Journal Pre-proof

Gauging the performance of CMIP5 historical simulation in reproducing observed gauge rainfall over Kenya

Lucia Mumo, Jinhua Yu

PII: S0169-8095(19)31163-9

DOI: https://doi.org/10.1016/j.atmosres.2019.104808

Reference: ATMOS 104808

To appear in: *Atmospheric Research*

Received date: 5 September 2019

Revised date: 4 November 2019

Accepted date: 19 December 2019

Please cite this article as: L. Mumo and J. Yu, Gauging the performance of CMIP5 historical simulation in reproducing observed gauge rainfall over Kenya, *Atmospheric Research*(2019), https://doi.org/10.1016/j.atmosres.2019.104808

Gauging the performance of CMIP5 historical simulation in reproducing observed gauge rainfall over Kenya

Lucia Mumo [1,2], **Jinhua Yu**[1,2*]

[1]Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disaster, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing 210044, China

[2]College of Atmospheric Science, Nanjing University of Information Science and Technology, 210044 Nanjing, Jiangsu, China

*Email: jhyu64@163.com

Acknowledgement

This work was funded by the Nation Basic Research Program of China (2018YFC1507704) and NSFC Grant (41730961 and 41575083). The authors are grateful to Nanjing University of Information Science and Technology (NUIST), China for establishing a suitable environment for research. We express our sincere gratitude to Kenya Meteorology Department which provided the observed rainfall data, Met Office Hadley Centre for the observed SST data and Program on Climate Model Diagnosis and Intercomparison (PCMDI) group for publicly availing CMIP5 rainfall data which was used in the analysis. We acknowledge the efforts made by the editor to contact the learned reviewers and we acknowledge their constructive comments and suggestions which immensely improved the original version of the manuscript.

Abstract

The detrimental impacts of climate change have drawn the interest of many climate scientists towards understanding the past climate for the sake of preparing for the future. The current study evaluates the performance of 12 rainfall models available in Coupled Model Intercomparison Project Phase 5 (CMIP5) in reproducing observed rainfall over Kenya from 1979-2005. Several statistical metrics were deployed in quantifying the disparities between CMIP5 models, in situ, and the Global Precipitation Climatology Centre (GPCC v7) rainfall datasets. The results show satisfactory skill of CMIP5 models' in simulating the bimodal rainfall regime despite exhibiting dry (wet) bias during March-May (MAM) and October-November (OND) season, respectively. The models' skills in reproducing the interannual variability is relatively weak. However, majority of the models captures the temporal pattern with reasonable skills in OND than in MAM and annual rainfall. The impacts of Indian Ocean Dipole (IOD) and El Niño southern oscillation (ENSO) are marked in observed OND rains with no significance link in MAM rains. Overall, CMIP5 models' skills in replicating the mean statistics and teleconnection links are relatively weak. Remarkably, the performance of models at different time scale, metrics, simulation of dynamical and teleconnection patterns are inconsistent among models. Nevertheless, based on skill score, the models are listed from top to bottom as; MPI-ESM-MR, CSIRO-MK3-6-0, GISS-E2-R, MPI-CGCM2.3.3, EC-EARTH, MIROC-ESM-CHEM, FGOALS-g2, BCC-CSM1.1-M, HADGEM-AO, CanESM2, GFDL-ESM2G, IPSL-CM5A-MR, and MME model. This study sheds light on the use of statistical metrics and teleconnection pattern to rank CMIP5 models and forms basis on model selection. Model parameterization over tropics is prudent, and bias correction is paramount in future projections and impacts studies.

Keywords: CMIP5; Rainfall; Kenya; teleconnections

1. Introduction

Climate change is an unequivocal phenomenon which has drawn the interest of many researchers globally with a common goal of understanding the past, present, and future projections. According to a report by Intergovernmental Panel on Climate change (IPCC, 2014), poverty is elevating in developing countries, while in developed countries it has increased socio-economic inequality due to dreadful impacts of climate change. Kenya's economy depends on rain-fed agriculture (World Bank, 2012), and pronounced seasonality, drought within a season accompanied by high rainfall variability (Mupagwa et al., 2006) has threatened the livelihood of majority of the citizens. Rainfall is the key factor in crop production (Adhikari et al., 2015; Mumo et al., 2018), and in the hydrological cycle (Mehran et al., 2014). This calls for need to understand the performance of Global circulations models (GCMs) in simulating the historical rainfall. This is because, future projections depend on the selection of the models especially in rainfall since previous studies have shown existence of huge model divergence over Africa (Williams and Funk, 2011; Otieno and Anyah, 2013; Bhattacharjee and Zaitchik, 2015; Endris et al., 2016; Ongoma et al., 2018).

GCMs are crucial tool for climate science and have been vastly used by IPCC periodic future projections. Various studies have used GCMs to examine climate aspects in different parts of the globe. At the global scale, several scholars have shown that GCMs models have high biases over Africa and Asia but depict low biases in Europe (Braconnot et al., 2012; Otieno and Anyah. 2013; Blázquez and Nuñez.2013; Müller et al., 2014; Tierney et al., 2015; Raghavan et al., 2018;Ayugi et al., 2019). Wood et al. (2004), further elucidated that, most of the GCMs models exhibit significant bias in simulating rainfall than other climate variables. In Africa, different researchers have deployed various methods to rank CMIP3/5. For instance, Otieno and Anyah (2013), Jury (2015), and Ongoma et al. (2018) used statistical metrics whilst Brands et al. (2013) employed large scale atmospheric circulations. Similarly, Rowell (2013), Martin et al. (2014), and Endris et al. (2016) conducted validation through teleconnection patterns whereas, Bhattacharjee and Zaitchik (2015) and Blázquez and Nuñez (2013) combined all the methodologies mentioned above to rank the models. The results show that there exists a divergence view of GCMs description of Africa climatology and variability as a continent and, also over different countries. For example, in the Great Horn of Africa (GHA), southern Africa,

the Sahel regions due to their intricate climate, there exist extensive models' discrepancies (Williams and Funk, 2011; Otieno and Anyah, 2013). Additionally, areas that experience enhanced Hadley cell circulations like the Northern parts of Africa, and certain regions of tropical Africa, some models agree on the sign of the rainfall trend (IPCC, 2013). Regardless of some region having good models' agreement, the regions with much ambiguity add up to a larger part of the Africa continent than the regions with a significant agreement (Bhattacharjee and Zaitchik, 2015). This call for a revisit of the current GCMs output over Africa as a continent since the regions with poor models' performance are highly vulnerable to impacts of climate change and variability.

Over Kenya, the main rainy season (March- April) is drastically decreasing and threatening the economy at large (Tierney et al., 2015; Ayugi et al., 2016; Ongoma and Chen, 2017; Mumo et al., 2019). Concurrently, frequency and intensity of droughts have increased recently over the region (Williams and Funk 2011). Paradoxically, GCMs projection about future rainfall shows wetter conditions (Otieno and Anyah, 2013; Ongoma et al.,2018) which is attributed to the influx of high quantities of greenhouse gases (Chadwick et al., 2015). This inconsistency between GCMs and observed rainfall trend is termed as "East Africa Climate Paradox" (Rowell et al., 2015). Brands et al., (2013), has associated this discrepancy to the paucity of in situ data required for model parameterization over the region. Additionally, rainfall biases in CMIP3/5 have been attributed to several factors like excessive equatorial Pacific cold tongue, double intertropical convergence zone (ITCZ), and weak Atlantic meridional overturning circulation (Lee and Wang, 2014). These models' divergences can be reduced by using a big ensemble GCM simulation. However, due to the high cost of running models, a limited number of ensembles are run to produce one model output which is free online (Taylor et al., 2012). Despite the documented shortcomings of CMIP models, newer phase of CMIP models have shown improved ability to simulate climate (Reichler and Kim, 2008). For example, Anyah and Qiu, (2012) compared the performance of CMIP3/5 over GHA region and found out that, CMIP3 placed the peak of March - May seasonal rainfall in May instead of April. This discrepancy has been resolved in CMIP5 algorithms to capture the seasonal maxima correctly.

There is a dire need to assess the representation of the statistical link between rainfall and teleconnection patterns taking into consideration that, rainfall over the country is not only

controlled by synoptic and mesoscale factors but also teleconnection systems. This is crucial because teleconnection links are bound to change with time as the climate evolves, and their impacts have been projected to become stronger (Endris et al., 2018). The major teleconnection patterns governing rainfall over the study domain are; sea surface temperature variations over the Indian Ocean, forming the Indian Ocean Dipole (IOD) (Saji and Yamagata, 2003; Yamagata et al., 2004). The disparity of sea surface temperature over the equatorial Pacific Ocean, forming the El Niño-Southern Oscillation (ENSO) index (Ropelewski and Halpert, 1987). The strength of these two indices (IOD and ENSO) determines the occurrence of extremes events over the region (Camberlin and Okoola, 2003). Notably, failure of GCM rainfall model to capture the elementary climate dynamics and teleconnection links signify its inability to simulate future projections reliably (Bhattacharjee and Zaitchik, 2015). This is due to the diagnostic nature of rainfall, which can be bias-corrected if the model fails to capture statistical metrics, but it cannot be rectified if the model does not capture the dynamics aspects. To minimize models' discrepancy in the future projections, use of sophisticated techniques that goes beyond replicating the mean climatology of the variable of interest but evaluating the climate dynamics of the models and the associated known teleconnection is encouraged. Ranking of GCMs should be treated with caution, and the sole purpose of the model selection should be noted first. Based on this background, this study will assess and rank the performance of 12 historical CMIP5 rainfall models relative to in situ and GPCC rainfall datasets.

The rest of the paper will cover the following sections; section two, study domain, data, and methodology. Section 3 elaborates models' skill to capture mean statistics, seasonality, and teleconnection associations and ranking of the models. Conclusion and recommendation are drawn in section 4.

2. Study Domain, Data, and Methodology

2.1 Study Domain

Kenya is bounded on latitude 5° S - 5° N and longitude 34°E - 42° E. Different ecosystems exist in the country due to varying climate characteristics and heterogeneous landscape (see Fig. 1a, of Mumo et al., 2019). Rainfall over the region is highly erratic accompanied by droughts even within a rainy season (Mupangwa et al., 2006). The country receives two rainy seasons classically classified as; March - May (MAM) "Long rains", and October - December (OND)

"Short rains" with a peak in April and November respectively (Fig. 1). Despite the dominant biannual rainfall regime, some studies have shown that the western highlands receive quantifiable rainfall amounts during June - August (JJA) (Ogwang et al., 2015). MAM is usually known as the primary rainy season due to its great spatial extent and amounts especially in the lake regions and along the coast, while OND is the main rainy season at the central and lower eastern regions (Ayugi et al., 2016; Mumo et al., 2019)

Rainfall over the region is controlled by large scale features like ENSO index which is known to have a significant influence on OND season (Ntale and Gan, 2004; Endris et al., 2013; Bahaga et al., 2015) and is associated with wet conditions (El Niño) and dry conditions (La Niña). Moreover, IOD significantly influences OND than MAM seasonal rainfall (Behera et al., 2005). Large water bodies like Lake Victoria and the Indian ocean modulates the climate of the region (Ogwang et al., 2014). Intertropical Convergence Zone (ITCZ) controls rainfall seasonality over the country and the entire EA region (Black et al., 2003; Endris et al., 2013).. Despite the influence of large-scale systems on rainfall, its distributions are moderated by complex topography which varies from low lying areas at the coast to central highlands which host Mt. Kenya at 5199 meters. In addition, the Great Rift valley stretches from north to south bisecting the country into two. Its features entail the undulating valleys and ridges, and large scales maritime features such as lakes Victoria, Nakuru, Naivasha, and Bogoria which controls the transport of airmasses hence determining the rainfall distribution within the great rift valley region and central parts of Kenya (Wakachala et al., 2015). Elaborated details on the influence of topography on rainfall distribution over East Africa (see Oettli and Camberlin, 2005)

2.2 Data

The choice of CMIP5 models evaluated in this study was based on a study by Knutti et al. (2013) which described the models' genealogy. The same selection of models has been adopted by Bhattacharjee and Zaitchik, (2015). CMIP5 historical models span from 1850 referred to as the pre-industrial period to 2005. The CMIP5 historical rainfall output has a number of models at a different resolution, the regridded dataset at a resolution of $1.0° \times 1.0°$ was drawn from http://climate-scenarios.canada.ca/?page= gridded-data. According to Taylor et al., (2012), each model has multiple ensembles runs which are forced by different greenhouse gas concentration, observed aerosols, variations in land use, and solar forcing. The couple models are assigned an

identification number by the World Climate Research Programme (WRCP) as *r1i1p1*, where *r* represents the first realization number, *i* is the initialization procedure used, and *p* refers to the perturbed physics number. More details of the CMIP5 models can be obtained from Taylor et al. (2012). The CMIP5 rainfall models used, their resolution, greenhouse gas composition and modeling center are highlighted in Table 1.

For consistent purpose, this study used the first run ensemble of each model as presented in Table 1. The choice of the study period (1979 - 2005) was based on the number of synoptic stations which had homogeneous and consistent rainfall data. Thirty-three synoptic stations were chosen for comparison with CMIP5 rainfall dataset. Point to point comparison is prudent due to geographical complexity of Kenya, as the mechanisms which control weather differ from locale to another and models' skill to capture diverse feature are different (Woldemeskel et al., 2016). Monthly rainfall in situ data was sourced from Kenya Meteorological Department (KMD) and is available upon request. In addition, observed gridded based dataset, Global Precipitation and Climatology Centre (GPCCv7) (Schneider et al., 2014, 2016) was used to evaluate the models' ability to mimic the movement of ITCZ. Observed ENSO and IOD indices were calculated from mean monthly Sea Surface Temperature recorded at Met Office Hadley Centre (HadISST) (Rayner et al., 2003). For consistency purposes, the same period was applied for both models' simulations, GPCCv7, and in situ data.

Table 1. The CMIP5 models used, their greenhouse gas composition, and the modeling center.

2.3 Methodology

The CMIP5 rainfall and SST output were nudged to the reference datasets through bilinear interpolation (common resolution of $1.0° \times 1.0°$). Twelve CMIP5 rainfall models' performance was assessed by comparing the models' output with in situ rainfall and GPCC datasets from 1979-2005 over Kenya. For point to point comparison between station data and CMIP5 models, bilinear interpolation was used to extract the model data which corresponds to each station coordinate. A minimum of 33 points corresponding to synoptic stations, which are ununiformly distributed within the country were used for comparison (Mumo et al., 2019, Fig. 1b). Standard normal homogeneity test (SNHT), introduced by Alexanderson, (1986) was used to test the homogeneity and consistency of the in-situ data. The results show that, all the stations were

consistency. We adopted Casanova and Ahrens, (2009) methodology to build the multimodal ensemble model (MME) by assigning equal weight to each model and then averaging the values. In this study, the skills of the models were assessed based on four categories; 1). Mean climate, 2). Seasonality, 3). Interannual variability, and 4). Teleconnection links. The following statistical metrics were deployed to examine the above-stated categories.

Normalized root means square error (nRMSE) was used to elucidate the relative average oddness of the GCMs models from the in-situ observations. After obtaining the RMSE (Eq.1), the values are scaled for comparison purposes by diving it by the mean/standard deviation of observation field (Randall, 2007). In this study, the RMSE was normalized by mean where μ_obs refers to the mean of the observation (Eq.2).

$$\mathrm{RMSE} = \frac{\sqrt{\sum_{i=1}^{n}(\mathrm{Obs_p} - \mathrm{Sim_p})^2}}{n} \quad (1)$$

$$\mathrm{nRMSE} = \frac{\mathrm{RMSE}}{\mu_{obs}} \quad (2)$$

A non-parametric Spearman's rank correlation method was utilized to establish the relationship between simulated rainfall and calculated ENSO/IOD indices. This method has been adopted from Endris et al., (2015), and Hauke and Kossowsi, (2011) since its does not assume linearity of variable as opposed to the Linear regression method. A two-tailed t-test was then applied to determine the significance of the correlation at a 90 % confidence level. The models' dynamics were evaluated through simulation of the movement of ITCZ as being controlled by the rainy bands of the monsoonal/trade winds. ITCZ is well documented as the main cause of rainfall seasonality over East Africa and was traced using a Hovmöller diagram. Hovmöller diagrams averages all variables under consideration, either along the longitude or latitude and put them in one axis while the other axis shows time (Hovmöller, 1949).

Kling-Gupta efficiency modified (KGE`) (Kling et al., 2012), is a modified form of Nash-Sutcliffe efficiency and considers the correlation coefficient *(r)*, bias factor, and the variability between observation and simulated data. KGE` solves the problems resulting from the interactions among variables like the case where variability is underestimated. KGE_α represents the variability ratio between the standard deviation of the model's simulation (σs) and the observed standard deviation (σo). If the KGE_α value is greater than 1, it implies overestimation

of the simulation model, and if is less than one, it signifies underestimation. KGE_β is used to represent the bias ratio between the mean of the simulated model (μs) and the observed mean (μo). KGE_β value ≥ 1 implies the presence of positive bias, while KGE_β value ≤ 1 signifies negative bias. KGE_r indicates the temporal dynamics relationship between the simulation and the observation expressed as Pearson correlation coefficient. It is worth noting that, the ideal value of KGE` is 1. Mathematically it can be illustrated as;

$$\text{KGE`} = 1 - \sqrt{(\alpha - 1)^2 + (\beta - 1)^2 + (r - 1)^2} \quad (3)$$

Where; $\alpha = \frac{\sigma_s}{\sigma_o}$ and, $\beta = \frac{\mu_s}{\mu_o}$

To measure the internal-model process which determines the physical processes of climate system like extreme events and trend, variability index *(VI)* was used (Scherrer et al., 2005; Palmer et al., 2008). Variability index which is obtained from the standard deviation of a model divided by the standard deviation of the observation exhibit bias since the model with tiny/huge *VI* mostly cancels out hence giving a false alarm of individual model (Scherrer, 2011). To evade that, this study chose to use *(VI)* introduced by (Gleckler et al., 2008). The new *VI* is calculated as the differences between the ratio of two models in question. The bounds in which to determine what to be called a good or bad *VI* is subjective simply because an ideal model should have a *VI* of zero. This signifies that the researcher can determine the limits to use in the study by estimating it from this ratio ($\frac{\sigma_{Sim}^i}{\sigma_{obs}^i}$). Mathematically, the *VI* is defined using Eq. 4;

$$\text{VI} = \left(\frac{\sigma_{\text{Sim}}^{\text{i}}}{\sigma_{\text{obs}}^{\text{i}}} - \frac{\sigma_{\text{obs}}^{\text{i}}}{\sigma_{\text{Sim}}^{\text{i}}}\right)^2 \quad (4)$$

Where, σ_{sim}^i represents the standard deviation of the simulation model rainfall at given station *i*, while σ_{obs}^i represents the standard deviation of the observed rainfall at a given station *i*.

Relative percentage bias *(R.B)* was used to quantify the magnitude of model deviation from the observation. Scaling model's bias helps to determine which model is better than the other on a uniform ground. Statistically, the percentage bias can be expressed as;

$$R.B = \frac{\frac{1}{N}\sum_{i=1}^{n} Sim_p - Obs_p}{\frac{1}{N}\sum_{i=1}^{n} Obs_p} * 100\% \quad (5)$$

Mann Kendall (Mann, 1945; Kendall, 1975) is a robust non-parametric trend detection method, and it was used in this study to detect rainfall trend. Theil Sen's slope estimator (Sen, 1968), is also a non-parametric method which quantifies the magnitude of the trend. More details of the method can be found in (Ongoma and Chen, 2017; Ullah et al., 2018, Mumo et al., 2019).

To compare how spatially the combined statistical metrics which include; centered-root mean square error (RMSE), a standard deviation which shows the amplitude of variations between models, and the pattern correlation, Taylor diagram was used (Taylor, 2001). The closer the model is to the observation (reference data point), the better the model performance, and the opposite is true. Taylor diagram also shows the relative performance of models against each other hence signifying the existence/non-existence of consensus of the models toward the reference data.

The combined model's skill to simulate the observed rainfall and the teleconnections links using the methodologies discussed above was tested using skill score. This technique was adopted from (Fu et al., 2013; Anandhi and Nanjundiah, 2015; Arora et al., 2017). The model's statistics obtained at seasonal and annual were subjected to relative absolute error to quantify their closeness to the observed magnitude of the same metric as proposed by Fu et al., (2013). Then all the relative absolute error values were summed up to get the skill score (SSC) at each time scale. The following statistics were used to build SSC; correlation (CORR), KGE` (Kling-Gupta-efficiency modified), nRMSE, VI, and absolute relative errors in; Mean (μ), standard deviation (σ), median (Med), skewness (cS), coefficient of kurtosis (cK), trend (MK), and Theil-slope estimator (TSA). The model with the smallest SSC value was termed as the best. Mathematically, it can be stated as;

$$SSC = (1 - CORR) + (1 - KGE`) + nRMSE + |\mu| + |\sigma| + |Med| + |cS| + |cK| + |MK| + |TSA| \quad (6)$$

The skill score of each model at each station was averaged over the whole country to get the mean skill score as expressed in Eq.7;

$$SSC_{MEAN} = \frac{\sum_{n=1}^{i} SSC}{P} \quad (7)$$

Where, SSC is the skill score at a given station *(i)*, and *P* is the total number of stations used.

3. Results and Discussions

3.1 Mean annual cycle

It was paramount to evaluate the mean annual cycle since a good model is regarded as the one that can capture the seasonal cycle of a certain weather parameter (Sperber and Palmer, 1996). Results indicate that, Kenya receives biannual rainfall regimes with two distinct seasons defined as MAM and OND with a peak in April and November, respectively (Fig. 1). All the models assessed exhibit dry bias during MAM season while 7/12 show wet bias during OND in comparison with the in-situ data (Fig. 1). The problem of CMIP rainfall models' tendency to depict dry (wet) bias during MAM (OND) rainy season over Equatorial East Africa and in extension the Great Horn of Africa (GHA) has been reported (Anyah and Qiu, 2012; Otieno and Anyah, 2013; Yang et al., 2015; Ongoma et al., 2018). The gridded data (GPCC), captured the observed mean annual cycle (r = 0.98) but underestimated its magnitude. Due to this high level of agreement between the two datasets, GPCC was used in analyzing the ITCZ movements. Since models' errors are bound to be carried forward in future projection, underestimation of the MAM season can adversely impact the economy due to lack of adequate preparedness to adapt probably to flood caused by unexpected heavy rainfall. On the other hand, overestimation of the OND rains can negatively impact the agriculture sector and consequently the whole economy at large due to overdependence on rainfed agriculture as people anticipate more rainfall, which turns to be less than expected.

Fig. 1. Mean annual rainfall cycle comparison between CMIP5 models, MME model in thick black color, GPCC in thick red color, and station data (thick blue color).

The model's ability to reproduce the seasonality of rainfall over Kenya is principally underlined on its ability to simulate the movements of ITCZ. Different scholars have associated the migration of ITCZ with the surface pressure and convergence, outgoing longwave radiation, cloud-top height, and rainfall, but unfortunately, these diagnostics rarely align (Nicholson, 2013). In this paper, we chose to deploy rainfall as a proxy for ITCZ since rainfall is the theme variable

and also it has been vastly used by previous studies to simulate ITCZ over East Africa region (Bhattacharjee and Zaitchik, 2015; Ongoma et al., 2018; Kisembe et al., 2018). The movements of ITCZ to the south bring OND rainy season, and as it retreats to the north, it brings MAM rainy season over the study domain. To simulate the three distinct features (onset, peak, and cessation) of seasonal rainfall over Kenya, Hovmöller diagram was used. Hovmöller diagram provides a framework to determine the model's ability to capture seasonality and intra-seasonality of rainfall (Hourdin et al., 2010). In this category, GPCC was used as the reference dataset to evaluate the performance of the models. The results as presented in Fig. 2 show that most of the models fairly captured the movement of the ITCZ by portraying the two wet seasons (MAM and OND), and one dry season (JJAS). GPCC shows the anticipated bimodal rainfall regime with pronounced seasonality along the Equator.

The simulation of ITCZ by CMIP5 models differs from one model to another with majority capturing the correct seasonality. Overall, the wettest model is BCC-CSM1.1-M while the driest model is IPSL-CM5A-MR. Apart from IPSL-CM5A-MR exhibiting the highest dry bias, the model has late onset of OND rains which starts in November and extends to January. CanESM2, EC-EARTH, CSIRO-Mk-6-0, GFDL-ESM2G, HADGEM-AO, MPI-ESM-MR, MIROC-ESM-CHEM, MRI-CGCM3, and MME replicate with reasonable skills the bimodal rainfall pattern. Notably, MAM season starting synchronously throughout the country in CSIRO-Mk-6-0. GISS-E2-R and FGOALS g2 do not simulate the intra-seasonal with satisfactory skill as the models do not capture the dry season but simulate rainfall from April – December along the Equator. These two models capture the JJA rainy season which is attributed to the influx of Congo air masses in the region (Ogwang et al., 2015). The highest rainfall amount during MAM (OND) is reproduced by GISS-E2-R (MRI-CGCM3). This is because the models push the rain bands far south leading to wet bias in the simulation. This can be attributed to the facts that, the southern oceans absorb a lot of incoming solar radiation leading to the well documented SST, wind, and rainfall biases (Williams et al., 2014). This phenomenon is termed as the Southern hemisphere albedo error (Haywood et al., 2016). Failure of a model to simulate a dynamical prominent phenomenon like ITCZ which is bound to change due to interaction with mesoscale and synoptic scale features is realistically unreliable to project rainfall in the future accurately (Bhattacharjee and Zaitchik, 2015)

It is worth noting that, most of the models' discrepancies occur in simulating the start, the peak and the cessation of the seasonal rainfall. These inconsistencies arise from their different abilities to simulate the mechanisms behind rainfall formation over the region. These mechanisms include topography, mesoscale and largescale features, and teleconnection patterns which influence rainfall over the region. This calls for the modeling group to check the internal-model physics, which shapes the model locality bias. The rest of the CMIP5 model comparison is based only on in situ rainfall data.

Fig. 2. Hovmöller diagram depicting the monthly average rainfall along latitude (5°S - 5°N) in each month. J - D denotes January - December respectively.

Multiple aspects of CMIP5 models in reproducing the observed mean annual cycle are summarized in Taylor diagram (Fig. 3). The dashed green lines show the unbiased centered RMSE while the spatial variability is measured in terms of standard deviation. The best performing model in reproducing the mean annual rainfall over Kenya is CSIRO-Mk3-6-0 which is in consensus with the previous study by Yang et al., (2014) over the broad EA region. This is because it has a significant correlation ($r = 0.9$) at a 95% confident band, coupled with small unbiased RMSE. Most of the models (9/12), in addition of the MME model, can simulate the annual rainfall cycle amplitude despite showing low spatial variability than the observation. Three models; BCC-CSM1.1-M, GFDL-ESM2G, MPI-ESM-MR overestimate the differences between dry and wet months over Kenya due to their high standard deviation (>1). The lowest performing model in this category is MPI-ESM-MR as shown by high standard deviation of 1.45 and low temporal correlation pattern of 0.38.

Fig. 3. Zonal average rainfall of CMIP5 models in comparison with the OBS (station data). The alphabetical letters represent the models; **A** - BCC-CSM1.1-M, **B** - CanESM2, **C**- CSIRO-Mk3-6-0, **D** - EC-EARTH, **E** - FGOALS-g2, **F**- GFDL-ESM2G, **G** - GISS-E2-R, **H** - HADGEM2-AO, **I**- IPSL-CM5A-MR, **J** - MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L** - MRI-CGCM3, and **M** - MME model.

3.2 Spatial distribution of annual rainfall and associated models' biases

Rainfall over the country is unevenly distributed with high amount recorded along the shores of Lake Victoria which exceeds 1800 mm/year, and along the coastline. These anomalous wet conditions in those regions are caused by enormous convective activities. Enhanced rainfall is also observed around Mt. Kenya and Mt. Elgon regions due to orographic effects (Fig. 4a). The Northern parts and northeastern parts receive below 800 mm/year and are delineated as Arid and Semi-Arid Lands (ASALs). Elaborated details about the mechanism behind the observed rainfall pattern (see Ayugi et al., 2018b; Camberlin, 2018; Mumo et al., 2019).

The ability of the CMIP5 models to capture the observed spatial pattern varies considerably from one station to the other on the bases of sign and magnitude (Fig. 4b - m). The models portray systematic relative dry bias along the shores of Lake Victoria and Mount Kenya. Similar results of dry conditions over Lake Victoria have been recorded (Jury, 2015; Ongoma et al., 2018; Kisembe et al., 2019). This dry bias can be linked with strong convective activity over the Lake basin and along the Indian Ocean. All the models show wet conditions over peculiar zone in the northwest region (Lodwar station) except GISS-E2-R. Similarly, all the models underestimate relief rainfall around Kenyan highlands i.e Mt. Elgon and Mt. Kenya regions. This can be linked to orographic effects which has been identified also in CORDEX models in simulating mean rainfall over south Africa (Favre et al., 2016). At an annual average, the wettest model is GISS-E2-R with a wet bias of 23.9 % and shows the most humid conditions over the northeastern parts of the country. Contrary, on average, the driest model is IPSL-CM5A-MR (-62.9 %) and depicts dry conditions in most parts of the country. It is worth noting that, the models exhibit huge biases in areas that are climatologically dry (wet) by showing they are wet (dry). This has been attributed to the facts that, in the tropospheric column of GCMs, there is larger warming in areas that experience descending motions according to climatology while the areas with high convective activities (updrafts) climatologically has lower warming under both greenhouse gas forcing and uniform sea surface temperature (SST) experiments (Ma et al., 2012)

Conclusively, the CMIP5 models evaluated underestimated annual rainfall (9/12) over Kenya (Table 2). The remaining three models which show wet biases are in this order; GISS-E2-R, BCC-CSM1.1-M, and FGOALS-g2. Their wet biases range from 2.2% - 23.9% coupled with normalized RMSE of 0.2 and 0.3 mm/year, respectively (Table 2). The errors in CMIP5 models

in simulating the observed rainfall have been associated with convective parameterization schemes (Hohenegger et al., 2008). These schemes are designed to show the average effect of convection on a model grid which is passed to the grid-scale as an increase of momentum, temperature, and moisture (Kendon et al., 2012). According to Hohenegger et al., (2008), these convective schemes were developed for tropics and coarse resolution models hence not suitable for extratropic regions and high-resolution models. This calls for the modeling groups to reparametrize the convective schemes well to deal with uncertain convective activity in the atmospheric column, which leads to unrealistic results in the models.

Fig. 4. Spatial distribution of mean annual rainfall based on station data (a), and relative bias of CMIP5 models in relation to station data. The models are as follows; BCC-CSM1.1-M (b), CanESM2 (c), CSIRO-MK3-6-0 (d), EC-EARTH (e), FGOALS-g2 (f), GFDL-ESM2G (g), GISS-E2-R (h), HADGEM2-AO (i), IPSL-CM5A-MR (j), MIROC-ESM-CHEM (k), MPI-ESM-MR (l), MRI-CGCM3 (m), and MME model (n).

CMIP5 models do not show satisfactory skills in estimating the interannual variability of annual rainfall over Kenya as seen by low values of KGE` and insignificant correlation coefficient ($\alpha \leq 0.1$) (Table 2). Mann-Kendall trend test show that, observed annual rainfall was decreasing at a rate of - 0.04 mm/year but (9/12) CMIP5 models show an increasing trend with only (GFDL-ESM2G, GISS-E2-R, and MRI-CGCM3) simulating the observed negative trend correctly. To summarize different facets of the models, Taylor diagram show that, there is little consensus among the models in reproducing the interannual variability, spatial pattern and amplitude of the observed annual rainfall (Fig. 5). Almost all the models can reproduce the correct amplitude despite exhibiting low spatial variability (low standard deviation than the observation) except CanESM2, which has high variability than the observation.

Table 2. Annual statistics for CMIP5 models and observation.

Fig. 5. Taylor diagram showing performance of CMIP5 models against station data at annual scale. The alphabetical letters represent the models; **A** - BCC-CSM1.1-M, **B** - CanESM2, **C**- CSIRO-MK3-6-0, **D** - EC-EARTH, **E** - FGOALS-g2, **F**- GFDL-ESM2G, **G** - GISS-E2-R, **H**- HADGEM2-AO, **I** - IPSL-CM5A-MR, **J**- MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L**- MRI-CGCM3, and **M**- MME model.

3.3 Interannual variability of MAM seasonal rainfall

Understanding interannual variability of the seasonal rainfall is crucial since the occurrence of catastrophically extreme events (droughts/floods) over the region are strongly linked with coupled ocean-atmospheric features. The mechanisms which control the magnitude of rainfall predominantly underlies on the drivers of interannual variability which include ENSO, IOD, Madden Julian Oscillation (MJO), Quesi-biennial Oscillation (QBO) (Indeje et al., 2000; Manatsa et al., 2014; Ogwang et al., 2015). In this regard, Pearson correlation was used to evaluation CMIP5 models' skill to simulate temporal dynamics of seasonal rainfall over Kenya. During MAM (Fig. 6), results show that, BCC-CSM1.1-M, MRI-ESM-MR, and IPSL-CM5A-MR, depicted good model agreement with reference data as portrayed by positive correlation in most parts of the country. These models show similar dynamics of simulating MAM seasonal rainfall i.e southwest to northeast positive slant pattern while the coastal area is dominated by negative correlation. This can be attributed to the incompetence of the models to simulate convective activities along the coast and lake basin regions. The CSIRO-Mk3-6-0, FGOALS-g2, GISS-E2-R, HADGEM-AO, GFDL-ESM2G, MIROC-ESM-CHEM, MRI-CGCM3 and MME portray a low level of agreement in most parts of the country as depicted by negative correlation values and unrealistic rainfall pattern. It is worth noting that, most of the models evaluated depict insufficient skill in simulating the dynamics which govern rainfall formation along the coastal region except CanESM2 and EC-EARTH. Moreover, 9/12 of the models have KGE` values less than zero, and the ensemble model is the worst with a KGE` value of - 0.5. This is attributed to a distinct teleconnection pattern impacts in each month during this season, as reported by Nicholson (2017).

Fig. 6. Spatial distribution of temporal agreement between observed MAM rains and CMIP5 models from 1979-2005 over Kenya.

Additionally, all models underestimated MAM seasonal rainfall with dry bias ranging from -1.1% in BCC-CSM1.1-M to -68% in GFDL-ESM2G (Table 3). The linear trend and the magnitude of the slope were assessed using Mann Kendall test and Sens slope estimators, respectively. The results show that observed MAM seasonal rainfall was decreasing at a rate of -0.2 mm/year, but unfortunately, only 2/12 models (CanESM2 and GISS-E2-R) simulated the

correct sign of the trend but underestimating its magnitude. The poor performance of the CMIP5 models to simulate the correct rainfall trend, especially during the primary rainy season threatens people livelihood. Despite the facts that, the causes of rainfall trend are outside the scope of this recent work, multidecadal variability of SST anomaly extending from the Indian Ocean to the Pacific Ocean are the leading causes (Bahaga et al., 2019)

Table 3. MAM seasonal rainfall statistics for CMIP5 models and observation. The values in bold depicts significant correlation values at $\alpha \leq 0.1$.

Taylor diagram of MAM seasonal rainfall in relation to the observation shows that most of the models' temporal patterns are different from the observation, including even the ensemble model. This is supported by an insignificant correlation for most of the models coupled with big RMSE. There is no consistency among models in reproducing the observed rainfall pattern since the models are spread apart, and far from the reference data (Fig. 7). With the exception of CanESM2 and CSIRO-Mk3-6-0, all the other models could reproduce the amplitude of the observed rainfall but with low spatial variability. This presents the inability of the models to simulate skillfully the mechanism that governs interannual variability of this seasonal rainfall. Similar large model divergence and inability of CMIP5 rainfall to capture interannual variability has been reported in different parts of the globe (Ongoma et al., 2018; Raghavan et al., 2018)

Fig. 7. Fig. caption same as Fig. 5, but during MAM rainy season.

3.4 Interannual variability of OND seasonal rainfall

The impact of climate change is bound to increase the frequency and intensity of extreme ENSO signal (Cai et al., 2014). This signal is associated with anomalously wet(dry) rainfall episode during OND than in MAM (Bahaga et al., 2015; Endris et al., 2016; Kisembe et al., 2018), hence the need to understand the interannual variability during this season. During OND, the ability of the models to simulate the interannual variability of the observed patterns differed from one model to another. Most of the models during OND were overestimating rainfall (8/12) in addition to the MME model. The wet bias ranged from 5.3% in IPSL-CM5A-MR to 67.7% in MIROC-ESM-CHEM (Table 4). Despite the models exhibiting wet bias, the models depicted reasonable skills in capturing the long-term variability of the in-situ rainfall i.e (7/12) (Fig.8). The FGOALS-g2, GISS-E2-R**,** IPSL-CM5A-MR, MPI-ESM-MR, and MRI-CGCM3 models are incompetent in capturing long-term dynamics of OND seasonal rainfall (Fig.8). Overall, all the

models' pattern correlation is insignificant except MIROC-ESM-CHEM (r = 0.4) according to a two-sided t-test (Table 4).

Fig. 8. Spatial distribution of temporal agreement between observed OND rains and CMIP5 models from 1979-2005 over Kenya.

There is considerable models' consensus on right skewness of data that is similar to the observations and it implies occurrence of wet conditions. In addition, Taylor diagram (Fig. 9), shows that most of the models can simulate the right amplitude irrespective of having low spatial variability except CanESM2 and MRI-CGCM3. MK results elucidate that, observed rainfall was increasing at a rate of 0.1 mm/year (Table 4). This elevating trend has been associated with increasing intensity of ENSO signals and positive IOD (Shongwe et al., 2011). Interestingly, 10/12 models except GFDL-ESM2G and HADGEM-AO simulated the correct sign of the trend, though the majority overestimated the magnitude. Particularly, CanESM2 highly overestimated the observed magnitude (6.5 mm/year) (Table 4).

Notably, if one had to rank the models based on the level of temporal agreement, MIROC-ESM-CHEM could be termed as the best model as portrayed by dominant positive correlation throughout the country and the model rationally showing OND seasonal distributed with high concentration at the central regions and lower eastern parts (Ayugi et al., 2016; Mumo et al., 2019). The HADGEM2-AO showed rational skill in capturing the temporal dynamics of the in-situ data but failed to capture the seasonal distribution of OND rainfall. The poorest model in this category is IPSL-CM5A-MR as shown by dominance of negative correlation (Table 4) and contrary to OND rainfall distribution (Fig.8).

Fig. 9. Fig. caption same as Fig. 5, but during OND rainy season.

Remarkably, CMIP5 models performed better in OND than in MAM rainy season. Similar findings have been recorded (Ongoma et al., 2018; Kisembe et al., 2019) using reanalysis data. These differences in performance can be explained by the mechanisms modulating seasonal rainfall over the region. During MAM, rainfall is mainly controlled by local and mesoscale factors rather than large-scale factors (Camberlin and Philippon, 2002), which are not well captured by the GCMs. Amid OND season, significant large-scale features play a key role in rainfall distribution and are mimicked well by the GCMs hence the better performance of the

models during that season. This implies that, CMIP5 models are more reliable in projecting future changes in OND than MAM rainy season as depicted by their reasonable skills in capturing the historical patterns. These findings are congruent with previous studies over the extended EA (Hastenrath et al., 1993; Black et al., 2003)

Table 4. OND seasonal rainfall statistics for the CMIP5 models and in situ data. The values in bold show significant correlation values at $\alpha \leq 0.1$.

3.5 Statistical links between rainfall and teleconnection patterns

It was paramount to evaluate the skills of the models in relation to the known teleconnection. This is because GCMs have been attributed to capture well large scale climate phenomena more reliably than local processes. For future climate projection, its crucial for a GCM to capture regional variability associated with major climate features than how it simulates the exact mean rainfall amount or seasonal timing of rainfall at the location of interest (Bhattacharjee and Zaitchik, 2015). According to Wilby et al. (2002) GCMs that are able to associate local variability with large scale climate systems drivers can be reliably used to dynamically/statistically downscale rainfall in the region of interest. In Kenya and EA region at large, SST anomalies over the Indian Ocean (IOD) according to Yamagata et al. (2004) and ENSO as reported by Nicholson (2017) are strongly associated with interannual variability of rainfall over EA region. To assess the statistical relationship between observed ENSO or IOD indices with in situ data and simulated rainfall, Spearman rank correlation was performed at seasonal, and annual scale. Results show that observed OND and annual rainfall exhibit significant positive correlation at 95% confidence level with ENSO and IOD indices. The relationship of these indices with observed MAM rains is very weak with an insignificant negative (positive) correlation with ENSO (IOD) index (r = - 0.04 (0.05)). This implies that the two teleconnection patterns are dominant during OND than MAM rainy season. It is worth noting that IOD has a major influence on OND interannual variability (r = 0.77) than ENSO link (r = 0.5). This signifies that anomalously wet (dry) conditions during OND are associated with positive (negative) IOD signals. These findings are consistent with previous studies over the region (Clark et al., 2003; Endris et al., 2013; Bahaga et al., 2015; Endris et al., 2016; Kisembe et al., 2018).

The models' performance with the ENSO index at both seasonal and annual scale is poor since no model showed a significant relationship with the index. Furthermore, During MAM (OND), only 7 (9) models, including the MME model captured the signature of the ENSO event (Table 5). While at an annual scale, 8 models simulated the sign, but no significant relationship was shown. Nevertheless, if one had to rank the GCMs based on their capability to simulate ENSO index association with rainfall; during MAM season, GFDL-ESM2G model leads, while in OND and at annual scale FGOALS-g2 outperforms all other models including the MME model. The models' ability to reproduce the relationship between the IOD index and simulated MAM rainfall was good since 10 models, including the MME model reproduced the sign of the correlation coefficient during MAM. Notably, 3 models showed significant correlation at $\alpha \leq 0.1$ (CSIRO-Mk3-6-0, r = 0.43, MME, r = 0.4 and, EC-EARTH, r = 0.36) during MAM. CSIRO-Mk3-6-0 is the best model in capturing the IOD links with MAM seasonal rainfall and has been identified also by Yang et al., (2014) over EA. During OND, 9 models including the MME simulated correctly the sign of the statistical relation between the simulated OND rains and IOD signal with MIROC-ESM-CHEM outpacing all with r = 0.34 (Table 5). Amid annual scale, most of the models (10) inclusive of the MME replicated the sign of the statistical relationship with IOD except HADGEM2-AO and IPSL-CM5A-MR, which had a significant negative correlation. This can be associated with the anomalously dry bias of the models in simulating annual rainfall over most parts of the country (see Fig. 4). Conclusively, historical CMIP5 models poorly captures the statistical relationship with the known teleconnection patterns.

Table 5. Statistical links between observed rainfall, and CMIP5 rainfall simulations with ENSO/ IOD indices. The values in bold show significant correlation values $\alpha \leq 0.1$.

3.6 Ranking of Models

To determine the model which had a reasonable level of agreement with the observation, ranking of the models using skill score was inevitable. Ranking of the models is of great importance in model selection. Model selection is crucial for weather prediction purpose as some models have proved to perform better than MME model. Skill score results show that different models performed differently at different time scale (Fig. 10). During MAM season, the best models are in this order; CanESM2, GISS-E2-R, and MPI-ESM-MR, and the least performing model is FGOALS-g2. The best performing models in OND season are as follows; MPI-ESM-MR, CSIRO-MK3-6-0, and EC-EARTH, while IPSL-CM5A-MR and the MME model are the

poorest. At annual scale, the models can be ranked as; MPI-ESM-MR, CSIRO-MK3-6-0, and EC-EARTH while the poorest model in this category is IPSL-CM5A-MR and MME model. This is a clear indication that model performance is not consistent at different time scale. Based on that, the final ranking was done by averaging each model performance within the 3-time scale considered (See Fig. 10, Mean-SS). Overall, the models are arranged from the top to the least as follows; MPI-ESM-MR, CSIRO-MK3-6-0, GISS-E2-R, MRI-CGCM3, EC-EARTH, MIROC-ESM-CHEM, FGOALS-g2, BCC-CSM1.1-M, HADGEM-AO, CanESM2, GFDL-ESM2G, IPSL-CM5A-MR, and MME model. The MME model was the least performing model hence all the individual models outperformed it. This concurs with IPCC (2007) report that rainfall errors are typically large at the equatorial belt, hence making some models to outperform the MME model. Conclusively, CMIP5 models' performance in comparison with the in-situ data is big, i.e. overestimating or underestimating. The complex topography of the study region, the influence of teleconnection patterns which are difficult to represent in a coupled ocean-atmosphere modeling system and progression of circulation patterns during monsoonal transition presents great challenges to modelers (Hastenrath 1985) hence poor model performance in the region.

Fig. 10. Skill score of CMIP5 models during MAM, OND, and annual rainfall. Mean skill (Mean-SS) is in yellow color. The alphabetical letters represent the models; **A**- BCC-CSM1.1-M, **B**- CanESM2, **C**- CSIRO-Mk3-6-0, **D**- EC-EARTH, **E**- FGOALS-g2, **F**- GFDL-ESM2G, **G**- GISS-E2-R, **H**- HADGEM2-AO, **I**- IPSL-CM5A-MR, **J**- MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L**-MRI-CGCM3, and **M**- MME model.

4.0 Conclusion and Recommendation

This study evaluated the performance of twelve CMIP5 historical rainfall simulation in reproducing the observed rainfall and teleconnection links over Kenya from 1979 - 2005 using different metrics. The results reveal that, different statistical metrics lead to a different ranking of the models and the performance of an individual model is not consistent at different time scale. CMIP5 historical models show satisfactory skills in simulating the mean annual cycle as compared to interannual variability. However, all the models underestimated MAM seasonal rainfall with majority overestimating OND season. CMIP5 models show reasonable skills in replicating the mean rainfall statistics, trend and magnitude in OND than MAM and annual rainfall. There is huge modal divergence in simulating historical MAM than any other time scale

considered. Further analysis shows that Indian Ocean Dipole (IOD) and El Niño southern oscillation (ENSO) exhibits significance influence on observed OND than MAM rains. A significant statistical link is also seen between observed annual rainfall with IOD index.

The models' performance in replicating the statistical links between teleconnection patterns and simulated rainfall is poor. Most models showed rational skill in replicating the IOD than ENSO links with simulated rainfall, but the relationship is weakest in MAM than OND and annual rains. Overall, CMIP5 models show unsatisfactory skills in replicating the mean statistics and the teleconnection links. Regardless of that, to capture the long-term internal climate dynamics, we combined both model's skill to replicate mean rainfall statistics and teleconnections links and ranked the models. According to skill score test, the models from top to the bottom can be listed as; MPI-ESM-MR, CSIRO-MK3-6-0, GISS-E2-R, MRI-CGCM3, EC-EARTH, MIROC-ESM-CHEM, FGOALS-g2, BCC-CSM1.1-M, HADGEM-AO, CanESM2, GFDL-ESM2G, IPSL-CM5A-MR, and MME model. Due to the evidenced inconsistency and inherent biases of CMIP5 historical simulation in reproducing observed rainfall over Kenya, model ranking should be approached with caution and the aim of model selection should be identified first. We recommend utilization of bias correction methods when using the rainfall models for impact studies and future projection due to the diagnostic nature of rainfall. Future detailed studies are warranted over EA region and Africa as a continent. Improvement of model parameterization schemes in the next phase of CMIP is prudent over the tropics since CMIP3/5 models have been reported to misrepresent the seasonal rainfall over vulnerable regions which can catastrophically impact the susceptible economy.

Conflicts of interest

We wish to state that there is no conflict of interest.

References

Adhikari, U., Nejadhashemi, A.P., Woznicki, S.A., 2015. Climate change and eastern Africa: a review of impact on major crops. Food Energy Secur. 4, 110–132. https://doi.org/10.1002/fes3.61.

Alexandersson, H., 1986. A homogeneity test applied to precipitaion data. J Clim. 6, 661-675.

Anandhi, A., Nanjundiah, R. S., 2015. Performance evaluation of AR4 Climate Models in simulating daily precipitaion over the Indian region using skill scores. Int J Climatol. 119, 551-566. https://doi:10.1007/s00704-013-1043-5.

Anyah, R. O., Qiu, W., 2012. Characteristic 20th and 21st century precipitation and temperature patterns and changes over the Greater Horn of Africa. Int J Clim. 32, 347-363. https://doi:10.1002/joc.2270.

Arora, H., Ojha, C. S. P., Buytaert, W., Suryaprakash, G., 2017. Spatio-temporal trends in observed and downscaled precipitation over Ganga Basin. Hydrol. Earth Syst. Sci. https://doi:10.5194/hess-2017-388.

Ayugi, B. O., Wen, W., Chepkemoi, D., 2016. Analysis of spatial and temporal patterns of rainfall variations over Kenya. J. Environ. Earth Sci. 6.

Ayugi, B., Tan, G., Gnitou, G. T., Ojara, M., & Ongoma, V., 2019. Historical evaluations and simulations of precipitation over East Africa from Rossby centre regional climate model. *Atmospheric Res*, 104705. doi:https://doi.org/10.1016/j.atmosres.2019.104705

Bahaga, T., Mengistu Tsidu, G., Kucharski, F., Diro, G., 2015. Potential predictability of the sea- surface temperature forced equatorial East African short rains interannual variability in the 20th century. Quart. J. R. Meteorol. Soc. 141, 16-26. https://doi:10.1002/qj.2338.

Behera, S. K., Luo, J.-J., Masson, S., Delecluse, P., Gualdi, S., Navarra, A., Yamagata, T., 2005. Paramount impact of the Indian Ocean dipole on the East African short rains: A CGCM study. J Clim. 18, 4514-4530. https://doi:10.1175/JCLI3541.1.

Bhattacharjee, P. S., Zaroohik, B. F., 2015. Perspectives on CMIP5 model performance in the Nile River headwaters regions. Int J Clim. 35, 4262-4275. https://doi:10.1002/joc.4284.

Black, E., Slingo, J., Sperber, K. R., 2003. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. Mon. Weather Rev. 31, 74-94. https://doi:10.1175/15200493(2003)131\0074:AOSOTR[2.0.CO;2

Blázquez, J., Nuñez, M. N., 2013. Performance of a high resolution global model over southern South America. Int J Clim. 33, 904-919. https://doi:10.1002/joc.3478.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., . . . Zhao, Y., 2012. Evaluation of climate models using palaeoclimatic data. Nat. Clim. Change. 2, 417. https://doi:10.1038/nclimate1456.

Brands, S., Herrera, S., Fernández, J., Gutiérrez, J. M., 2013. How well do CMIP5 Earth System Models simulate present climate conditions in Europe and Africa? Clim. Dyn. 41, 803-817. https://doi:10.1007/s00382-013-1742-8.

Camberlin, P., 2018. In: Climate of Eastern Africa. Oxford University Press.

Camberlin, P., Okoola, R.E., 2003. The onset and cessation of the "long rains" in eastern Africa and their interannual variability. Theor. Appl Climatol. 54, 43–54. https://doi.org/10.1007/s00704-002-0721-5

Casanova, S., Ahrens, B., 2009. On the weighting of multimodel ensembles in seasonal and short-range weather forecasting. Mon. Weather Rev. 137, 3811-3822. https://doi:10.1175/2009MWR2893.1.

Clark, C. O., Webster, P. J., Cole, J. E., 2003. Interdecadal variability of the relationship between the Indian Ocean zonal mode and East African coastal rainfall anomalies. J Clim. 16, 548-554. https://doi:10.1175/1520-0442(2003)016<0548:IVOTRB>2.0.CO;2.

Endris, H. S., Lennard, C., Hewitson, B., Dosio, A., Nikulin, G., Artan, G. A., 2018. Future changes in precipitaion associated with ENSO, IOD and changes in the mean state over Eastern Africa. Clim. Dyn. 52, 2029-2053.

Endris, H. S., Lennard, C., Hewitson, B., Dosio, A., Nikulin, G., Panitz, H.-J., 2016. Teleconnection responses in multi-GCM driven CORDEX RCMs over Eastern Africa. Clim. Dyn. 46, 2821-2846. https://doi:10.1007/s00382-015-2734-7.

Endris, H. S., Omondi, P., Jain, S., Lennard, C., Hewitson, B., Chang'a, L., . . . Nikulin, G., 2013. Assessment of the performance of CORDEX regional climate models in simulating East African rainfall. J Clim. 26, 8453-8475. https://doi:10.1175/JCLI-D-12-00708.1.

Favre, A., Philippon, N., Pohl, B., Kalognomou, E.-A., Lennard, C., Hewitson, B., . . . Cerezo-Mota, R., 2016. Spatial distribution of rainfall annual cycles over South Africa in 10 CORDEX regional climate model present-day simulations. Clim. Dyn. 46, 1799-1818. https://doi:10.1007/s00382-015-2677-z.

Fu, G., Liu, Z., Charles, S. P., Xu, Z., Yao, Z., 2013. A score- based method for assessing the performance of GCMs: A case study of southeastern Australia. J. Geophys. Res. Atmos. 118, 4154-4167. https://doi:10.1002/jgrd.50269, 2013.

Giorgi, F., Mearns, L. O., 2002. Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging"(REA) method. J Clim. 15, 1141-1158. https://doi:10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2.

Gleckler, P. J., Taylor, K. E., Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res. Atmos. 113. https://doi:10.1029/2007JD008972

Hauke J, Kossowski T (2011) Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. Quaest Geogr. 30, 87–93

Haywood, J. M., Jones, A., Dunstone, N., Milton, S., Vellinga, M., Bodas- Salcedo, A., . . . Watanabe, S. 2016. The impact of equilibrating hemispheric albedos on tropical performance in the HadGEM2- ES coupled climate model. Geophys Res Lett. 43, 395-403. https://doi:10.1002/2015GL066803

Hohenegger, C., Brockhaus, P., Schar, C., 2008. Towards climate simulations at cloud-resolving scales. Meteorol Z. 17, 383-394. https://doi:10.1127/0941-2948/2008/0303

Hourdin, F., Musat, I., Guichard, F. s., Ruti, P. M., Favot, F., Filiberti, M.-A., . . . Marquet, P. (2010). AMMA-model intercomparison project. B. