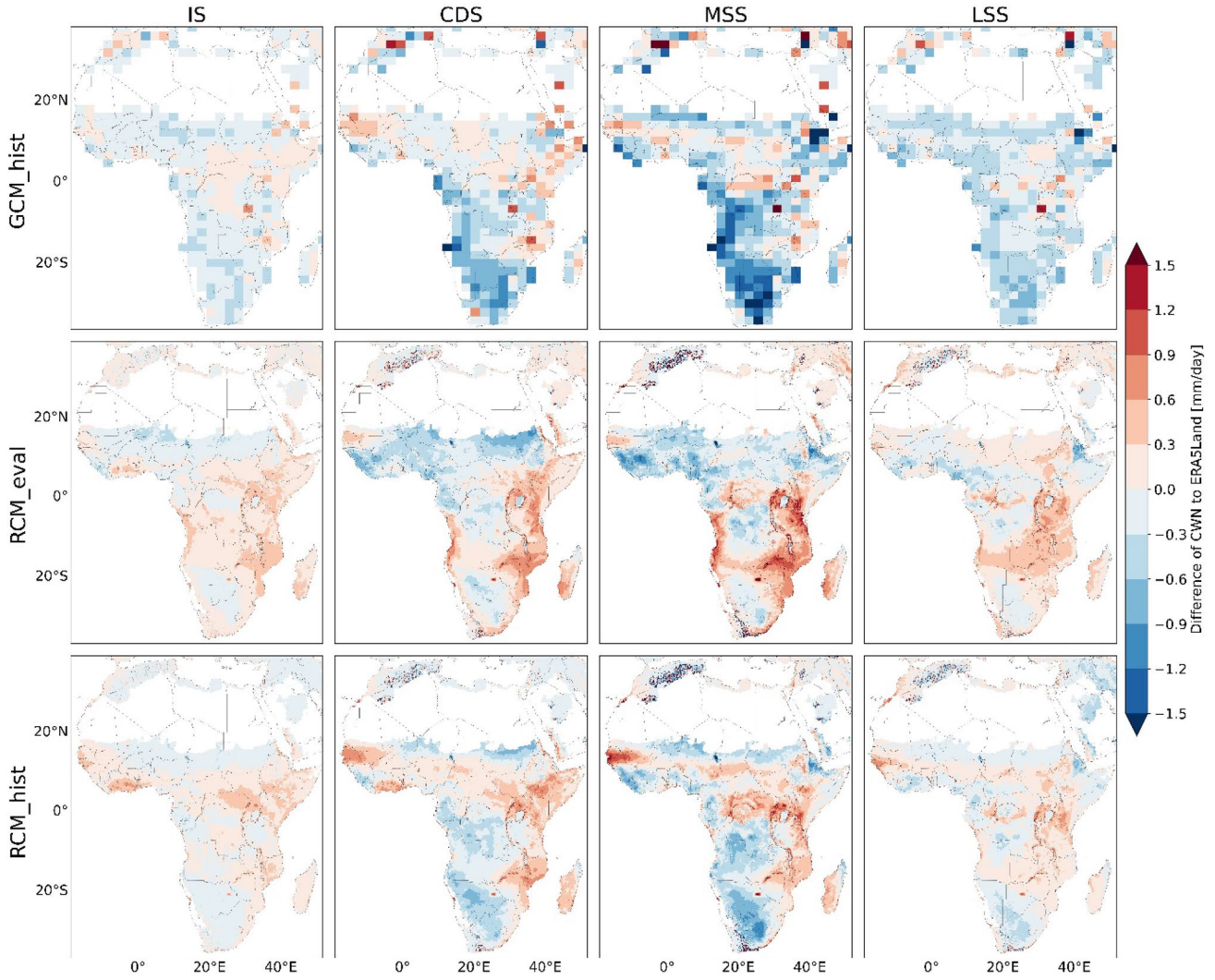

Fig. 13 Absolute differences of the model ensembles' CWN to ERA5Land over the four crop stages in the respective model resolution

patterns are similar to CWN in most areas and stages. This is reasonable as IR is a function of CWN and creates a relation to precipitation. The underestimation by GCM_hist south of the equator originates from the combined effect of underestimating CWN and overestimating precipitation which favors lower IR values. For areas which show no or small IR like CAF, an underestimation is neglectable for irrigation purposes as there either is no need to irrigate or the need lies within the uncertainty. Underestimation becomes more critical in SAF or parts of GN_C as these regions are marked by an irrigation need over all stages in ERA5Land but show consistently lower values in the GCMs. This can result in a misleading interpretation on the level of decision making. The underestimation is present in the Sahel as well but to a lower extent.

In RCM_eval, IR is simulated slightly too low in SAF. Opposing to that and to GCM_hist, the underestimation in the Sahel is stronger in this model ensemble. In CAF and EAF, generally higher values and thus overestimations are present while the Congo basin is marked by various changes in sign over the growing stages. The overestimations might lead to a simulated irrigation need although enough water is available. Underestimations in GN_C during CDS might lead to a wrong decision regarding the general need of irrigation as this shows a high spatio-temporal variability in that area. This is also true for RCM_hist where IR is even lower. Especially the diverging behavior of the RCM-ensembles during MSS reveals large uncertainties. Thus, the differences among the historical RCM-scenario relate to a combination of the forcing data and the RCMs' model physics, depending on the region as it can be seen at the western coast of CAF as well. Like GCM_hist, RCM_hist leads to an underestimation of IR in SAF. However, its extent is lower, hence demonstrating the RCMs' added value.

