Understanding CMIP6 Biases in the Representation of the Greater Horn of Africa Long and Short Rains

Kevin Schwarzwald (kschwarzwald@iri.columbia.edu)
Columbia University International Research Institute for Climate and Society https://orcid.org/0000-0001-8309-7124
Lisa Goddard
Columbia University International Research Institute for Climate and Society
Richard Seager
Lamont-Doherty Earth Observatory
Mingfang Ting
Lamont-Doherty Earth Observatory
Kate Marvel
Goddard Institute for Space Studies

Research Article

Keywords: precipitation variability, climate models, Indian Ocean Dipole, Walker Circulation, East Africa

Posted Date: February 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1366281/v1

climatedynamics manuscript No.
(will be inserted by the editor)

Understanding CMIP6 biases in the representation of the Greater Horn of Africa long and short rains

Kevin Schwarzwald · Lisa Goddard · Richard Seager · Mingfang Ting · Kate Marvel

Received: date / Accepted: date

Abstract The societies of the Greater Horn of Africa (GHA) are vulnerable to variability in two climatologically distinct rainy seasons, the March-May 'long' rains and the October-December 'short' rains. Recent trends in both rainy seasons, possibly related to patterns of low-frequency variability, have increased interest in future climate projections from General Circulation Models (GCMs). However, previous generations of GCMs historically have a poor record in simulating the regional hydroclimate. This study conducts a process-based evaluation of simulations of the GHA long and short rains in CMIP6, the latest generation of GCMs. Key biases in CMIP5 models remain or are worsened, including long rains that are too short and weak and short rains that are too long and strong. Model biases are driven by a complex set of related oceanic and atmospheric factors. A too strong climatological zonal sea surface temperature gradient in the Indian Ocean and convection over the GHA that is too deep in particular are connected with erroneously powerful short rains in models. Model mean state biases in the timing of the western Indian Ocean sea surface temperature seasonal cycle are associated with certain GHA rainfall timing biases; this connection is however not replicated in interannual variability within models, suggesting there may be a common driver of both biases. Ocean biases cannot explain rainfall biases on their own; simulations driven by historical SSTs (AMIP runs) often have larger biases than fully coupled runs. A path towards using biases to better understand uncertainty in projections of GHA rainfall is suggested.

Keywords precipitation variability · climate models · Indian Ocean Dipole · Walker Circulation · East Africa

This work is undertaken as part of the Columbia World Project, ACToday, Columbia University in the City of New York. The authors have no relevant financial or non-financial interests to disclose.

Kevin Schwarzwald
International Research Institute for Climate and Society
E-mail: kschwarzwald@iri.columbia.edu

Mathematics Subject Classification (2020) MSC code1 · MSC code2 · more

1 Introduction

The Greater Horn of Africa (GHA), comprising eleven countries in East Africa, is a region of both climatic extremes and related societal vulnerability. It comprises the driest area of the tropics, while its societies are heavily dependent on the rainfall cycle. Around 75% of the population in Ethiopia, Kenya, and Tanzania are smallholder farmers primarily working on rainfed lands (Salami et al 2019; Biazin et al 2012), and around 60% of the Somali population practice pastoralism in arid and semi-arid water-stressed regions (UNDP 2019). Consequently, droughts are often associated with threats to food security – for example, the 2011 East African Drought led to the United Nations declaring a famine in southern Somalia, where 2.8 million people needed 'life-saving assistance' (NASA Earth Observatory 2011).

A notable characteristic of the regional climate is the presence of two distinct rainy seasons in the coastal plains of Ethiopia, Somalia, Kenya, and Tanzania: the stronger 'long' rains, known locally as the *gu'* in Somali or *masika* in Swahili, occur in the boreal spring, and the generally weaker but more variable 'short' rains, known locally as the *deyr* in Somali or *vuli* in Swahili, occur in the boreal fall (these will be referred to as the 'long' and 'short' rains, respectively, throughout this paper). Drought extremes that contribute to famines often result from a mistiming or a complete loss of a rainy season such as during the fall 2010 drought (FEWSNET 2011), in which the 'short' rains largely failed. Conversely, particularly wet seasons can cause destructive flooding, such as during the record 'short' rains associated with the 1997-1998 El Niño, which resulted in over 1,300 deaths and 270,000 displacements in Somalia alone (IRIN 97).

Recent trends in the observational records in both rainy seasons have heightened concerns about the impact of climate change on rainfall variability in the GHA region. Declines in total seasonal rainfall since 1983 have been found in studies examining satellite data, station records, satellite-station hybrid datasets, and in farmer recollections (Diem et al 2014, 2019; Ssentongo et al 2018; Cattani et al 2018; Salerno et al 2019), together with a decrease in the rainy season length, with both later onsets and earlier demises (Wainwright et al 2019).The frequency of 'long' rain droughts seems to have particularly increased since 1998, though this is likely the consequence of natural variability attributable to the Pacific Decadal Oscillation (Lyon 2014).

Consequently, many recent studies have used climate models to project changes in rainfall characteristics under global warming scenarios. Modeling studies predict wetter and more intense 'short' rains (e.g. Dunning et al (2018); Otieno and Anyah (2013); Wainwright et al (2021)) and later and wetter 'long' rains (e.g., Wainwright et al (2021)). These projections are incompatible with recent decreases in rainfall, a 'paradox' likely related to simulations of internal variability in GHA rainfall (e.g. Lyon and Vigaud (2017)) or other modeling deficits.

Climate models are increasingly used to project the impacts of regional climate change into the future (e.g. Hsiang et al (2017); Carleton et al (2019)). In East Africa, recent studies have for example used CMIP5-era models to project the impact of global warming on maize and beans production in Ethiopia (Abera et al 2018; Thornton et al 2010), groundwater resources (Taylor et al 2013), and metrics of fisheries, flood management, urban infrastructure, and urban health (Bornemann et al 2019), among others. Climate model studies are also routinely cited in government documents such as Kenya's National Climate Action Plan Government of Kenya (2018), Ethiopia's National Adaptation Plan (Federal Democratic Republic of Ethiopia 2019), or Somalia's communications to the UN Framework Climate Change Convention (Office of the Prime Minister, the Federal Republic of Somalia 2018).

However, despite their heavy use in both academic and government sources, climate models historically have a poor record in simulating rainfall in East Africa. CMIP5 models have well-known biases in simulating both the strength and the timing of the 'long' and 'short' rains in East Africa. The 'long' rains in CMIP5 models start 19 days later on average than in observations (Dunning et al 2017); the 'long' rains are generally too weak and the 'short' rains too strong in models, leading to the 'short' rains being stronger than the 'long' rains (Yang et al 2014).

A process-based model evaluation is however particularly complex in the GHA due to the many regional and large-scale processes that affect local rainfall. Both the 'long' and 'short' rains in the GHA are strongly dependent on the behavior of the large-scale circulation over the Indian Ocean basin. In its long-term average state, the atmosphere above the Indian Ocean is formed into a zonal overturning circulation referred to in the recent literature as the Indian Ocean Walker Cell or Walker-type circulation due to its similarities with the Pacific Ocean Walker Cell pattern over the Pacific Ocean. The Indian Ocean pattern mirrors its Pacific Ocean counterpart; the climatological circulation involves near-surface westerlies, high-level easterlies, ascent over the eastern Indian Ocean and Indo-Pacific Warm Pool, and descent over the GHA (Nicholson 2017). This descent suppresses convection and is present to a certain extent even during the climatological average 'short' rain period (Nicholson 2017; King et al 2019).

The 'long' and 'short' rains occur during the temporary reprieve of this climatological descent in the 'shoulder' seasons between the summer and winter monsoons. The 'long' rains generally begin in late March or early April as the Arabian High dissipates and the strong surface northerlies of the boreal winter weaken, and end as the Mascarene High intensifies, reversing the low-level meridional geopotential height gradient, and turning the offshore winds southerly as part of the broader Indian Monsoon circulation (Vizy and Cook 2020; Camberlin et al 2010). The 'short' rains generally begin in late September, as these strong southerly winds weaken and reverse once more (Vizy and Cook 2020).

The wet seasons are both characterized by seasonal peaks in offshore sea surface temperatures (SSTs) and rising motion in the atmosphere above the GHA. They feature weak, onshore surface winds bringing warm, wet air onto the GHA. The dry seasons are characterized by seasonal minima in offshore SSTs, large-scale descent through the middle and upper troposphere, and surface winds that are both parallel to the shore and dry (Yang et al 2015a; Nicholson 2017).

This complex system suggests the influence of both oceanic and atmospheric factors; studies tracing the interannual variability of the 'long' and 'short' rains have found corresponding influences from both. This variability is particularly strong in the 'short' rains, which, despite being weaker on average than the 'long' rains, contribute more to the overall interannual precipitation variability in the region (Camberlin and Philippon 2002).

Anomalies representing a strengthening of the mean structure of the Indian Ocean Walker Cell are associated with drier rainy seasons in the GHA and vice-versa. Stronger low-level westerlies are negatively correlated with the strength of the 'short' rains (Nicholson 2017). Conversely, low-level easterlies, often associated with the positive phase of the Indian Ocean Dipole (IOD), a metric of the zonal SST gradient, are often associated with particularly strong 'short' rains (Liebmann et al 2014; Nicholson 2017; Blau and Ha 2020). Mid-tropospheric vertical velocity, corresponding to the descending limb of the Walker Cell and local convection, has also been connected to regional rainfall; for example, models that overestimate the strength of the descending limb tend to be biased dry (King et al 2019) and models that explicitly resolve convection over the GHA reduce timing biases in the seasons (Wainwright et al 2021). The influence of the direction of the high-level zonal winds above the GHA is complex; though weaker easterlies may indicate a weaker Walker Cell (e.g. King et al (2019); Hastenrath et al (2011)), above land they may indicate divergence aloft associated with convective activity in the western Indian Ocean (Camberlin and Philippon 2002; Limbu and Tan 2019).

Given their connection to the interannual variability in the GHA rainy seasons, simulations of the surface SSTs and the Indian Ocean Walker Circulation are therefore logical targets to search for the sources of model biases. Consequently, we develop diagnostic metrics based on two aspects of the oceanic state, the zonal SST gradient represented through the IOD and western Indian Ocean SSTs (WIOSSTs); and two aspects of the atmospheric circulation, zonal winds aloft and ascent over the GHA, to identify these sources.

Warmer WIOSSTs and more positive values of the IOD index are expected to correlate with stronger long and short rains; later peaks of the SST seasonal cycle in both variables are expected to correlate with later peaks in the long and short rains. Given the IOD's connection to interannual variability in the 'short' rains in particular, metrics of the IOD are expected to particularly correlate with metrics of the 'short' rains.

Stronger high-level easterlies above the GHA may be an indicator of the development of a convective center in the western Indian Ocean, particularly in the short rains when the coherence of the Walker Cell is stronger. In this paradigm, stronger easterlies are expected to be correlated with stronger 'short' rains. By a similar argument, zonal winds aloft are expected to have weaker correlations with metrics of the 'long' rains; though Camberlin and Philippon (2002) find westerly anomalies aloft for May (towards the end of the 'long' rains) and easterly anomalies aloft for March-April in years in which the 'long' rains seem to be particularly affected by the ENSO cycle.

Stronger ascent, an indicator of convective activity, is expected to be tightly correlated with both stronger 'long' and 'short' rains, as is later ascent with later rainy seasons; biases in these metrics could diagnose problems with model convection simulations.

CMIP6 models are now available, and offer higher resolutions, more explicitly modeled physical processes, and improvements in key dynamics for the Indian Ocean basin (e.g., Gusain et al (2020)). A necessary but insufficient condition for users of CMIP6 output to be confident in their projections is the models' ability to reproduce key aspects of the climate variability in the historical record related to the task at hand (see e.g., the discussion in Nissan et al (2020)). Since these models will likely be extensively used to create projections of the impacts of climate change on East Africa in the coming years, this paper seeks to understand whether these models accurately represent the characteristics of the seasonal cycles in the double rainy season area of the GHA, and whether they replicate key physical drivers of regional rainfall gleaned from the literature and derived from observations.

The rest of this paper is structured as follows: Section 2 will introduce the daily observational and CMIP6 data used; Section 3 introduces the methodology for calculating seasonal and dynamical metrics. Section 4 will detail issues in CMIP6 representations of GHA rainfall. Sections 5 and 6 will investigate to what extent metrics of the ocean and the atmospheric circulation can explain biases in seasonal characteristics, respectively. Finally, Section 7 summarizes conclusions and charts a path forward for how this information can be used to interpret projections of the rainy seasons in the GHA.

2 Data

Daily data are used throughout this study to accurately characterize the timing of the rainy seasons (see e.g., Camberlin and Okoola (2003)). Not only are rainy seasons often less than two months long, but sub-seasonal variability apparent even in monthly data suggests that higher resolution data are needed to fully resolve the relevant dynamics (e.g., Camberlin and Philippon (2002)).

To cover the longest timeframe included in all observational and modeling data products used, all analysis is conducted over the years 1981-2014 for climatological averages. Analysis for individual years is limited to the period 1981-2013, to account for the demise of the short rains sometimes occurring after the Gregorian New Year.

2.1 Observational data

To characterize precipitation in the Horn of Africa, we use daily rainfall data from the Climate Hazards Infrared Precipitation with Stations (CHIRPS) dataset (Funk et al 2015). CHIRPS combines satellite data from the TRMM satellite with interpolated rain gauge products and an elevation model. Though evaluation is complicated by the lack of a dense rain gauge network in the region (e.g. Dinku (2018)), studies have shown CHIRPS to outperform other commonly used datasets in the GHA; while it overestimates the occurrence of rainfall, rainfall in those extra events tends to be minimal (e.g. Diem et al (2019); Ayehu et al (2018)).

Daily sea surface temperatures (SSTs) from the Daily Optimum Interpolation Sea Surface Temperature (DOISST) record, version 2.1 (Huang et al 2021) are used to construct the ocean metrics. DOISST is a 0.25-degree gridded product blending in situ ship and buoy measurements with satellite-derived estimates from the Advanced Very High Resolution Radiometer (AVHRR). Though the Indian Ocean in DOISST is biased slightly low compared to in situ measurements (e.g., $\sim 0.08°$ C vs. Argo floats in Huang et al (2021)), having gridded daily data allows for a direct comparison to model output.

250 hPa zonal velocity and 250 hPa and 500 hPa vertical pressure velocity from the ERA5 reanalysis product are used to analyze historical circulation patterns (Hersbach et al 2020). Data were downloaded in ERA5's native hourly format and daily averages were taken to obtain daily data.

2.2 Model data

This study examines biases in models from the 6th edition of the Coupled Model Intercomparison Project (CMIP6; Eyring et al (2016)). Compared to the previous generation of climate models (CMIP5), CMIP6 models on average have slightly higher resolution and directly simulate more physical processes.

While comprehensive analyses of the newer generation of models are still being performed, studies have begun to evaluate model behavior in the Indian Ocean region. For example, Gusain et al (2020) showed that CMIP6 models have improved representations of the Indian Monsoon compared to CMIP5 models, which may suggest improvements in the simulation of tropical precipitation generally, and dynamics in the Indian Ocean in particular.

Precipitation, SSTs, zonal velocity at 250 mb, and vertical pressure velocity at 250 hPa and 500 hPa from any CMIP6 model with daily data for that variable (not every model has daily data for each variable, see Table 1) are used.

To isolate the impact of SST biases on the biases in the GHA rainy seasons, daily precipitation data from CMIP6 model runs forced by historical SSTs are used, and referred to as 'AMIP' runs ('atmospheric model intercomparison project') throughout.

To illustrate how information about current biases may be used to partition future model projections, precipitation from model runs using the SSP3 scenario (O'Neill et al 2016), representing high challenges to mitigation and adaptation, are used as well.

3 Methods

3.1 Study area

This study focuses on the area of the GHA that experiences a bimodal rainfall climatology (hereafter referred to as the "bimodal region"). In calculations of seasonal statistics, we consider every land grid cell in observations or models between 32° E and the Indian Ocean and between −3° S and 12.5° N for which the second harmonic is larger than the first harmonic. This region is similar to commonly-used geographic subsets for studies of East African rainfall, see e.g., the regions studied by Wainwright et al (2021) or Yang et al (2014). Some authors use a smaller region centered on Southern Somalia (e.g. Camberlin et al (2010); Liebmann et al (2014)); we show our results are robust to the particular region studied.

Each model is evaluated based on its own reality – i.e., the study area is calculated separately for each model and for observations. Models do differ in the exact geographic area in which a bimodal rainy season is simulated (Figure 1); however, models generally place this region in the coastal plains of Somalia, southeast Ethiopia, and northern Kenya, consistent with observations. The factors causing these differences may be important for understanding model behavior in this region, but are beyond the scope of this paper.

3.2 Seasonal definitions

Throughout this paper, we use the seasonal definitions by Dunning et al (2016) based on inflection points in the cumulative precipitation rate. This method was specifically designed for African regions with bimodal rainy seasons, and is designed to reduce the likelihood of 'false starts' – early-season storms followed by prolonged periods of dryness – that may be particularly damaging to recently-planted crops (Huho et al 2012; Dunning et al 2016). Notably, however, it is derived from the data itself, and therefore may not overlap with local agricultural or pastoral definitions of the seasons. These may emphasize different aspects of the season, other variables such as soil moisture content, or use threshold-based definitions that are easier to measure using local information (e.g., Goddard et al (2010); Lala et al (2020)).

For each grid cell in the study area, the onset and demise of the 1981-2014 climatological rainfall is determined using the Dunning et al (2016) method, as is the onset and demise of the rainy seasons in each individual year from 1981 to 2013 (see Section S1 for full details).

3.3 Seasonal metrics

For each season, seasonal characteristics are calculated based on the onset and demise determined using the methodology detailed above. The 'duration' of each season is defined as the simple difference in days between the onset and demise, and the 'total integrated rainfall' or 'strength' as the total sum of daily rainfall between the onset and demise. The 'peak timing' is the day of peak rainfall, while the 'peak amount' is the amount of rain on that day.

3.4 Circulation variables

We develop diagnostic metrics based on two aspects of the oceanic state - IOD and western Indian Ocean SSTs, and two aspects of the atmospheric circulation - zonal winds aloft and ascent over the GHA.

Connections between statistics of the rainy seasons as defined above and diagnostic statistics of the broader circulation are investigated. Each variable has a similar bimodal seasonal cycle to the rainy seasons (Figures 2, 3). The analysis focuses primarily on two metrics defined independently from the rainy seasons – the day on which the variable peaks, referred to as the 'peak timing,' and the value of the variable on that peak day, referred to as the 'peak amount,' for either the first or second portion of the calendar year. For each metric, this cutoff point between the boreal spring and fall seasons is chosen *ad hoc* to encompass the inflection points for each CMIP6 model and the observations. The boreal spring peak timing and amount values are compared to metrics for the 'long' rains, and the boreal fall values with the metrics for the 'short' rains. All metrics are calculated both as a climatological mean and individually for all years in the sample, after each time series has been smoothed using a Gaussian filter with a 30-day width.

To avoid defining explanatory variables using characteristics of the rainy seasons they may be imperfectly related to, the analysis is limited to variables that peak with a bimodal seasonal cycle. We can therefore use nonparametric variables, such as the peak day or peak value, that are robust to the limits of the rainy seasons. This approach may overlook several key processes, chief among them near-surface zonal winds in the Indian Ocean, which have historically been connected to the rainy seasons (e.g. Hastenrath et al (1993)), but nevertheless allows an analysis of aspects of the primary oceanic and atmospheric dynamics of the region.

Western Indian Ocean SSTs (WIOSSTs) Following the region used in Yang et al (2015a), average SSTs in the western Indian Ocean (referred to as WIOSSTs) are calculated as the average from -10° S to 12° N and 38° E to 55° E. For each year, the day of peak WIOSSTs and the peak WIOSSTs are calculated using daily OISST data, for the days 30 to 250 to compare to the long rains, and 250 to 30 of the following year to compare to the short rains.

Indian Ocean dipole mode index The IOD is characterized by the Dipole Mode Index (DMI) developed by Saji et al (1999) and used e.g. in Lyon (2020). The DMI is the difference between SSTs in the West (-10° S to 10° N, 50°-70° E) and East (-10° S - 0° S, 90°-110° E) Indian Ocean. If the DMI is positive, then SSTs in the western Indian Ocean are higher than in the eastern Indian Ocean. For each year, the day of peak DMI and the peak DMI are calculated using daily OISST data, for the days 30 to 230 to compare to the long rains, and 230 to 30 of the following year to compare to the short rains.

Zonal winds aloft The average 250 hPa zonal velocity above the study area (-3° S to 12.5° N and 32° E to 52° E) is used to characterize the zonal circulation aloft. For each year, the day of peak westerlies and the peak westerly strength are calculated using daily ERA5 data, for the days 30 to 230 to compare to the long rains, and 230 to 30 of the following year to compare to the short rains.

Ascent The average 500 and 250 hPa vertical pressure velocities in the bimodal region are used to characterize mid-level and upper-level ascent, respectively. For each year, the day of peak ascent and the peak vertical velocity using daily ERA5 data, for the days 50 to 250 to compare to the long rains, and 250 to 50 of the following year to compare to the short rains.

3.5 Analysis

The timing and strength of the circulation variables are compared with the timing and strength of the rainy seasons in both models and observations. For the rest of this paper, ‘correlations’ refer to Pearson’s correlation coefficients. First, interannual correlations in observations ρ^{OI} between these circulation metrics and their precipitation counterparts are calculated, which reveals whether these facets of the circulation are associated with characteristics of the rainy seasons in the historical record (see Section S3.1 for a detailed derivation). Significance is reported based on two-sided confidence 95% confidence intervals for correlation calculations.

Interannual correlations between the circulation metrics and their precipitation counterparts $\rho^{MI,mod}$ for each individual model *mod* are then calculated, revealing whether these facets of the circulation are associated with characteristics of the rainy seasons within a given model (Section S3.2). Finally, the correlation between model climatological means of these circulation and rainy season metrics ρ^{MM} is calculated, which gives insight into whether the mean state of the model is associated with the biases in these metrics (Section S3.3).

Whether a model is truly simulating the right processes for the right reasons is a combination of both low biases in variables of interest and good performance at replicating the dynamical factors that affect these variables in the observational record. Relationships robustly mirrored in both models and observations may therefore point to metrics useful for diagnosing model performance.

4 Precipitation biases in CMIP6 models

Previous generations of models tended to begin the ‘long’ rains too late, produce too little rain in the ‘long’ rains, and produce too much rain in the ‘short’ rains (Yang et al 2014; Dunning et al 2017). These biases remain largely unchanged in the CMIP6 generation of models.

4.1 Timing biases

The average model ‘long’ rains across CMIP6 models begin 24 $\pm$ 18 days late (with $\pm$ expressing one standard deviation) compared to the average onset in the study area in CHIRPS data (Figure 4a). This bias is of similar magnitude to biases in CMIP5 (19 ± 13 in Dunning et al (2017)). The bias in the onset of the ‘short’ rains, on the other hand, is minor across models; the ensemble model-year bias is 2 ± 9 days too early.

The peak day of the rainy seasons is also too late in both the ‘long’ and the ‘short’ rains, but more consistently so between rainy seasons than in the onset (19 ± 18 days and 14 ± 13 days, respectively; Figure 4d), with more late outliers during the ‘short’ rains.

Models tend to be late on the demise of both rainy seasons - and similarly so; models that are late on the demise in the ‘long’ rains also tend to be late on the demise of the ‘short’ rains. Given that the demise of the ‘long’ rains has been connected in observations to the onset of the Indian Monsoon (Camberlin et al 2010), a pattern unique to the boreal summer, the robust connection with the demise of the ‘short’ rains before the boreal winter is surprising.

These factors combine to make model ‘long’ rains slightly too short on average, and ‘short’ rains significantly too long on average (Figure 4c), and may be connected to the biases in relative strength of the rainy seasons, since rainy season strength is largely modulated by its *length* rather than average rate (Camberlin et al 2009).

4.2 Strength biases

As in CMIP5 models (Yang et al 2015b), CMIP6 models also overestimate the strength of the ‘short’ rains and underestimate the strength of the ‘long’ rains (Figure 4f). The average ratio of the amount of rain in the ‘short’ rains to the ‘long’ rains in models is 2.0, compared to 0.8 in the observations. Like in CMIP5 models, this discrepancy does come both from an underestimation of the strength of the ‘long’ rains (29 $\pm$ 93 mm too dry) *and* an overestimation of the strength of the ‘short’ rains (129 $\pm$ 152 mm too wet). In the amount of both 'long' and 'short' rains, there is however substantial overlap with the range of observations (Figure 4f).

Models tend to underestimate peak rainfall of both rainy seasons (Figure 4e), which is consistent with existing biases in CMIP3 and CMIP5-generation models (e.g. Sun et al (2015)). However, peak rainfall may more generally be related to the model's treatment of rainfall extremes, which is beyond the scope of this study.

4.3 Resolution

Increased CMIP6 model resolution does not remedy biases in precipitation over East Africa (Akinsanola et al 2021), suggesting that orography is not the primary driver of biases, at least within the resolution range of CMIP6 models ($0.70° - 2.8°$ per grid cell). The rest of this study will therefore focus on ocean-atmosphere dynamic processes in the Indian Ocean Basin alone as sources of rainfall biases in the bimodal region.

5 SST representations

5.1 Expected impact of SSTs

To diagnose the impact of model SST biases on GHA rainfall biases, the relationships between WIOSSTs and the IOD and the GHA rainy seasons are investigated. Given connections found between the interannual variability of SSTs and the GHA rainy seasons in observations, models with WIOSSTs and IODs that are too strong or peak too late may be expected to have rainy seasons that are biased wet and late, or vice-versa.

Like rainfall in the bimodal region, both variables climatologically peak twice a year (see Figure 2 for climatologies, and Figure 3 for composite climatologies relative to the onset of each season), though the average SST peak during the 'short' rains is notably a few weeks after the average end of the season. Since most of the interannual variability of the IOD is concentrated in the boreal fall, analyses have generally focused on its impact on the 'short' rains; however, a west-east temperature gradient generally also forms in the boreal spring, peaking along with the average 'long' rains.

Biases in the IOD and in WIOSSTs may point to errors in different, but related underlying processes. The IOD is closely related to the structure of the Indian Ocean Walker Cell – a positive IOD (warm west, cool east) is generally associated with low pressure in the western Indian Ocean and surface easterly winds that advect warm, moist air onto the GHA. A positive IOD generally involves anomalously positive WIOSSTs; however, several studies have also suggested a role for offshore SSTs in encouraging moisture convergence over central East Africa, regardless of the presence of a dipole event (e.g., Liu et al (2020)).

5.2 SSTs and the rainy seasons

For each variable and each rainy season, six correlations are calculated – the interannual correlation in observations ρ^{OI}, the interannual correlation in an individual model $\rho^{MI,mod}$ for every model *mod* separately, and correlations across model means ρ^{MM} for 'strength' (correlation between peak value of the variable and total rainfall in a season) and the 'timing' of the rainy seasons (correlation between peak timing of the variable and peak timing of the rainy season) (see Section S3 for derivations). The correlations across model means ρ^{MM} are based on climatological values of each metric, while the interannual correlations ρ^{OI} in observations and ρ^{MI} in models are calculated across values for each individual year. Figure 5a-b shows ρ^{OI}, $\rho^{MI,mod}$, and ρ^{MM} between the two diagnostic SST metrics and the strength and timing of the rainy seasons. Correlations are relatively robust to the GHA subset chosen (Figure S5).

5.2.1 Mean state biases in WIOSSTs correlate with mean state biases in model rainy seasons

The average model WIOSSTs tend to peak too late during the long rains and too early during the short rains, in line with the average model onset being too late for the long rains and too early for the short rains; the average model WIOSSTs also peak too high in both seasons (Figure S2). Models whose SSTs peak later on average have rainy seasons that peak later, and models whose peak WIOSSTs are higher have stronger rainy seasons, for both the 'long' and the 'short' rains – i.e., the correlation across model means $\rho^{MM,mod}$ in Figure 5 is high for both timing and amount in the 'short' and 'long' rains (Figure 5a). However, apart from the association between warmer SSTs and stronger 'short' rains, this signal is not mirrored across years in observations as a significant interannual correlation in observations ρ^{OI}, nor is it present across years in most individual models as a significant ρ^{MI}. This combination suggests that while the direct relationship between WIOSSTs and the rainy seasons may be weak, mean-state biases easily visible in the SST seasonal cycle may nevertheless be indicative of common drivers of both SST and GHA rainfall biases. A model that is particularly suggestive of this mean-bias relationship is KIOST-ESM, which has the lowest mean state bias in the timing of the boreal spring SST peak, one of the lowest biases in the strength of the boreal spring SST peak, and one of the lowest biases in the timing and strength of the 'long' rains (Figure S2).

5.2.2 IOD strength biases associated with model short rain biases

Generally, the strength of the IOD – meaning how much warmer the western Indian Ocean is than the eastern Indian Ocean – and the WIOSSTs peak are strongly correlated with the strength of the 'short' rains in all metrics - ρ^{OI},

ρ^{MM}, and ρ^{IM} for many models are generally positive and significant (rightmost columns in Figure 5a-b). The correlation is stronger with the IOD than with WIOSSTs by themselves, in line with previous studies connecting the IOD to the short rains in observations on interannual timescales. The high correlation between the dipole mode index and 'short' rain strength across model years in particular suggests that models on average are reproducing this well-known ocean-atmosphere relationship. One notable outlier to this strong relationship is AWI-ESM-1-1-LR, for which the strength of the IOD is *negatively* correlated with the strength of the short rains (though this is insignificant); AWI-ESM-1-1-LR also has the largest dry bias in the short rains among models studied.

These relationships in both models and observations seem to suggest the use of the IOD as a diagnostic variable for model simulation of processes that affect the strength of the 'short' rains in the bimodal region. In particular, some of the models with the most prominent mean state wet biases in the 'short' rains also systematically create climatological IODs that are too powerful;[1] i.e., the western Indian Ocean is much too warm compared to the eastern Indian Ocean. Models with low mean state IOD strength biases in the boreal fall tend to also have low biases in the strength of the 'short' rains. A notable exception is IPSL-CM6A-LR, which has a low strength bias in the dipole mode index despite overestimating the strength of the 'short' rains by a factor of more than 2, suggesting that this low bias may mask structural errors in the model simulation of the region (Figure S3]).

The strong relationship between model mean-state biases in the IOD and corresponding biases in the rainy seasons are in line with the findings of Hirons and Turner (2018), who show that many CMIP5 models have climatological low-level equatorial *easterlies* in the Indian Ocean instead of observed westerlies and associated zonal SST gradients that are too strong during the 'short' rains; these models subsequently cannot correctly capture the dynamics of moisture advection onto East Africa during IOD events in the boreal fall.

Interestingly, the timing of the dipole mode index in the boreal spring is positively correlated with the timing of the 'long' rains in several models, though it is insignificant in observations. The correlation is once again strongest for model means (i.e., $\rho^{MM} > \rho^{OI}, \rho^{MI}$), further suggesting that the mean state of the SST seasonal cycle may be related to rainfall biases in the bimodal region. However, since the corresponding timing correlation across model means for WIOSSTs is larger, it is possible that this correlation may be capturing the effect of WIOSSTs by themselves, which are generally higher during IOD events and are more frequently connected to 'long' rain variability (e.g., Yang et al (2015a)).

[1] e.g. MRI-ESM2-0, BCC-ESM1, the EC-Earth3 models, SAM0-UNICOM, BCC-CSM2-MR, see models highlighted in green in Figure S2

5.3 Evidence from atmosphere-only runs

Do the mean state correlations imply that the SST biases are the primary driver of rainy season biases or, perhaps, that both SST and rainy season biases are affected by a common driver? To investigate this connection, we take advantage of "AMIP" runs – versions of the studied CMIP6 models that replace their ocean component with historical SSTs. These runs can simulate to a certain extent how the model would behave if it perfectly simulated the ocean, though important atmosphere-ocean feedbacks are removed by prescribing SSTs. We recalculate model biases in the rainy seasons and compare them to biases in the same models run in their fully coupled mode (Figure 6). Forcing CMIP6 models with historical SSTs does not uniformly improve biases; rather, for many models and metrics, biases are *increased*, particularly concerning the peak timing and demises of the GHA rainy seasons (Figure 7), suggesting the relationship between SST and rainfall biases may be more complex.

The average WIOSSTs in a CMIP6 model's coupled run peak too late compared to observations during the long rains, but too early during the short rains (Figure S2); however, forcing models with historical SSTs does not lead to a consistent shift in the timing of the rains, nor a consistent reduction in the magnitude of the bias across metrics of the seasonal cycle. The bias in the demise of the long rains is in fact worsened, from 9 days on average in coupled runs to 30 days in AMIP runs on average (Figure 6b, x axis) and up to a factor of 10 in one model, BCC-CSM2-MR (Figure 7, 'demise' column in left panel). Given that the demise of the long rains is tightly correlated with the start of the Indian Monsoon in Kerala in observations (e.g., Camberlin et al (2010)), this bias may therefore be related to changes in monsoon dynamics brought on by the lack of interactive atmosphere-ocean coupling in AMIP runs. In fact, Yang et al (2015b) show that coupling-induced biases in GHA rainy seasons in CMIP5 models can appear jointly with dry biases in the Indian Monsoon.

Unlike the fully coupled models, which especially overestimate the duration and intensity of the short rains, AMIP runs have long rains that are longer and stronger than the short rains, a reversal of a key bias in CMIP6 models; however, AMIP long rains are now too long and too strong compared to observations. (Figure 6c, f; dots to the right and below the dotted 1:1 line show model- or observation-years in which the long rains are lower than the short rains). These two processes are likely linked; the total rainfall in a season is more modulated by the rainy season's length than the average intensity of rainfall in observations (Wainwright et al 2019). It is important to emphasize however that these improvements in the biases in the strength of the rainy seasons in AMIP runs were the result of *increased* timing biases, and underscores the point that models may produce the right metrics, but for the 'wrong' reasons.

Correlations between the model mean SST peak timing and the model mean rainy season timing in coupled models have suggested a role for mean state biases in the seasonal cycle in the modulation of the rainy seasons (Figure

5). AMIP results are consistent with this interpretation, at least for the onset of the rainy seasons. In particular, in one of the few robust improvements in biases in AMIP runs, coupled models that are most biased in the early year peak of WIOSSTs also tend to have the largest reductions in the bias of the onset of the long rains in their AMIP runs (Figure 8, L panel). A similar pattern is seen during the short rains; in AMIP runs, the SST peak in the second half of the year is pushed back compared to coupled runs (WIOSSTs are biased early; Figure S2), and the onset of the short rains is biased late on average in AMIP runs instead of early in coupled runs.

Furthermore, the coupled models with the largest boreal fall IOD biases have the largest reduction in the strength biases of the short rains in their AMIP runs (Figure 8, R panel). These coupled models have zonal SST gradients that are substantially too strong, and correspondingly tend to produce short rains that are too powerful as well. As a result, particularly unphysical values of the mean state IOD in a model may be a useful diagnostic to determine model skill in simulating the East African short rains.

5.4 Conclusions on ocean-driven biases

SST biases play a role in some, but not all facets of the biases in the rainy seasons in East Africa, in line with the highly coupled nature of the regional dynamics. Model mean correlations between mean state biases in SSTs and rainfall metrics that are not borne out in model interannual correlations suggest certain SST biases are driven by the same underlying patterns that produce erroneous long and short rains in the GHA in models. AMIP runs tend to substantially reduce biases only in limited situations, for example, in models whose climatological IODs are 100-400% too strong during the short rains, or for models whose SST seasonal cycle is particularly out of phase with observations. Instead, many biases are worsened in AMIP runs, implying either that the coupling between the atmosphere and ocean is crucial to the regional dynamics affecting those aspects of the GHA rainy seasons, or that competing atmosphere-ocean biases in different aspects of the model may have fortuitously 'cancelled out' in the fully coupled runs.

These findings are only partially consistent with those of Lyon (2020), which suggested that SST biases are the primary driver of both timing and strength biases of the East African rainy seasons, though that study used only one model which was not present in our sample. They are, however, consistent with previous studies showing that AMIP runs did not substantially fix biases in the rainy seasons in CMIP5, the previous generation of models (Hirons and Turner 2018; King et al 2019).

6 Circulation representations

6.1 Expected impact of circulation biases

Circulation metrics have been found to explain more variability in GHA precipitation than ocean variables (Nicholson 2017), and the moisture budget is affected more by the circulation than the humidity cycle in observations (Yang et al 2015a), making aspects of the circulation useful foci for diagnosing biases in the processes driving the rainy seasons.

The climatological circulation pattern over the Indian Ocean Basin, particularly during the dry seasons, consists of ascent in the East over the Maritime Continent, easterlies aloft, descent over the western Indian Ocean and GHA, and surface westerlies along the equator. Strong descent over the GHA inhibits convection for most of the year.

During the rainy seasons, this pattern reverses around the GHA: there is anomalous ascent over the GHA, anomalous westerlies aloft, and anomalous easterlies close to the GHA coast. The seasonal reversal of the winds aloft and the vertical motion over the GHA roughly track GHA rainfall, exhibiting a clear bimodal structure (Figures 2c-d, 3c-d); both of these are associated with an eastward shift in the descending arm of the Indian Ocean Walker Circulation during the rainy seasons, reducing its ability to suppress convection over the GHA (e.g. King et al (2019); Hastenrath et al (2011)). Surface easterlies are also strongly correlated with the 'short' rains (e.g. ~ 0.85 in Hastenrath et al (1993)) but do not peak during the rainy seasons. Zonal velocity aloft (at 250mb) and vertical velocity over the GHA are therefore examined to diagnose biases in the circulation processes associated with the rainy seasons.

6.2 High-level zonal winds are associated with the strength of short rains

This study primarily examines the relationship between high-level zonal winds and the 'short' rains, since the zonal circulation cell is not as coherent during the 'long' rains and therefore plays a smaller role in interannual variability (Hastenrath et al 2011). The strength of the 'short' rains are significantly negatively correlated with the peak zonal wind value in the second half of the year (Figure 5c); i.e., wetter 'short' rains are associated with stronger easterly anomalies. Strong easterlies directly above the GHA may be related to a reversal of the structure of the Indian Ocean Walker Cell, with a convective center in the Indian Ocean off the coast of the GHA and upper-level divergence, as Limbu and Tan (2019) found in the OND climatology. This relationship is robust across model means as well (significant ρ^{MM}), and is present across years in most models, though it is only significant in 6 models. The one model with a significant positive correlation (a positive ρ^{MI}), i.e., where wetter 'short' rains are associated with weaker easterlies across years, INM-CM4-8, also has the largest wet bias in the 'short' rains. Furthermore, only one model, BCC-CSM2-MR, has strong westerlies during the 'short' rains on average together with a substantial wet bias. Models however simulate the range of peak 250 hPa zonal winds relatively well (Figure S4); biases in the upper-level zonal winds are therefore not a good diagnostic for GHA rainfall biases.

6.3 Models overestimate the depth of short rain convection

Vertical velocity is closely related to convection processes in observations and models. Rainfall in the bimodal region tends to occur when the processes that inhibit convection, such as descent associated with the Walker Cell or the import of cool, dry air leading to strong static stability, weaken (King et al 2019; Hastenrath et al 2011; Yang et al 2015a). Correspondingly, ascent, especially in the mid-troposphere, tends to closely track the development of both rainy seasons (Figures 2d, 3d).

As expected, peak ascent at both 500mb and 250mb is strongly correlated with the strength of both the long and short rains in observations (Figure 5d, red column and S6 for 500 mb), and peak timing of ascent with the timing of the long rains. As with other metrics, the timing of the short rains tends to not be strongly correlated with the timing of ascent in observations or models, though one model is a particularly prominent outlier (CanESM5), for which later onsets are significantly associated with *earlier* peaking of ascent.

Models generally replicate this strong relationship between peak ascent and peak strength of the rainy seasons, both across years in individual models and across mean states in different models (blue dots and bars in Figure 3d). Biases in ascent are therefore expected to translate directly to rainfall biases; this seems particularly relevant in the case of biases in the depth of convection during the ‘short’ rains. In observations, convection during the average ‘short’ rains is much shallower than during the ‘long’ rains and tends to not reach 250 hPa (Figure 9, red bars). The average model, however, produces ascent at 250 hPa during the ‘short’ rains (Figure 9); in particular, models that produce climatological ascent at 250 hPa during the ‘short’ rains are on average 150 mm too wet, compared to 11 mm too dry for those that don’t. Models overestimate ascent in the ‘long’ rains on average as well (left panel in Figure 9), but this discrepancy is weaker.

n other words, the strength biases of the ‘short’ rains may be related to model convection being too deep. In particular, models whose convection is not too deep tend to have ‘short’ rains closer to observed strengths, though even within this group, biases range from 140 mm too dry to 157 mm too wet. This signal is visible in other metrics as well; for example, the same models that are particularly biased in their vertical velocity also tend to be the models producing an IOD that is too powerful (see above in Section 5). Since the strength bias in the short rains is reduced in those models’ AMIP runs, the deep convection in the short rains is likely connected to the same overall structural error in these models that produces too much boreal fall convection in the western Indian Ocean.

7 Conclusions and discussion

In conclusion, models continue to produce poor simulations of the rainy seasons in the GHA bimodal region. As in the CMIP5 generation of models, in CMIP6 the timing of both the 'long' and 'short' rains tends to be late, the 'short' rains tend to be too strong, and the 'long' rains tend to be too weak. These biases decrease confidence in projections of the evolution of future rainfall in the GHA, particularly since many are connected to problems simulating the underlying large-scale processes of the Indian Ocean Basin.

In particular, these biases are correlated with biases in model representations of four metrics of the ocean and atmosphere circulations in the Indian Ocean Basin – western Indian Ocean SSTs, the Indian Ocean dipole mode index, zonal winds aloft above the GHA, and ascent over the bimodal region. Mean state biases in the timing and strength of peak SSTs in the boreal spring and fall are correlated with biases in the timing and strength of the 'long' and 'short' rains, respectively. Most models replicate the observed year-to-year relationship between the dipole mode index and the strength of the 'short' rains, though the average model produces an IOD that is too strong in the boreal fall.

However, most rainy season biases in models are not reduced by fixing ocean biases; timing biases in particular are increased in AMIP runs in many models. Consequently, improvements to the performance of ocean models in GCMs alone may not be sufficient to improve model performance over the GHA. Nevertheless, due to their connection with both the 'long' and 'short' rains, particularly significant mean state biases in the timing of the WIOSST cycle and the strength of the IOD may still be used as diagnostics for general biases in the simulation of the overall seasonal cycle of the basin.

Model biases may therefore be particularly susceptible to issues in the simulation of broader circulation patterns. Peak zonal winds in the boreal fall aloft above the GHA are indeed significantly correlated with the strength of the 'short' rains across years in observations, across model means, and across years in a subset of models, with stronger easterlies or weaker westerlies associated with wetter seasons. This suggests that biased simulation of the Indian Ocean Walker Cell, which is particularly coherent during the 'short' rains, may exacerbate rainfall biases. This is consistent with the findings of King et al (2019) for CMIP5 models, who also highlight the importance of improving Walker Cell dynamics in future modeling efforts.

Finally, ascent over the bimodal region itself, which is predictably connected with the strength of both the 'long' and 'short' rains, is a useful diagnostic of biases in model representations of convection. Despite the average 250 hPa vertical pressure velocity in observations being positive, i.e., descending, models produce high-level ascent on average, that is, convection that is too deep. The models with the biggest ascent bias are also the models with the largest positive bias in the Indian Ocean dipole mode index, suggesting an anomalously strong Bjerknes-type feedback, as had previously been found in CMIP3 and CMIP5 models by Cai and Cowan (2013). The models are also those with the largest strength bias in the 'short' rains.

Like in previous studies, it is easier to identify meaningful diagnostic metrics for the 'short' rains, since these are more strongly coupled to large-scale patterns due to the stronger coherence of the Indian Ocean zonal circulation cell in this season (e.g., Hastenrath et al (2011)). Until a better understanding of the physical processes underlying the dynamics and interannual variability of the 'long' rains is developed, process-based model evaluations will continue to be more difficult to produce for the 'long' rains.

A process-based model evaluation such as this one can be used to diagnose whether models are simulating the rainy seasons correctly for the 'right' reasons. A logical direction for future research would be to determine whether CMIP6 models that replicate observed relationships between the rainy seasons in East Africa and aspects of the atmospheric and ocean circulations produce different projections of future rainfall than those that don't. For example, models that associate stronger IODs or stronger easterlies aloft over the GHA with weaker 'short' rains run counter to robust relationships found in the observations and backed by literature; their projections may be flawed. Similarly, models with particularly large biases in key variables in the historical period, such as the IOD, may produce less trustworthy projections. An example of such a partitioning, based on models with the largest historical bias in the strength of the IOD is shown in Figure 10, which shows changes in rainy season metrics between the historical period and end of century (2066-2098) in SSP370 (see Figure S8 for future values). These models' changes are relatively clustered in the short rains (as would be expected given the increased relevance of the IOD to the short rains), particularly in their onset, demise and total rainy season amount changes. Similar to historical simulations, these models tend to be show the some of the wettest future short rains as well (Figure S1). Though part of this clustering may be due to the fact that these models are not all independent (three of the six highlighted models are variants of the EC-Earth model), they may hint at particularly untrustworthy future outcomes. Further study will be needed to fully interpret these results, and compare them to model partitioning schemes based on different metrics.

Finally, studying the biases in underlying processes is particularly crucial to identifying models that may have a low bias in the rainy seasons despite having an unrealistic simulation of the broader circulation; these models may have the 'right' rainy seasons, but for the 'wrong' reasons.

More generally, studies that use climate model projections to estimate the impact of climate change on society should verify that the models are adept at simulating not just the variables of interest, but the processes that affect them. This is particularly important for rainfall, which is often poorly simulated, and in regions with complex dynamics such as the GHA, where biases in rainy seasons may have many causes.

Data Availability

All rainfall statistics, circulation metrics, and correlations calculated for the research in this paper are available in the "gha_rainfall_cmip6" repository at `https://github.com/ks905383/gha_rainfall_cmip6`. All other data and code is available by request.

Acknowledgements

This work is undertaken as part of the Columbia World Project, ACToday, Columbia University in the City of New York. The authors have no relevant financial or non-financial interests to disclose. We are grateful for helpful feedback by Weston Anderson, Rebecca Herman, and Azhar Ehsan, and for the International Research Institute for Climate and Society (IRI) for institutional and computational support.

Declarations

This work is undertaken as part of the Columbia World Project, ACToday, Columbia University in the City of New York. The authors have no relevant financial or non-financial interests to disclose.

References

Abera K, Crespo O, Seid J, Mequanent F (2018) Simulating the impact of climate change on maize production in Ethiopia, East Africa. Environmental Systems Research 7(1):4, DOI 10.1186/s40068-018-0107-z

Akinsanola AA, Ongoma V, Kooperman GJ (2021) Evaluation of CMIP6 models in simulating the statistics of extreme precipitation over Eastern Africa. Atmospheric Research 254:105,509, DOI 10.1016/j.atmosres.2021.105509

Ayehu G, Tadesse T, Gessesse B, Dinku T (2018) Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia. Drought Mitigation Center Faculty Publications

Biazin B, Sterk G, Temesgen M, Abdulkedir A, Stroosnijder L (2012) Rainwater harvesting and management in rainfed agricultural systems in sub-Saharan Africa - A review. Physics and Chemistry of the Earth 47:139–151, DOI 10.1016/j.pce.2011.08.015

Blau MT, Ha KJ (2020) The Indian Ocean Dipole and its Impact on East African Short Rains in Two CMIP5 Historical Scenarios With and Without Anthropogenic Influence. Journal of Geophysical Research: Atmospheres 125(16):e2020JD033,121, DOI 10.1029/2020JD033121

Bornemann FJ, Rowell DP, Evans B, Lapworth DJ, Lwiza K, Macdonald DM, Marsham JH, Tesfaye K, Ascott MJ, Way C (2019) Future changes and uncertainty in decision-relevant measures of East African climate. Climatic Change 156(3):365–384, DOI 10.1007/s10584-019-02499-2

Cai W, Cowan T (2013) Why is the amplitude of the Indian Ocean Dipole overly large in CMIP3 and CMIP5 climate models? Geophysical Research Letters 40(6):1200–1205, DOI 10.1002/grl.50208

Table 1 Models analyzed in this study. A '-' marks models where we do not have data for a given variable / run. See Table S1 for modeling group information.

	precipitation (experiment)			circulation (variable)			
model	historical	AMIP	SSP370	WIOSST	IOD	u	ω
ACCESS-CM2	X	X	X	X	X	X	X
ACCESS-ESM1-5	X	X	X	X	X	X	X
AWI-ESM-1-1-LR	X	-	X	X	X	X	X
BCC-CSM2-MR	X	X	X	X	X	X	X
BCC-ESM1	X	X	-	X	X	X	X
CESM2	X	X	X	X	X	X	X
CESM2-FV2	X	-	-	X	X	X	X
CESM2-WACCM	X	X	X	X	X	X	X
CESM2-WACCM-FV2	X	X	-	X	X	-	X
CMCC-CM2-HR4	X	X	-	X	X	-	-
CMCC-CM2-SR5	X	X	X	X	X	X	X
CMCC-ESM2	X	-	X	X	X	-	-
CNRM-ESM2-1	X	X	X	-	-	-	-
CanESM5	X	X	X	X	X	X	X
EC-Earth3	X	X	X	X	X	X	X
EC-Earth3-AerChem	X	X	X	X	X	-	-
EC-Earth3-CC	X	X	X	X	X	-	-
EC-Earth3-Veg	X	X	X	X	X	X	X
EC-Earth3-Veg-LR	X	-	X	X	X	X	X
FGOALS-f3-L	X	X	-	X	-	X	X
FGOALS-g3	X	X	X	X	-	X	X
GFDL-CM4	X	X	-	X	X	X	X
GFDL-ESM4	X	X	X	X	X	-	X
IITM-ESM	X	X	X	X	-	-	X
INM-CM4-8	X	X	X	X	-	X	X
INM-CM5-0	X	X	X	X	-	X	X
IPSL-CM5A2-INCA	X	-	X	X	X	-	-
IPSL-CM6A-LR	X	X	X	X	X	X	X
IPSL-CM6A-LR-INCA	X	-	-	X	X	-	-
KACE-1-0-G	X	X	X	X	-	X	X
KIOST-ESM	X	X	-	X	X	-	X
MIROC6	X	X	X	X	X	X	X
MPI-ESM-1-2-HAM	X	X	-	X	X	X	X
MPI-ESM1-2-HR	X	X	X	X	X	X	X
MPI-ESM1-2-LR	X	X	X	X	X	X	X
MRI-ESM2-0	X	X	X	X	X	X	X
NESM3	X	X	-	X	X	X	X
NorCPM1	X	X	-	X	-	-	X
NorESM2-LM	X	-	X	X	X	X	X
NorESM2-MM	X	-	X	X	X	X	X
SAM0-UNICON	X	X	-	X	X	X	X
TaiESM1	X	X	X	X	-	X	X
UKESM1-0-LL	X	-	-	-	-	-	-

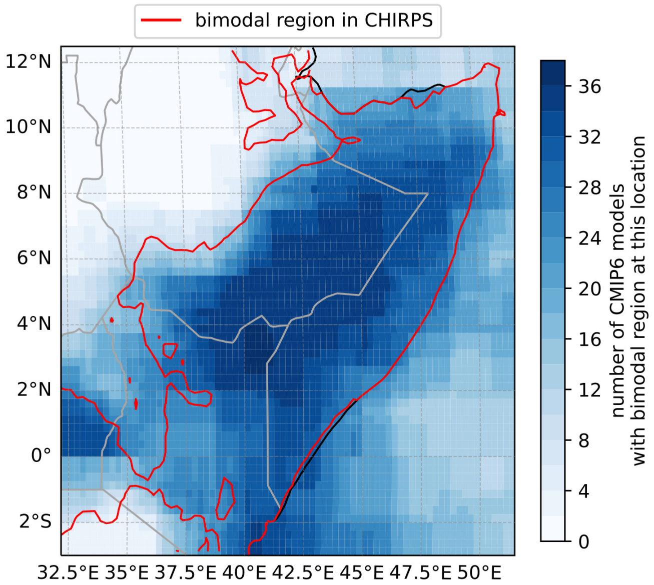

Fig. 1 Study area in CHIRPS observations and CMIP6 models. The red contour shows the area with a bimodal rainy season structure over GHA land in CHIRPS; note that CHIRPS is a land-only data product and rainfall observations over the ocean are not considered in this study. Darker shading means more CMIP6 models have a bimodal rainy season in that location. All models are shown at their native resolutions; grid cells may only partially overlap between models. Most models place the bimodal region along the coastal plains of Somalia, Kenya, and southeastern Ethiopia, consistent with observations. For the remainder of this study, statistics of the rainy seasons (and ascent) are averaged over each data product's bimodal region over land.

Camberlin P, Okoola RE (2003) The onset and cessation of the "long rains" in eastern Africa and their interannual variability. Theoretical and Applied Climatology 75(1):43–54, DOI 10.1007/s00704-002-0721-5

Camberlin P, Philippon N (2002) The East African March–May Rainy Season: Associated Atmospheric Dynamics and Predictability over the 1968–97 Period. Journal of Climate 15(9):1002–1019, DOI 10.1175/1520-0442(2002)015⟨1002:TEAMMR⟩2.0.CO;2

Camberlin P, Moron V, Okoola R, Philippon N, Gitau W (2009) Components of rainy seasons' variability in Equatorial East Africa: Onset, cessation, rainfall frequency and intensity. Theoretical and Applied Climatology 98(3):237–249, DOI 10.1007/s00704-009-0113-1

Camberlin P, Fontaine B, Louvet S, Oettli P, Valimba P (2010) Climate Adjustments over Africa Accompanying the Indian Monsoon Onset. Journal of Climate 23(8):2047–2064, DOI 10.1175/2009JCLI3302.1

Carleton T, Delgado M, Greenstone M, Houser T, Hsiang S, Hultgren A, Jina A, Kopp RE, McCusker K, Nath I, Rising J, Rode A, Seo HK, Simcock J, Viaene A, Yuan J, Zhang AT (2019) Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits. SSRN Scholarly Paper ID 3224365, Social Science Research Network, Rochester, NY, DOI 10.2139/ssrn.3224365

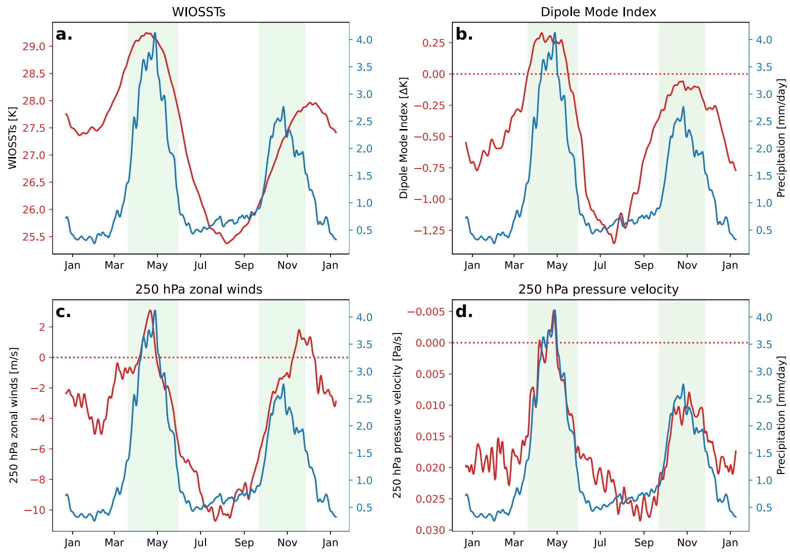

Fig. 2 Rainfall (blue) and key variable (red) climatologies in observations (CHIRPS for rainfall, OISST for SST variables) or reanalysis (ERA5 for circulation variables). Light green shading is the geographical average long (centered on May) and short (centered on October) rainy seasons. Since climatology shows the study area average rainfall, but seasonal onset and demises were calculated using local rainfall before averaging, correspondence between light green shaded area and local rainfall climatologies is not perfect. See Figure 3 for composite climatologies.

Cattani E, Merino A, Guijarro JA, Levizzani V (2018) East Africa Rainfall Trends and Variability 1983–2015 Using Three Long-Term Satellite Products. Remote Sensing 10(6):931, DOI 10.3390/rs10060931

Diem JE, Ryan SJ, Hartter J, Palace MW (2014) Satellite-based rainfall data reveal a recent drying trend in central equatorial Africa. Climatic Change 126(1):263–272, DOI 10.1007/s10584-014-1217-x

Diem JE, Konecky BL, Salerno J, Hartter J (2019) Is equatorial Africa getting wetter or drier? Insights from an evaluation of long-term, satellite-based rainfall estimates for western Uganda. International Journal of Climatology 39(7):3334–3347, DOI 10.1002/joc.6023

Dinku T (2018) Overcoming challenges in the availability and use of climate data in Africa. ICT Update

Dunning CM, Black ECL, Allan RP (2016) The onset and cessation of seasonal rainfall over Africa. Journal of Geophysical Research: Atmospheres 121(19):11,405–11,424, DOI 10.1002/2016JD025428

Dunning CM, Allan RP, Black E (2017) Identification of deficiencies in seasonal rainfall simulated by CMIP5 climate models. Environmental Research Letters 12(11):114,001, DOI 10.1088/1748-9326/aa869e

Dunning CM, Black E, Allan RP (2018) Later Wet Seasons with More Intense Rainfall over Africa under Future Climate Change. Journal of Climate 31(23):9719–9738, DOI 10.1175/JCLI-D-18-0102.1

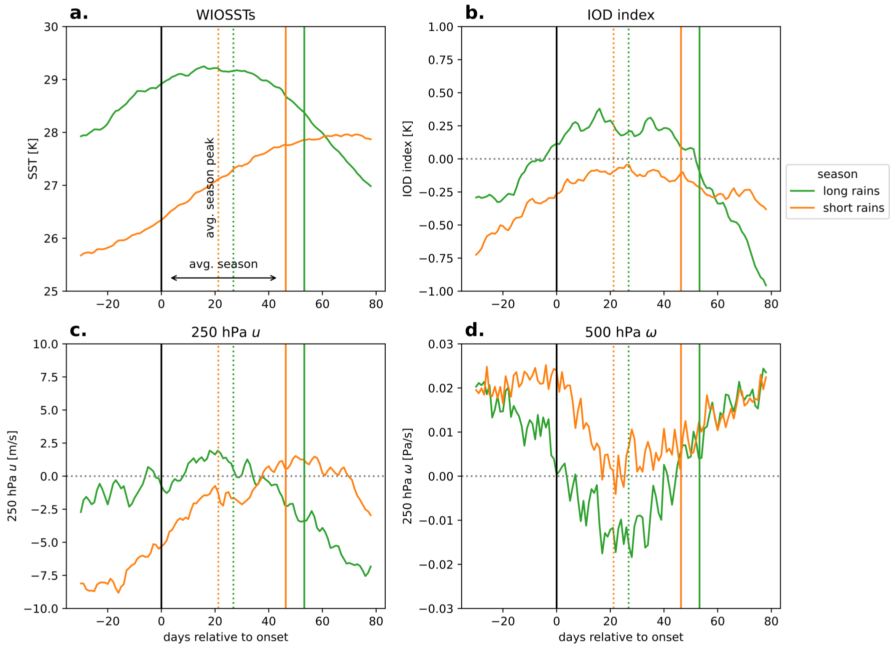

Fig. 3 Seasonal composites of WIOSST, IOD, 250 hPa zonal winds (u), and 250 hPa pressure velocity (ω) observations / reanalysis (top row: OISST, bottom row: ERA5). Values are the average across years relative to CHIRPS seasonal onset (1981-2013); the average peak day of each season is shown in dotted lines and the average end of each season in a solid line. All variables peak roughly around the GHA bimodal rainy seasons, though peaks generally correspond more closely to rainfall peaks during the long rains. See Figure 2 for raw (not composite) climatologies.

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development 9(5):1937–1958, DOI 10.5194/gmd-9-1937-2016

Federal Democratic Republic of Ethiopia (2019) Ethiopia's Climate Resilient Green Economy: National Adaptation Plan. Tech. rep., Addis Ababa, Ethiopia

FEWSNET (2011) Eastern Africa: Drought - Humanitarian Snapshot (as of 24 Jun 2011). Humanitarian Snapshot

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data 2(1):150,066, DOI 10.1038/sdata.2015.66

Goddard L, Aitchellouche Y, Baethgen W, Dettinger M, Graham R, Hayman P, Kadi M, Martínez R, Meinke H (2010) Providing Seasonal-to-Interannual Climate Information for Risk Management and Decision-making. Procedia Environmental Sciences 1:81–101, DOI 10.1016/j.proenv.2010.09.007

Government of Kenya (2018) National Climate Change Action Plan 2018-2022 (NCCAP). Tech. rep., Ministry of Environment and Forestry, Nairobi, Kenya

Gusain A, Ghosh S, Karmakar S (2020) Added value of CMIP6 over CMIP5 models in simulating Indian summer monsoon rainfall. Atmospheric Research 232:104,680, DOI 10.1016/j.atmosres.2019.104680

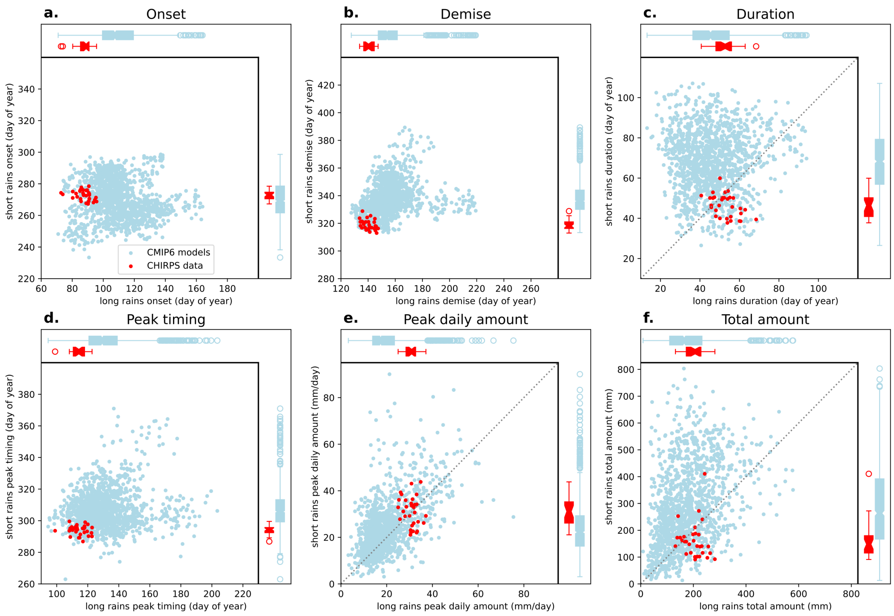

Fig. 4 Key characteristics of the 'long' and 'short' rains in the study region in CMIP6 models (light blue) and CHIRPS observations (red). Each dot shows a model-year (CMIP6) or an observation-year (CHIRPS) between 1981 and 2013. Box plots show the median (notch), 0.25 and 0.75 quartiles (box), up to $1.5 \times IQR$ beyond the 0.25 and 0.75 quartiles (whiskers), and outliers beyond this limit (circles). The range of models is biased versus observations for almost every characteristic, except for the onset of the 'short' rains (panel a, x-axis). Otherwise, models tend to be too late in their demise and peak timing, rain too little on the wettest days, overestimate the length and strength of the 'short' rains, and underestimate the length and strength of the 'long' rains.

Hastenrath S, Nicklis A, Greischar L (1993) Atmospheric-hydrospheric mechanisms of climate anomalies in the western equatorial Indian Ocean. Journal of Geophysical Research: Oceans 98(C11):20,219–20,235, DOI 10.1029/93JC02330

Hastenrath S, Polzin D, Mutai C (2011) Circulation Mechanisms of Kenya Rainfall Anomalies. Journal of Climate 24(2):404–412, DOI 10.1175/2010JCLI3599.1

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, Chiara GD, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146(730):1999–2049, DOI 10.1002/qj.3803

Hirons L, Turner A (2018) The Impact of Indian Ocean Mean-State Biases in Climate Models on the Representation of the East African Short Rains. Journal of Climate 31(16):6611–6631, DOI 10.1175/JCLI-D-17-0804.1

Hsiang S, Kopp R, Jina A, Rising J, Delgado M, Mohan S, Rasmussen DJ, Muir-Wood R, Wilson P, Oppenheimer M, Larsen K, Houser T (2017) Estimating economic damage from climate change in the United States. Science 356(6345):1362–1369, DOI 10.1126/science.aal4369

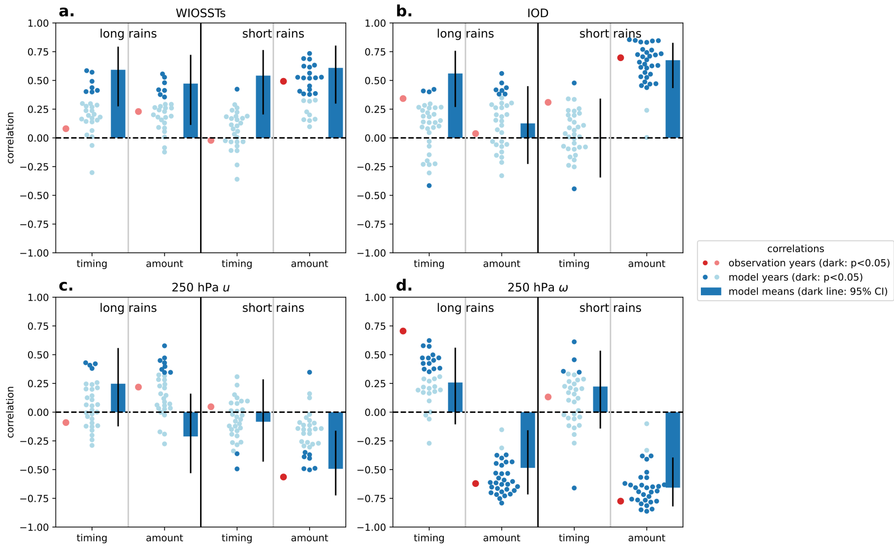

Fig. 5 Correlations between statistics of the GHA rainy seasons and statistics of (clockwise from top) WIOSSTs, the IOD, upper-level pressure velocity, and upper-level zonal winds in models and observations. For each sub-panel, the leftmost column (red dot) shows the correlation between years of the variable and the rainy season in observations ('observation-year' correlation), the center column (blue dots) shows the correlation between years of the variable and the rainy season for each model ('model-year' correlation), and the rightmost column (blue bar) shows the correlation between model means of the variable and the rainy season ('model-means' correlation). Black vertical lines show 95% confidence intervals; for individual models and observations, darker blue dots show significant Pearson's correlation coefficients at the $p < 0.05$ level. For each variable and season, correlations between two sets of statistics are shown: 'timing' means the correlation between the peak day of the rainy season and the peak day of the variable, 'amount' means the correlation between the *total* amount of rain in that season and the *peak* amount of the variable. Correlations are robust to different subsets of the GHA; see Figure S5 for the same calculations over a smaller box centered on southeastern Somalia.

Huang B, Liu C, Banzon V, Freeman E, Graham G, Hankins B, Smith T, Zhang HM (2021) Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. Journal of Climate 34(8):2923–2939, DOI 10.1175/JCLI-D-20-0166.1

Huho JM, Ngaira JKW, Ogindo HO, Masayi N (2012) The changing rainfall pattern and the associated impacts on subsistence agriculture in Laikipia East District, Kenya. Journal of Geography and Regional Planning 5(7):198–206, DOI 10.5897/JGRP12.018

IRIN (97) Central and Eastern Africa: Background brief on floods, 11/25/97. IRIN background brief, UN Department of Humanitarian Affairs, Nairobi, Kenya

King JA, Washington R, Engelstaedter S (2019) Representation of the Indian Ocean Walker circulation in climate models and links to Kenyan rainfall. International Journal of Climatology n/a(n/a), DOI 10.1002/joc.6714

Lala J, Tilahun S, Block P (2020) Predicting Rainy Season Onset in the Ethiopian Highlands for Agricultural Planning. Journal of Hydrometeorology 21(7):1675–1688, DOI 10.1175/JHM-D-20-0058.1

Liebmann B, Hoerling MP, Funk C, Bladé I, Dole RM, Allured D, Quan X, Pegion P, Eischeid JK (2014) Understanding Recent Eastern Horn of Africa Rainfall Variability and Change. Journal of Climate 27(23):8630–8645, DOI 10.1175/JCLI-D-13-00714.1

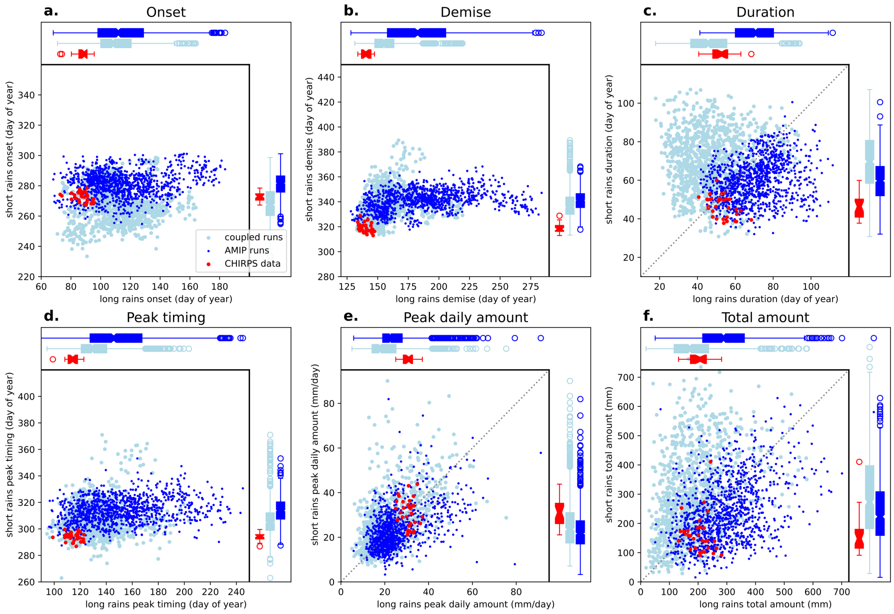

Fig. 6 Key characteristics of the 'long' and 'short' rains in the study region (as in Figure 4), for models with available daily rainfall data from both fully coupled runs (light blue) and runs forced with historical SSTs (dark blue). Coupling doesn't uniformly reduce biases. AMIP runs tend to end the long rains later, leading to an increase in the duration bias, and begin the short rains later than fully coupled runs, leading to a decrease in the duration bias. In both rainy seasons, the late bias in the timing of the rainy season peak is *increased* compared to the fully coupled runs. In line with the changes in duration bias, the average model-year total amount is too strong in the AMIP long rains, but the positive rainfall bias is decreased in the short rains. In line with observations, the AMIP long rains are now stronger than the short rains.

Limbu PTS, Tan G (2019) Relationship between the October–December rainfall in Tanzania and the Walker circulation cell over the Indian Ocean. Meteorologische Zeitschrift pp 453–469, DOI 10.1127/metz/2019/0939

Liu W, Cook KH, Vizy EK (2020) Influence of Indian Ocean SST regionality on the East African short rains. Climate Dynamics 54(11):4991–5011, DOI 10.1007/s00382-020-05265-8

Lyon B (2014) Seasonal Drought in the Greater Horn of Africa and Its Recent Increase during the March–May Long Rains. Journal of Climate 27(21):7953–7975, DOI 10.1175/JCLI-D-13-00459.1

Lyon B (2020) Biases in CMIP5 Sea Surface Temperature and the Annual Cycle of East African Rainfall. Journal of Climate 33(19):8209–8223, DOI 10.1175/JCLI-D-20-0092.1

Lyon B, Vigaud N (2017) Unraveling East Africa's Climate Paradox. In: Climate Extremes, American Geophysical Union (AGU), chap 16, pp 265–281, DOI 10.1002/9781119068020.ch16

NASA Earth Observatory (2011) Severe Drought Causes Famine in East Africa. https://earthobservatory.nasa.gov/images/51411/severe-drought-causes-famine-in-east-africa

Nicholson SE (2017) Climate and climatic variability of rainfall over eastern Africa. Reviews of Geophysics 55(3):590–635, DOI 10.1002/2016RG000544

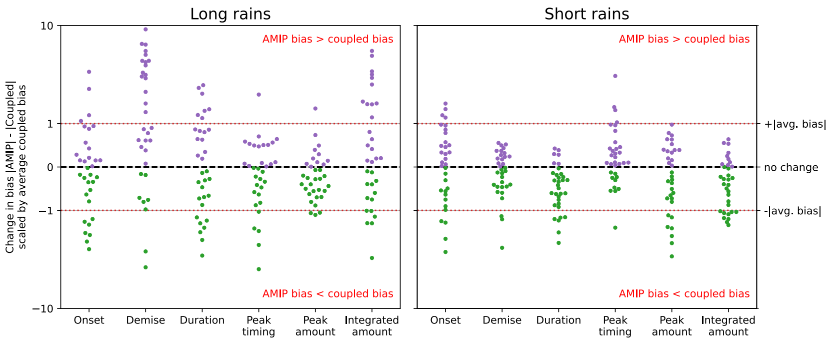

Fig. 7 Change in bias between fully coupled and AMIP runs. Each dot represents the climatological bias difference |AMIP| - |coupled|, scaled by the average climatological bias of the fully coupled runs for the long (L) and short (R) rains. A value of 0 means the AMIP and coupled biases are identical; a value of 1 means the AMIP bias is larger than the coupled bias by an amount equal to the average coupled bias, a value of -1 means the opposite. AMIP models do not uniformly decrease (or increase) biases; the long rain demise bias in particular is worsened in most models.

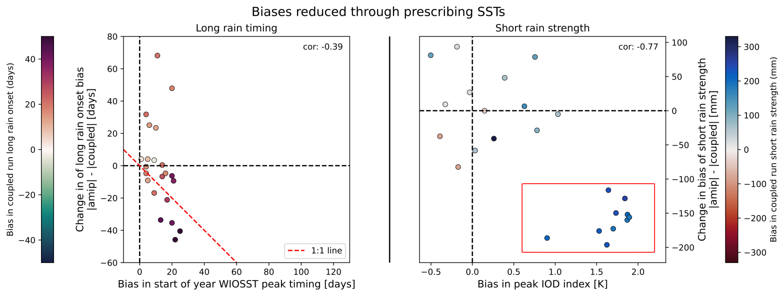

Fig. 8 Examples of metrics in which fully coupled runs tend to have stronger biases than AMIP runs. Points show model means. Y axes represent the change in the absolute bias between AMIP and coupled runs (negative values mean AMIP runs have lower biases) in a given metric; x axes represent the early year WIOSST timing bias (L panel) or the late year IOD strength bias (R panel) in the fully coupled run. Shading shows coupled model bias. Models whose western Indian Ocean SSTs (WIOSSTs) peak the latest compared to observations tend to see the biggest improvements in the late onset bias seen in the most models' long rains (L panel). Similarly, the models with the largest positive IOD biases (the 8 models in the red box in the R panel) show the largest improvements in short rain strength biases when forced with historical SSTs.

Nissan H, Muñoz ÁG, Mason SJ (2020) Targeted model evaluations for climate services: A case study on heat waves in Bangladesh. Climate Risk Management 28:100,213, DOI 10.1016/j.crm.2020.100213

Office of the Prime Minister, the Federal Republic of Somalia (2018) The Initial National Communication for Somalia to the United Nations Framework Convention on Climate Change (UNFCCC). Tech. rep.

O'Neill BC, Tebaldi C, van Vuuren DP, Eyring V, Friedlingstein P, Hurtt G, Knutti R, Kriegler E, Lamarque JF, Lowe J, Meehl GA, Moss R, Riahi K, Sanderson BM (2016) The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geoscientific

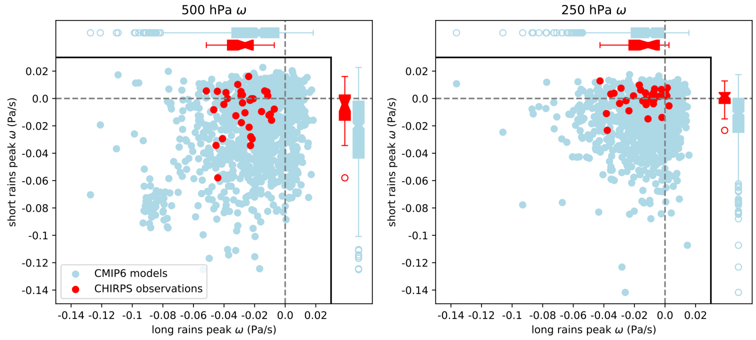

Fig. 9 Peak strength of pressure velocity (ω) over study region in models and observations. Models tend to produce deeper convection than observations in the short rains (vertical axes; model bias in pressure velocity is stronger at 250 hPa, where observations rarely show strong upward motion, than at 500 hPa).

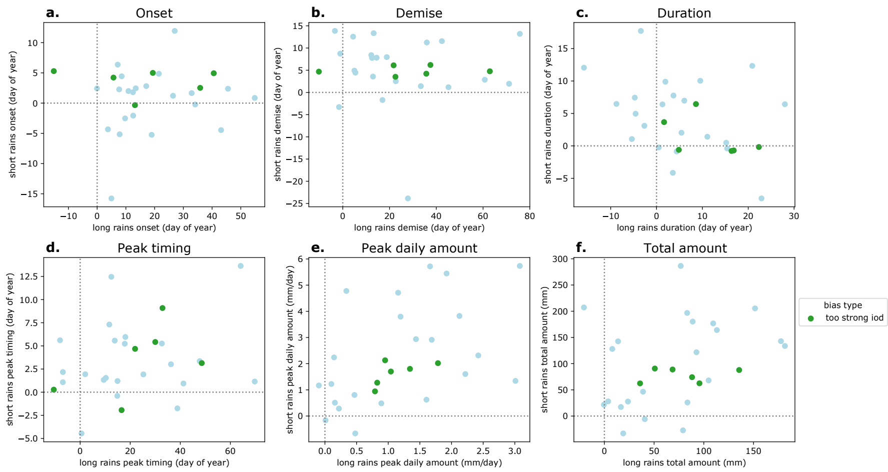

Fig. 10 Changes in GHA rainy season characteristic between models' historical runs (1981-2013) and SSP370 (2066-2098) runs. Green dots highlight models with IOD biases above 1.5 K. Projections of future changes in onset, demise, peak daily amount, and total amount for the short rains in particular seem to be similar across models with particularly biased historical IODs; however, three of these models are variants from the same modeling group (EC-Earth), which may explain the clustering.

Model Development 9(9):3461–3482, DOI 10.5194/gmd-9-3461-2016

Otieno VO, Anyah RO (2013) CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part II: Projected climate. Climate Dynamics 41(7):2099–2113, DOI 10.1007/s00382-013-1694-z

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401(6751):360–363, DOI 10.1038/43854

Salami A, Kamara AB, Brixiova Z (2019) Smallholder Agriculture in East Africa: Trends, Constraints and Opportunities. Working Paper 105, African Development Bank Group

Salerno J, Diem JE, Konecky BL, Hartter J (2019) Recent intensification of the seasonal rainfall cycle in equatorial Africa revealed by farmer perceptions, satellite-based estimates, and ground-based station measurements. Climatic Change 153(1):123–139, DOI 10.1007/s10584-019-02370-4

Ssentongo P, Muwanguzi AJB, Eden U, Sauer T, Bwanga G, Kateregga G, Aribo L, Ojara M, Mugerwa WK, Schiff SJ (2018) Changes in Ugandan Climate Rainfall at the Village and Forest Level. Scientific Reports 8(1):3551, DOI 10.1038/s41598-018-21427-5

Sun Q, Miao C, Duan Q (2015) Comparative analysis of CMIP3 and CMIP5 global climate models for simulating the daily mean, maximum, and minimum temperatures and daily precipitation over China. Journal of Geophysical Research: Atmospheres 120(10):4806–4824, DOI 10.1002/2014JD022994

Taylor RG, Todd MC, Kongola L, Maurice L, Nahozya E, Sanga H, MacDonald AM (2013) Evidence of the dependence of groundwater resources on extreme rainfall in East Africa. Nature Climate Change 3(4):374–378, DOI 10.1038/nclimate1731

Thornton PK, Jones PG, Alagarswamy G, Andresen J, Herrero M (2010) Adapting to climate change: Agricultural system and household impacts in East Africa. Agricultural Systems 103(2):73–82, DOI 10.1016/j.agsy.2009.09.003

UNDP (2019) More than 360,000 farmers and pastoralists are set to benefit from GEF-funded project - Somalia. https://reliefweb.int/report/somalia/more-360000-farmers-and-pastoralists-are-set-benefit-gef-funded-project

Vizy EK, Cook KH (2020) Interannual variability of East African rainfall: Role of seasonal transitions of the low-level cross-equatorial flow. Climate Dynamics 54(11):4563–4587, DOI 10.1007/s00382-020-05244-z

Wainwright CM, Marsham JH, Keane RJ, Rowell DP, Finney DL, Black E, Allan RP (2019) 'Eastern African Paradox' rainfall decline due to shorter not less intense Long Rains. npj Climate and Atmospheric Science 2(1):1–9, DOI 10.1038/s41612-019-0091-7

Wainwright CM, Black E, Allan RP (2021) Future Changes in Wet and Dry Season Characteristics in CMIP5 and CMIP6 Simulations. Journal of Hydrometeorology 22(9):2339–2357, DOI 10.1175/JHM-D-21-0017.1

Yang W, Seager R, Cane MA, Lyon B (2014) The East African Long Rains in Observations and Models. Journal of Climate 27(19):7185–7202, DOI 10.1175/JCLI-D-13-00447.1

Yang W, Seager R, Cane MA, Lyon B (2015a) The Annual Cycle of East African Precipitation. Journal of Climate 28(6):2385–2404, DOI 10.1175/JCLI-D-14-00484.1

Yang W, Seager R, Cane MA, Lyon B (2015b) The Rainfall Annual Cycle Bias over East Africa in CMIP5 Coupled Climate Models. Journal of Climate 28(24):9789–9802, DOI 10.1175/JCLI-D-15-0323.1

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- gharainfallsuppdocs.pdf