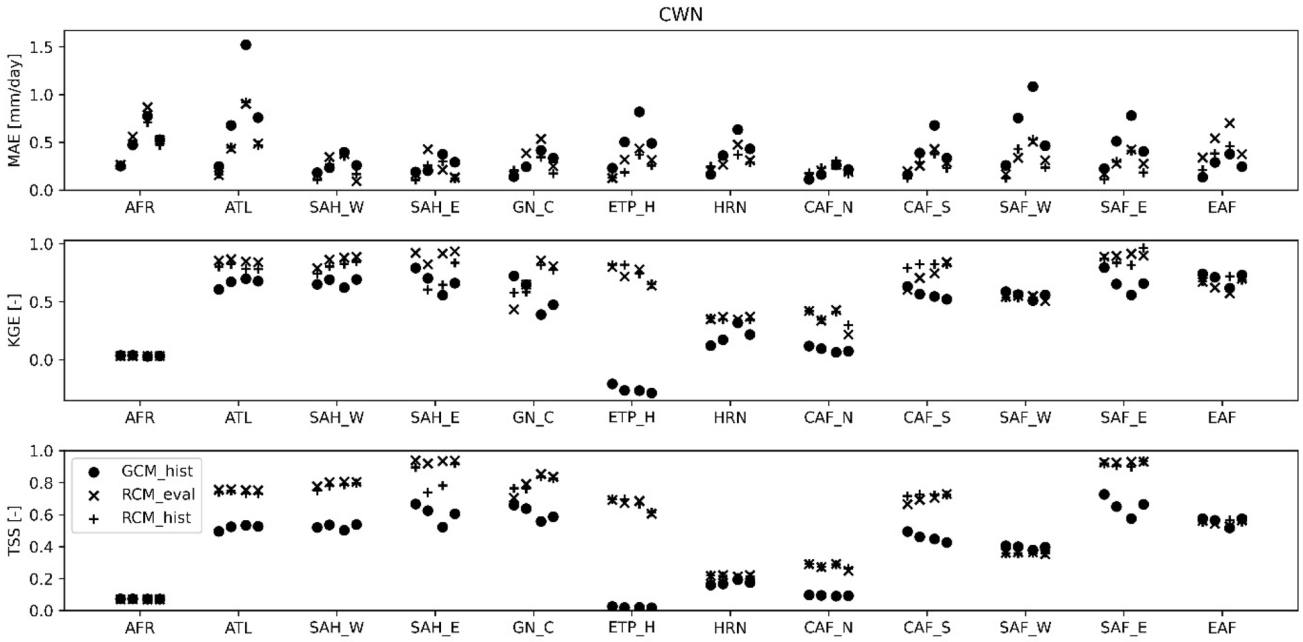

Fig. 14 MAE (top), KGE (middle), and TSS (bottom) of CWN in AFR and its subregions. The stages are represented by the shifted symbols of above each subregion. Their order follows the chronology of the stages and, thus, shows IS, CDS, MSS, and LSS from left to right

WA is strongly and systematically underestimated by all models in all subregions (not shown). The relative underestimation around the equator and in South Africa is around 40% and more while the subtropical regions of Sahel and Southern and East Africa show underestimations of more than 100% which leads to negative WA values. The underestimation results from too high actual evapotranspiration values based on the latent heat flux at the surface, which are subtracted from the precipitation to calculate WA. The overestimation of the latent heat flux over Africa is a known issue in some GCMs, including NorESM1-M (Bentsen et al. 2013) and MPI-ESM (Dosio and Panitz 2016). This bias is also present in RegCM, forced by reanalysis data (Sylla et al. 2012), and CLM, forced by GCMs (Dosio and Panitz 2016), although Dosio and Panitz (2016) showed that the RCM reduces the bias compared to the forcing GCM. This leads to the urgent necessity to improve land surface-atmosphere interactions in climate models.

4 Discussion and conclusion

This study examined the ability of CMIP5 and CORDEX-CORE models to simulate climate indices related to the rainy season and to agricultural issues in Africa during the historical period of 1981–2010. As a precondition, we compared various gridded precipitation datasets. We found a notable spread between individual datasets regarding their precipitation sums, trends, plant phase relation, and over different subregions. The three selected datasets represent the upper range of the spread in most regions which could be caused by their high spatial resolution and the resulting consideration of the topography. These findings are in line with other studies comparing gridded precipitation data over Africa (e.g., Akinsanola and Ogunjobi 2017; Dembélé et al. 2020; Satgé et al. 2020; Dosio et al. 2021b). However, the first three studies focus on African subregions and do not compare a climatological period. Thus, the long-term trends of datasets, e.g., the erroneous positive one of ARC2, and the reduction of the spread among the datasets with time due to the incorporation of more advanced satellite products since the beginning of this millennium are not detected. Dosio et al. (2021b) investigated a climatological period and a wider range of datasets compared to our study. However, the authors include coarser resolved ones but no ERA5-products. For the ERA5-products, we observe a strong negative and erroneous trend ("abrupt transition", Hersbach et al. 2020) in most subregions. The inhomogeneity over time due to the assimilation of different satellite data is a general limitation of reanalyses. Additionally, there are limitations in representing the ITCZ and, thus, tropical precipitation (Lavers et al. 2022) caused by an underrepresentation of the northward propagation of the rainy season in West Africa (Quagraine et al. 2020). A further limitation of reanalyses – which is true for climate model data as well – is that processes on a subgrid scale have to be parameterized (Sun et al. 2018). This can introduce errors which can affect

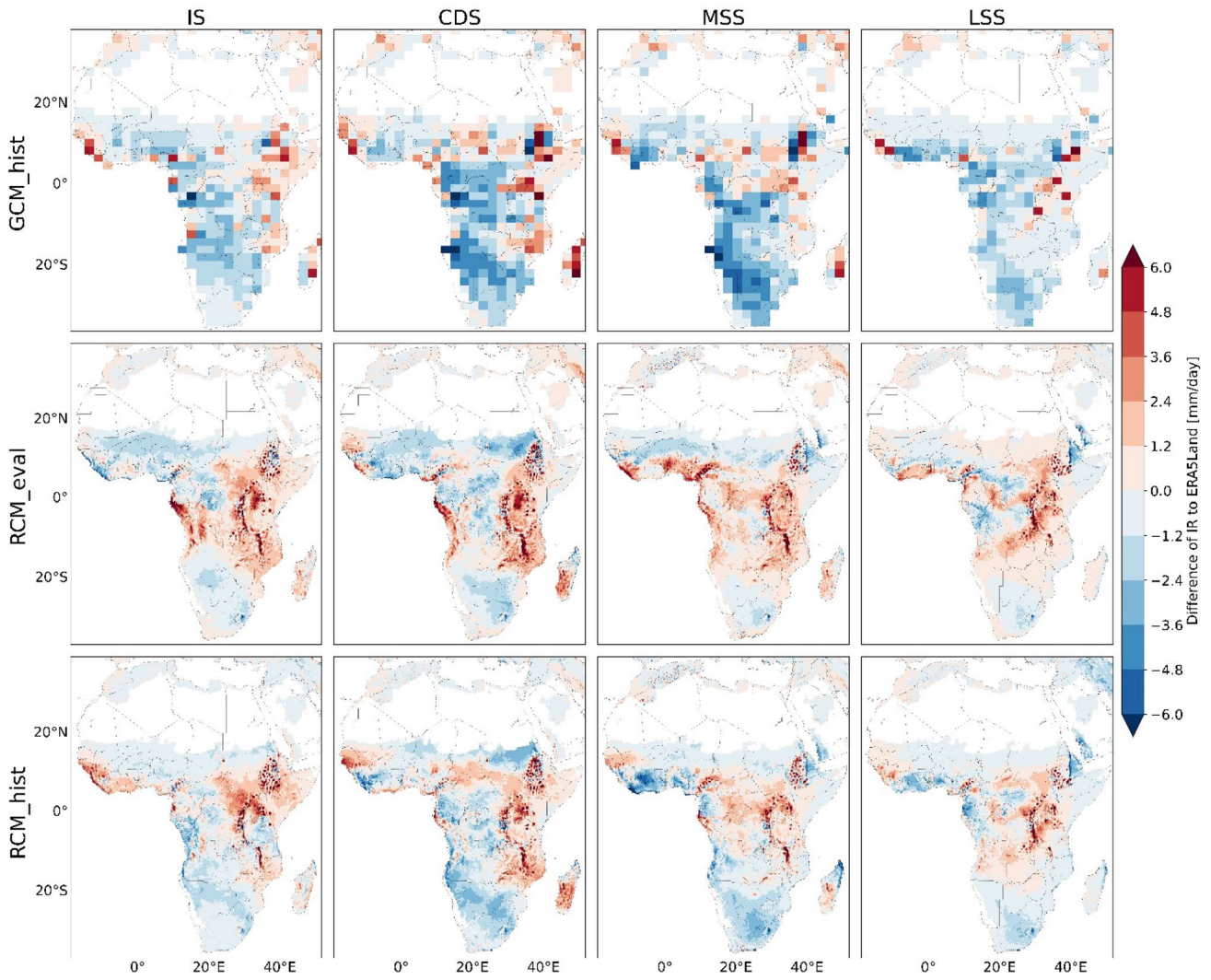

Fig. 15 Same as Fig. 13 but for IR. The skill scores from Fig. 14 but for IR are shown in Supplementary 5

several variables of the system as highlighted for ERA5 by Lavers et al. (2022). However, ERA5Land is assumed to be more adequate than ERA5 due to its consideration of more advanced parameterizations of the land surface (Muñoz-Sabater et al. 2021). Furthermore, the broad range of available variables and its consistency among each other is a strong benefit and makes reanalysis data a valuable source for process understanding and model assessment. Regarding the station data, it has to be highlighted that the decreasing number of in situ measurements in Africa (Bliefernicht et al. 2022; Kaspar et al. 2022) is problematic for the generation and validation of gridded precipitation products (e.g., Prein and Gobiet 2017). Focusing on the spatial distribution, TAMSAT and CHIRPS are representing the occurrence of precipitation well, although there are larger differences in EAF. ERA5Land shows a strong overestimation around coastal areas and mountain ranges which is typical for modeled data.

When it comes to the comparison of the model ensembles with the reference data, we observe that RCM_hist is able to add value to GCM_hist as the higher resolution represents more land surface features and allows for a better representation of related processes (e.g., in ETP_H and West Africa). In this study, this is especially true for parameterized convective (e.g., Rummukainen et al. 2015; Prein et al. 2016) and large-scale, i.e., monsoonal (Dosio et al. 2015, 2019) precipitation. On the one hand, this enables the RCMs to reduce systematic errors introduced by the GCM-forcing (e.g., Di Luca et al. 2016; Sørland et al. 2021; Karypidou et al. 2022). On the other hand, this is not true everywhere to the same extent. In some cases, the forcing overrules the dynamics of the RCMs (e.g., SAH_E and Southern Africa) (e.g., Panitz et al. 2014; Sørland et al. 2021). Generally, the

RCM ensembles are closer to the reference data than the GCMs.

Considering the representation of the rainy season using the method by Dunning et al. (2016), the spatial occurrence of one and two rainy seasons differs strongly among the datasets as it also was shown by Chapman et al. (2020). Thus, we use the mask from CHIRPS to be spatially consistent. Because of this constant assumption, we are not able to examine spatio-temporal changes associated with climate change. However, this limitation plays a minor role considering this study solely focuses on the historical period and does not aim to detect temporal changes. These will be investigated in a follow up study.

We showed that the timing of the rainy season at locations with solely one rainy season is reproduced quite well by the climate models (e.g., SAH_W). At locations with two rainy seasons, the discrepancies among the models as well as among the reference data (not shown) are higher. This increases the uncertainty on the model side which is partly introduced by applying the CHIRPS mask to all datasets. However, we consider CHIRPS to be of high accuracy due to its various data sources. Hence, we argue that, if a dataset does not overlap with that mask, it has limitations in this regard. A good example for this is ERA5Land which does not represent the two rainy seasons in GN_C in an appropriate way. This is the case for all ERA5 products (not shown) although there are improvements compared to ERA-Interim (Quagraine et al. 2020; Nogueira 2020). Focusing on the models, RCM_eval shows a generally good agreement in terms of the onset of the two rainy seasons (e.g., in GN_C and HRN), yet with some delay. The historical simulations are not able to capture two rainy seasons (GN_C) or show a strong delay regarding the onset of the first rainy season (GN_C and HRN). The delay of GCM_hist is increased by RCM_hist so that there is no overlap of the first rainy season with CHIRPS in HRN. Consequently, we state that RCM_hist is not able to reproduce the rainy season adequately in this region. This is caused by the forcing GCMs, since RCM_eval is in line with the reference data.

As a consequence, this inadequacy is also shown in the rainy season-dependent precipitation indices (e.g., RTOT) where the relative contribution of the rainy season to the annual precipitation is strongly underestimated. SAH_W and GN_C show a much better representation by the models. In GN_C, the RCM_hist models are more adequate than their forcing GCMs. The timing of precipitation during the rainy season differs strongly among the datasets and depends on the examined data and subregions as shown for CDD_rs and nCDD_rs. A limitation of our approach is the temporal rainy season mask of CHIRPS which is applied to all datasets but assumed to be a reasonable baseline.

Regarding the agricultural indices, we considered maize (grain) of the long growing season as an exemplary crop. We show that the timing of the precipitation occurrence differs between the models which has effects on the match of the CWN. However, in SAH_W and GN_C the historical simulations are able to represent CWN and tp in an adequate way. This is also found by Gbode et al. (2022). While RCM_eval strongly overestimates precipitation in SAH_W and, hence, overestimates the available amount of water for maize (grain). This is also underlined by the validation metrics. The underestimation of tp in HRN by all model ensembles is more stringent and only ERA5Land provides enough precipitation to fulfill the CWN at least until the middle of the CDS. A view on the absolute values shows that the CWN is highest during MSS. With the cessation of the rainy season during that phase, the irrigation requirement becomes highest compared to the other stages. Especially the Sahel, HRN, and Southern Africa depend on high irrigation amounts for maize in that stage. For the CWN, we compared the models with ERA5Land, revealing that the models' quality differs between subregions and that RCM_hist is partially dominated by the forcing data. Nevertheless, the difference between the GCMs and RCMs is remarkable and highlighted by the validation metrics indicating an added value RCM_hist brings to GCM_hist. While for some subregions GCM_hist is even closer to ERA5Land, it shows strong differences in SAF. In general, the coarse resolution is not able to represent important features and processes in an adequate way, especially in ETP_H, which, in our opinion, should be a strong argument against the consideration of GCMs for agricultural applications.

The considered agricultural indices show some limitations as well. E.g., constant factors like the Kc and the lengths of the growing stages for CWN are very strict and do not consider local variations of plant stages or land surface characteristics like the soil and contained nutrients that have strong effects on crop yield. Per definition, IR has the disadvantage that abundant precipitation events can create strong negative values anticipating that there is no irrigation need. However, at most locations this additional amount of water cannot be stored to bridge subsequent shortcomings. Additionally, most climate models are not able to simulate a potential storage in reservoirs. This can only be integrated by forcing a hydrological model with the climate model's output. Hence, high resolution earth system models considering more land surface and hydrological processes are required. This is also highlighted by the inability of the climate models to simulate WA due to the systematic overestimation of the latent heat flux (Sylla et al. 2012; Bentsen et al. 2013; Dosio and Panitz 2016). A further weakness of climate models is the general underestimation of the temperature range (Lindvall and Svensson 2015; Wang and Clow 2020; Top et al. 2021) – defined by the difference between maximum and

minimum temperatures – affecting the potential evapotranspiration considered for the calculation of CWN. Furthermore, aside from climate models, reanalyses like ERA5 show a bias in the surface heat fluxes as well (Martens et al. 2020). This bias also persists in ERA5Land (Muñoz-Sabater et al. 2021). However, they are an indispensable source due to their spatio-temporal and intervariable consistency which is not possible using observational data. While the size of the considered model ensembles should be larger to produce more robust results, it is currently limited by the available and usable data in the framework of CORDEX-CORE.

Despite the ability of the models to represent the investigated indices in most situations, we want to underline the uncertainties of the model assessment in this study and for Africa in general:

(1) The spread of reference data.
(2) The coarse resolution of GCMs does not consider land surface features and processes.
(3) The RCMs' dependence on the forcing data.
(4) The uncertainty added by different parameterizations of individual models.
(5) The spatially heterogeneous pattern of model biases and quality.

Taking these limitations and biases into account, it is now possible to analyze the indices' behavior and spatio-temporal changes under climate change considering different emission scenarios. Moreover, a more detailed analysis of the rainy season's representation in individual datasets is important because this is a prerequisite for the assessment of indices relevant for agricultural issues. For the rainy season representation within climate models, a process-based investigation of precipitation is important to understand and reduce individual model biases, like it has been done by Tamoffo et al. (2022, 2023) for Central and Western Africa, respectively. However, this is beyond the scope of the current study. Our results have shown that the RCMs are able to represent most indices well depending on the respective subregion. Nevertheless, before using the simulated indices for impact models and studies, a thorough bias correction is still required to reduce the systematic model biases (e.g., Dieng et al. 2022; Steininger et al. 2023) and avoid the transmission of these biases and potentially misleading conclusions on adaptational needs in the agricultural sector. Furthermore and despite the added value provided by state-of-the-art RCMs, additional model development is necessary since hydrological, biological, and oceanic processes are underrepresented in RCMs although they play key roles in the climate system. Therefore, their incorporation would lead to improved climate simulations and more reliable input to impact and adaptation studies (e.g., Guimberteau et al. 2014; Zhang et al. 2019; Drüke et al. 2021; Weber et al. 2023a).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-023-06956-8.

Acknowledgements We thank the ECMWF for providing the ERA5-products, the ESGF and DKRZ for providing the climate model data, the NOAA NCEP CPC FEWS for providing ARC2, the CHC for providing CHIRPS, the DWD for providing GPCC, the NOAA NCEI for providing PERSIANN-CDR, and the University of Reading for providing TAMSAT. We further thank four anonymous reviewers for their valuable comments and Judith N. Claassen for her input.

Author contributions Conceptualization: DA, KZ, IEG, TW, VOA, SBT, HP. Methodology: DA, KZ, IEG, TW. Data and software preparation: DA, KZ, IEG, TW. Formal analysis, investigation, visualization: DA, KZ. Writing - original draft: DA, KZ. Writing - review and editing: DA, KZ, IEG, TW, VOA, SBT, HP. Funding acquisition: VOA, SBT, HP.

Funding Open Access funding enabled and organized by Projekt DEAL. This study has been conducted in the frame of the research project WASCAL WRAP2.0: LANDSURF (FKZ: 01LG2080A) funded by the German Federal Ministry for Education and Research (BMBF) via the project carrier at the German Aerospace Center (DLR Projektträger).

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abiodun BJ, Makhanya N, Petja B et al (2019) Future projection of droughts over major river basins in Southern Africa at specific global warming levels. Theor Appl Climatol 137:1785–1799. https://doi.org/10.1007/s00704-018-2693-0

Ahmadalipour A, Moradkhani H, Castelletti A, Magliocca N (2019) Future drought risk in Africa: integrating vulnerability, climate change, and population growth. Sci Total Environ 662:672–686. https://doi.org/10.1016/j.scitotenv.2019.01.278

Akinsanola AA, Ogunjobi KO (2017) Evaluation of present-day rainfall simulations over West Africa in CORDEX regional climate

models. Environ Earth Sci 76:1–20. https://doi.org/10.1007/s12665-017-6691-9
Akinseye FM, Agele SO, Traore PCS et al (2016) Evaluation of the onset and length of growing season to define planting date—'a case study for Mali (West Africa).' Theor Appl Climatol 124:973–983. https://doi.org/10.1007/s00704-015-1460-8
Ali M, Mubarak S (2017) Effective rainfall calculation methods for field crops: an overview, analysis and New formulation. Asian Res J Agric 7:1–12. https://doi.org/10.9734/arja/2017/36812
Allen R, Pereira L, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements—FAO irrigation and drainage paper 56. FAO Irrig Drain Paper 56:48
Ashouri H, Hsu K-L, Sorooshian S et al (2015) PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull Am Meteorol Soc 96:69–83. https://doi.org/10.1175/BAMS-D-13-00068.1
Awoye OHR, Pollinger F, Agbossou EK, Paeth H (2017) Dynamical-statistical projections of the climate change impact on agricultural production in Benin by means of a cross-validated linear model combined with bayesian statistics. Agric for Meteorol 234–235:80–94. https://doi.org/10.1016/j.agrformet.2016.12.010
Ayugi B, Tan G, Gnitou GT et al (2020) Historical evaluations and simulations of precipitation over East Africa from Rossby centre regional climate model. Atmos Res 232:104705. https://doi.org/10.1016/j.atmosres.2019.104705
Ayugi B, Zhihong J, Zhu H et al (2021) Comparison of CMIP6 and CMIP5 models in simulating mean and extreme precipitation over East Africa. Int J Climatol 41:6474–6496. https://doi.org/10.1002/joc.7207
Beltran-Peña A, D'Odorico P (2022) Future food security in Africa under climate change. Earth's Future 10:e2022EF002651. https://doi.org/10.1029/2022EF002651
Bentsen M, Bethke I, Debernard JB et al (2013) The norwegian Earth System Model, NorESM1-M – part 1: description and basic evaluation of the physical climate. Geosci Model Dev 6:687–720. https://doi.org/10.5194/gmd-6-687-2013
Bliefernicht J, Salack S, Waongo M et al (2022) Towards a historical precipitation database for West Africa: overview, quality control and harmonization. Int J Climatol 42:4001–4023. https://doi.org/10.1002/joc.7467
Bombardi RJ, Moron V, Goodnight JS (2020) Detection, variability, and predictability of monsoon onset and withdrawal dates: a review. Int J Climatol 40:641–667. https://doi.org/10.1002/joc.6264
Bonetti S, Sutanudjaja EH, Mabhaudhi T et al (2022) Climate change impacts on water sustainability of south african crop production. Environ Res Lett 17:084017. https://doi.org/10.1088/1748-9326/ac80cf
Boogaard H, Schubert J, De Wit A et al (2022) Agrometeorological indicators from 1979 to present derived from reanalysis. In: Copernicus Clim Chang Serv Clim Data Store (CDS). https://cds.climate.copernicus.eu/cdsapp#!/dataset/ . https://doi.org/10.24381/cds.6c68c9bb?tab=overview . Accessed 25 Nov 2022
Cairns JE, Hellin J, Sonder K et al (2013) Adapting maize production to climate change in sub-saharan Africa. Food Secur 5:345–360. https://doi.org/10.1007/s12571-013-0256-x
Casanueva A, Herrera S, Iturbide M et al (2020) Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. Atmos Sci Lett 21:e978. https://doi.org/10.1002/asl.978
Chapman S, Birch E, Pope C et al (2020) Impact of climate change on crop suitability in sub-saharan Africa in parameterized and convection-permitting regional climate models. Environ Res Lett 15:094086. https://doi.org/10.1088/1748-9326/ab9daf
Ciarlo JM, Coppola E, Fantini A et al (2021) A new spatially distributed added value index for regional climate models: the EURO–CORDEX and the CORDEX–CORE highest resolution ensembles. Clim Dyn 57:1403–1424
Cleland J (2013) World population growth; past, present and future. Environ Resour Econ 55:543–554. https://doi.org/10.1007/s10640-013-9675-6
Coppola E, Raffaele F, Giorgi F et al (2021) Climate hazard indices projections based on CORDEX-CORE, CMIP5 and CMIP6 ensemble. Clim Dyn 57:1293–1383. https://doi.org/10.1007/s00382-021-05640-z
Degefu MA, Bewket W, Amha Y (2022) Evaluating performance of 20 global and quasi-global precipitation products in representing drought events in Ethiopia I: visual and correlation analysis. Weather Clim Extrem 35:100416. https://doi.org/10.1016/j.wace.2022.100416
Dembélé M, Schaefli B, van de Giesen N, Mariéthoz G (2020) Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. Hydrol Earth Syst Sci 24:5379–5409. https://doi.org/10.5194/hess-24-5379-2020
Deutscher W (2021) GPCC full data daily version 2020. https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-daily_v2020_doi_download.html. Accessed 12 Dec 2022
Di Luca A, Argüeso D, Evans JP et al (2016) Quantifying the overall added value of dynamical downscaling and the contribution from different spatial scales. J Geophys Res Atmos 121:1575–1590. https://doi.org/10.1002/2015JD024009
Di Luca A, Pitman AJ, de Elía R (2020) Decomposing temperature extremes errors in CMIP5 and CMIP6 models. Geophys Res Lett 47:e2020GL088031. https://doi.org/10.1029/2020GL088031
Dieng D, Laux P, Smiatek G et al (2018) Performance analysis and projected changes of Agroclimatological Indices across West Africa based on high-resolution Regional Climate Model Simulations. J Geophys Res Atmos 123:7950–7973. https://doi.org/10.1029/2018JD028536
Dieng D, Cannon AJ, Laux P et al (2022) Multivariate bias-correction of high-resolution regional climate change simulations for West Africa: performance and climate change implications. J Geophys Res Atmos 127:e2021JD034836. https://doi.org/10.1029/2021jd034836
Dinku T, Funk C, Peterson P et al (2018) Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q J R Meteorol Soc 144:292–312. https://doi.org/10.1002/qj.3244
Dosio A (2017) Projection of temperature and heat waves for Africa with an ensemble of CORDEX Regional Climate Models. Clim Dyn 49:493–519. https://doi.org/10.1007/s00382-016-3355-5
Dosio A, Panitz HJ (2016) Climate change projections for CORDEX-Africa with COSMO-CLM regional climate model and differences with the driving global climate models. Clim Dyn 46:1599–1625. https://doi.org/10.1007/s00382-015-2664-4
Dosio A, Panitz H-J, Schubert-Frisius M, Lüthi D (2015) Dynamical downscaling of CMIP5 global circulation models over CORDEX-Africa with COSMO-CLM: evaluation over the present climate and analysis of the added value. Clim Dyn 44:2637–2661. https://doi.org/10.1007/s00382-014-2262-x
Dosio A, Jones RG, Jack C et al (2019) What can we know about future precipitation in Africa? Robustness, significance and added value of projections from a large ensemble of regional climate models. Clim Dyn 53:5833–5858. https://doi.org/10.1007/s00382-019-04900-3
Dosio A, Turner AG, Tamoffo AT et al (2020) A tale of two futures: contrasting scenarios of future precipitation for West Africa from an ensemble of regional climate models. Environ Res Lett 15:064007. https://doi.org/10.1088/1748-9326/ab7fde
Dosio A, Jury MW, Almazroui M et al (2021a) Projected future daily characteristics of african precipitation based on global (CMIP5, CMIP6) and regional (CORDEX, CORDEX-CORE) climate

models. Clim Dyn 57:3135–3158. https://doi.org/10.1007/s00382-021-05859-w
Dosio A, Pinto I, Lennard C et al (2021b) What can we know about recent past precipitation over Africa? Daily characteristics of african precipitation from a large ensemble of observational products for model evaluation. Earth Space Sci 8:e2020EA001466. https://doi.org/10.1029/2020EA001466
Drüke M, von Bloh W, Petri S et al (2021) CM2Mc-LPJmL v1.0: biophysical coupling of a process-based dynamic vegetation model with managed land to a general circulation model. Geosci Model Dev 14:4117–4141
Du Y, Wang D, Zhu J et al (2022) Comprehensive assessment of CMIP5 and CMIP6 models in simulating and projecting precipitation over the global land. Int J Climatol Early View. https://doi.org/10.1002/joc.7616
Dunning CM, Black ECL, Allan RP (2016) The onset and cessation of seasonal rainfall over Africa. J Geophys Res Atmos 121:11405–11424. https://doi.org/10.1002/2016JD025428
Dunning CM, Black E, Allan RP (2018) Later wet seasons with more intense rainfall over Africa under future climate change. J Clim 31:9719–9738. https://doi.org/10.1175/JCLI-D-18-0102.1
Elkouk A, El Abidine El Morjani Z, Pokhrel Y et al (2021) Multi-model ensemble projections of soil moisture drought over North Africa and the Sahel region under 1.5, 2, and 3°C global warming. Clim Change 167:52. https://doi.org/10.1007/s10584-021-03202-0
Fant C, Gebretsadik Y, McCluskey A, Strzepek K (2015) An uncertainty approach to assessment of climate change impacts on the Zambezi River Basin. Clim Change 130:35–48. https://doi.org/10.1007/s10584-014-1314-x
Ferijal T, Batelaan O, Shanafield M (2021) Spatial and temporal variation in rainy season droughts in the indonesian Maritime Continent. J Hydrol 603:126999. https://doi.org/10.1016/j.jhydrol.2021.126999
Fotso–Nguemo TC, Vondou DA, Diallo I et al (2022) Potential impact of 1.5, 2 and 3°C global warming levels on heat and discomfort indices changes over Central Africa. Sci Total Environ 804:150099. https://doi.org/10.1016/j.scitotenv.2021.150099
Funk C, Peterson P, Landsfeld M et al (2015) The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. Sci Data 2:150066. https://doi.org/10.1038/sdata.2015.66
Gbode IE, Diro GT, Intsiful JD, Dudhia J (2022) Current conditions and projected changes in crop water demand, irrigation requirement, and water availability over West Africa. Atmos (Basel) 13:1155. https://doi.org/10.3390/atmos13071155
Giorgi F, Coppola E, Jacob D et al (2022) The CORDEX-CORE EXP-I initiative: description and highlight results from the initial analysis. Bull Am Meteorol Soc 103:E293–E310. https://doi.org/10.1175/BAMS-D-21-0119.1
Guimberteau M, Ducharne A, Ciais P et al (2014) Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin. Geosci Model Dev 7:1115–1136. https://doi.org/10.5194/gmd-7-1115-2014
Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J Hydrol 377:80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003
Gurara MA, Jilo NB, Tolche AD (2021) Impact of climate change on potential evapotranspiration and crop water requirement in Upper Wabe Bridge watershed, Wabe Shebele River Basin, Ethiopia. J Afr Earth Sci 180:104223. https://doi.org/10.1016/j.jafrearsci.2021.104223
IPCC (2021) Climate Change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York
IPCC (2022) Climate Change 2022: impacts, adaptation and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York
Haile GG, Tang Q, Hosseini-Moghari SM et al (2020) Projected impacts of climate change on drought patterns over East Africa. Earth's Future 8:e2020EF001502. https://doi.org/10.1029/2020EF001502
Hall C, Dawson TP, Macdiarmid JI et al (2017) The impact of population growth and climate change on food security in Africa: looking ahead to 2050. Int J Agric Sustain 15:124–135. https://doi.org/10.1080/14735903.2017.1293929
Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. Appl Eng Agric 1:96–99. https://doi.org/10.13031/2013.26773
Harrison L, Funk C, Peterson P (2019) Identifying changing precipitation extremes in Sub-Saharan Africa with gauge and satellite products. Environ Res Lett 14:085007. https://doi.org/10.1088/1748-9326/ab2cae
Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049. https://doi.org/10.1002/qj.3803
Incoom ABM, Adjei KA, Odai SN et al (2022) Impacts of climate change on crop and irrigation water requirement in the Savannah regions of Ghana. J Water Clim Change 13:3338–3356. https://doi.org/10.2166/wcc.2022.129
Iturbide M, Gutiérrez JM, Alves LM et al (2020) An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets. Earth Syst Sci Data 12:2959–2970. https://doi.org/10.5194/essd-12-2959-2020
Jones MR, Singels A, Ruane AC (2015) Simulated impacts of climate change on water use and yield of irrigated sugarcane in South Africa. Agric Syst 139:260–270. https://doi.org/10.1016/j.agsy.2015.07.007
Karam S, Seidou O, Nagabhatla N et al (2022) Assessing the impacts of climate change on climatic extremes in the Congo River Basin. Clim Change 170:40. https://doi.org/10.1007/s10584-022-03326-x
Karypidou MC, Katragkou E, Sobolowski SP (2022) Precipitation over southern Africa: is there consensus among global climate models (GCMs), regional climate models (RCMs) and observational data? Geosci Model Dev 15:3387–3404. https://doi.org/10.5194/gmd-15-3387-2022
Kaspar F, Andersson A, Ziese M, Hollmann R (2022) Contributions to the improvement of Climate Data availability and quality for Sub-Saharan Africa. Front Clim 3:12. https://doi.org/10.3389/fclim.2021.815043
Knox J, Hess T, Daccache A, Wheeler T (2012) Climate change impacts on crop productivity in Africa and South Asia. Environ Res Lett 7:034032. https://doi.org/10.1088/1748-9326/7/3/034032
Konzmann M, Gerten D, Heinke J (2013) Impacts climatiques selon 19 MCG sur les besoins globaux en irrigation simulés par un modèle d'hydrologie et de végétation. Hydrol Sci J 58:88–105. https://doi.org/10.1080/02626667.2013.746495
Kotir JH (2011) Climate change and variability in Sub-Saharan Africa: a review of current and future trends and impacts on agriculture and food security. Environ Dev Sustain 13:587–605. https://doi.org/10.1007/s10668-010-9278-0
Laux P, Kunstmann H, Bardossy A (2008) Predicting the regional onset of the rainy season in West Africa. Int J Climatol 28:329–342. https://doi.org/10.1002/joc.1542

Lavers DA, Simmons A, Vamborg F, Rodwell MJ (2022) An evaluation of ERA5 precipitation for climate monitoring. Q J R Meteorol Soc. https://doi.org/10.1002/qj.4351

Liebmann B, Bladé I, Kiladis GN et al (2012) Seasonality of african precipitation from 1996 to 2009. J Clim 25:4304–4322. https://doi.org/10.1175/JCLI-D-11-00157.1

Lindvall J, Svensson G (2015) The diurnal temperature range in the CMIP5 models. Clim Dyn 44:405–421. https://doi.org/10.1007/s00382-014-2144-2

Lottering S, Mafongoya P, Lottering R (2021) Drought and its impacts on small-scale farmers in sub-saharan Africa: a review. South Afr Geogr J 103:319–341. https://doi.org/10.1080/03736245.2020.1795914

Maidment RI, Allan RP, Black E (2015) Recent observed and simulated changes in precipitation over Africa. Geophys Res Lett 42:8155–8164. https://doi.org/10.1002/2015GL065765.Received

Maidment RI, Grimes D, Black E et al (2017) A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. Sci Data 4:1–19. https://doi.org/10.1038/sdata.2017.63

Martens B, Schumacher D, Wouters H et al (2020) Evaluating the surface energy partitioning in ERA5. Geosci Model Dev 13:4159–4181. https://doi.org/10.5194/gmd-2019-315

Masih I, Maskey S, Mussá FEF, Trambauer P (2014) A review of droughts on the african continent: a geospatial and long-term perspective. Hydrol Earth Syst Sci 18:3635–3649. https://doi.org/10.5194/hess-18-3635-2014

Mbokodo I, Bopape MJ, Chikoore H et al (2020) Heatwaves in the future warmer climate of South Africa. Atmos (Basel) 11:712. https://doi.org/10.3390/atmos11070712

Meza I, Siebert S, Döll P et al (2020) Global-scale drought risk assessment for agricultural systems. Nat Hazards Earth Syst Sci 20:695–712. https://doi.org/10.5194/nhess-20-695-2020

Muñoz-Sabater J, Dutra E, Agustí-Panareda A et al (2021) ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data 13:4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nikulin G, Jones C, Giorgi F et al (2012) Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. J Clim 25:6057–6078. https://doi.org/10.1175/JCLI-D-11-00375.1

Nogueira M (2020) Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: process-based analysis of systematic and random differences. J Hydrol 583:124632. https://doi.org/10.1016/j.jhydrol.2020.124632

Novella NS, Thiaw WM (2013) African rainfall climatology version 2 for famine early warning systems. J Appl Meteorol Climatol 52:588–606. https://doi.org/10.1175/JAMC-D-11-0238.1

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

Ongoma V, Chen H, Gao C (2019) Evaluation of CMIP5 twentieth century rainfall simulation over the equatorial East Africa. Theor Appl Climatol 135:893–910. https://doi.org/10.1007/s00704-018-2392-x

Paeth H, Hall NMJ, Gaertner MA et al (2011) Progress in regional downscaling of west african precipitation. Atmos Sci Lett 12:75–82. https://doi.org/10.1002/asl.306

Panitz HJ, Dosio A, Büchner M et al (2014) COSMO-CLM (CCLM) climate simulations over CORDEX-Africa domain: analysis of the ERA-Interim driven simulations at 0.44° and 0.22° resolution. Clim Dyn 42:3015–3038. https://doi.org/10.1007/s00382-013-1834-5

Prein AF, Gobiet A (2017) Impacts of uncertainties in european gridded precipitation observations on regional climate analysis. Int J Climatol 37:305–327. https://doi.org/10.1002/joc.4706

Prein AF, Gobiet A, Truhetz H et al (2016) Precipitation in the EURO-CORDEX 0.11◦ and 0.44◦ simulations: high resolution, high benefits? Clim Dyn 46:383–412. https://doi.org/10.1007/s00382-015-2589-y

Quagraine KA, Nkrumah F, Klein C et al (2020) West african summer Monsoon Precipitation variability as represented by Reanalysis Datasets. Climate 8:111. https://doi.org/10.3390/cli8100111

Rolle M, Tamea S, Claps P (2021) ERA5-based global assessment of irrigation requirement and validation. PLoS ONE 16:e0250979. https://doi.org/10.1371/journal.pone.0250979

Rolle M, Tamea S, Claps P (2022) Climate-driven trends in agricultural water requirement: an ERA5-based assessment at daily scale over 50 years. Environ Res Lett 17:044017. https://doi.org/10.1088/1748-9326/ac57e4

Roudier P, Sultan B, Quirion P, Berg A (2011) The impact of future climate change on west african crop yields: what does the recent literature say? Glob Environ Chang 21:1073–1083. https://doi.org/10.1016/j.gloenvcha.2011.04.007

Rummukainen M, Rockel B, Bärring L et al (2015) Twenty-first-century challenges in regional climate modeling. Bull Am Meteorol Soc rummukainen16:135–138

Sambou M-JG, Pohl B, Janicot S et al (2021) Heat waves in spring from Senegal to Sahel: evolution under climate change. Int J Climatol 41:6238–6253. https://doi.org/10.1002/joc.7176

Samuel S, Dosio A, Mphale K et al (2023) Comparison of multimodel ensembles of global and regional climate models projections for extreme precipitation over four major river basins in southern Africa - assessment of the historical simulations. Clim Change 176:57. https://doi.org/10.1007/S10584-023-03530-3

Satgé F, Defrance D, Sultan B et al (2020) Evaluation of 23 gridded precipitation datasets across West Africa. J Hydrol 581:124412. https://doi.org/10.1016/j.jhydrol.2019.124412

Schneider U, Becker A, Finger P et al (2020) GPCC full data monthly product version 2020 at 0.25°: Monthly land-surface precipitation fom rain-gauges built on GTS-based and historical data

Schulzweida U. CDO user guide. 2019; 1–206

Schwalm CR, Glendon S, Duffy PB (2020) RCP8.5 tracks cumulative CO2 emissions. Proc Natl Acad Sci USA 117:19656–19657. https://doi.org/10.1073/PNAS.2007117117

Shew AM, Tack JB, Nalley LL, Chaminuka P (2020) Yield reduction under climate warming varies among wheat cultivars in South Africa. Nat Commun 11:1–9. https://doi.org/10.1038/s41467-020-18317-8

Sillmann J, Kharin VV, Zhang X et al (2013) Climate extremes indices in the CMIP5 multimodel ensemble: part 1. Model evaluation in the present climate. J Geophys Res Atmos 118:1716–1733. https://doi.org/10.1002/jgrd.50203

Sørland SL, Brogli R, Pothapakula PK et al (2021) COSMO-CLM Regional Climate Simulations in the CORDEX framework: a review. Geosci Model Dev 14:5125–5154. https://doi.org/10.5194/gmd-2020-443

Sow M, Diakhaté M, Dixon RD et al (2020) Uncertainties in the annual cycle of rainfall characteristics over West Africa in CMIP5 models. Atmos (Basel) 11:216. https://doi.org/10.3390/atmos11020216

Steininger M, Abel D, Ziegler K et al (2023) ConvMOS: climate model output statistics with deep learning. Data Min Knowl Discov 37:136–166. https://doi.org/10.1007/s10618-022-00877-6

Sultan B, Defrance D, Iizumi T (2019) Evidence of crop production losses in West Africa due to historical global warming in two crop models. Sci Rep 9:1–15. https://doi.org/10.1038/s41598-019-49167-0

Sun Q, Miao C, Duan Q et al (2018) A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev Geophys 56:79–107. https://doi.org/10.1002/2017RG000574

Sylla MB, Giorgi F, Stordal F (2012) Large-scale origins of rainfall and temperature bias in high-resolution simulations over southern Africa. Clim Res 52:193–211. https://doi.org/10.3354/cr01044

Sylla MB, Nikiema PM, Gibba P et al (2016) Climate change over West Africa: recent trends and future projections. In: Yaro JA, Hesselberg J (eds) Adaptation to Climate Change and Variability in Rural West Africa. Springer International Publishing, pp 25–40

Sylla MB, Pal JS, Faye A et al (2018) Climate change to severely impact west african basin scale irrigation in 2°C and 1.5°C global warming scenarios. Sci Rep 8:1–9. https://doi.org/10.1038/s41598-018-32736-0

Tamoffo AT, Amekudzi LK, Weber T et al (2022) Mechanisms of Rainfall Biases in two CORDEX-CORE Regional Climate Models at rainfall peaks over Central Equatorial Africa. J Clim 35:639–668. https://doi.org/10.1175/JCLI-D-21-0487.1

Tamoffo AT, Dosio A, Amekudzi LK, Weber T (2023) Process-oriented evaluation of the west african monsoon system in CORDEX-CORE regional climate models. Clim Dyn 60:3187–3210. https://doi.org/10.1007/s00382-022-06502-y

Tarek M, Brissette F, Arsenault R (2021) Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. Hydrol Earth Syst Sci 25:3331–3350. https://doi.org/10.5194/hess-25-3331-2021

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 106:7183–7192. https://doi.org/10.1029/2000JD900719

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Teixeira EI, Fischer G, Van Velthuizen H et al (2013) Global hot-spots of heat stress on agricultural crops due to climate change. Agric for Meteorol 170:206–215. https://doi.org/10.1016/j.agrformet.2011.09.002

Thomas N, Nigam S (2018) Twentieth-century climate change over Africa: Seasonal hydroclimate trends and sahara desert expansion. J Clim 31:3349–3370. https://doi.org/10.1175/JCLI-D-17-0187.1

Top S, Kotova L, Cruz L, De et al (2021) Evaluation of regional climate models ALARO-0 and REMO2015 at 0.22° resolution over the CORDEX Central Asia domain. Geosci Model Dev 14:1267–1293. https://doi.org/10.5194/gmd-14-1267-2021

van Oort PAJ, Zwart SJ (2018) Impacts of climate change on rice production in Africa and causes of simulated yield changes. Glob Chang Biol 24:1029–1045. https://doi.org/10.1111/gcb.13967

van Vuuren DP, Edmonds J, Kainuma M et al (2011) The representative concentration pathways: an overview. Clim Change 109:5. https://doi.org/10.1007/s10584-011-0148-z

Waha K, Müller C, Rolinski S (2013) Separate and combined effects of temperature and precipitation change on maize yields in sub-saharan Africa for mid- to late-21st century. Glob Planet Change 106:1–12. https://doi.org/10.1016/j.gloplacha.2013.02.009

Wang K, Clow GD (2020) The diurnal temperature range in CMIP6 models: climatology, variability, and evolution. J Clim 33:8261–8279. https://doi.org/10.1175/JCLI-D-19-0897.1

Wang Y, Leung LR, McGregor JL et al (2004) Regional climate modeling: progress, challenges, and prospects. J Meteorol Soc Jpn 82:1599–1628. https://doi.org/10.2151/jmsj.82.1599

Waongo M, Laux P, Kunstmann H (2015) Adaptation to climate change: the impacts of optimized planting dates on attainable maize yields under rainfed conditions in Burkina Faso. Agric For Meteorol 205:23–39. https://doi.org/10.1016/j.agrformet.2015.02.006

Weber T, Haensler A, Rechid D et al (2018) Analyzing regional climate change in Africa in a 1.5, 2, and 3°C global warming world. Earth's Future 6:643–655. https://doi.org/10.1002/2017EF000714

Weber T, Cabos W, Dmitry, Sein V, Jacob D (2023a) Benefits of simulating precipitation characteristics over Africa with a regionally-coupled atmosphere–ocean model. Clim Dyn 60:1079–1102. https://doi.org/10.1007/S00382-022-06329-7

Weber T, Gbode IE, Ziegler K et al (2023b) Project LANDSURF—users' interaction protocol to identify specific climate indicators and end-user needs for the development of a decision support system (DSS). WASCAL WRAP2.0: LANDSURF project. https://doi.org/10.13140/RG.2.2.31247.46243

Wilks DS (2011) Statistical methods in the atmospheric sciences, 3rd edn. Academic Press, Oxford

Zebaze S, Jain S, Salunke P et al (2019) Assessment of CMIP5 multi-model mean for the historical climate of Africa. Atmos Sci Lett 20:e926. https://doi.org/10.1002/asl.926

Zhang X, Alexander L, Hegerl GC et al (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. Wiley Interdiscipl Rev Clim Chang 2:851–870. https://doi.org/10.1002/wcc.147

Zhang Z, Arnault J, Wagner S et al (2019) Impact of lateral Terrestrial Water Flow on land–atmosphere interactions in the Heihe River Basin in China: fully coupled modeling and precipitation recycling analysis. J Geophys Res Atmos 124:8401–8423. https://doi.org/10.1029/2018JD030174

Zhang Z, Kattel GR, Shang Y et al (2023) Steady decline in food self-sufficiency in Africa from 1961 to 2018. Reg Environ Change 232:1–12. https://doi.org/10.1007/S10113-023-02074-7

Zittis G, Hadjinicolaou P, Almazroui M et al (2021) Business-as-usual will lead to super and ultra-extreme heatwaves in the Middle East and North Africa. npj Clim Atmos Sci 4:1–9. https://doi.org/10.1038/s41612-021-00178-7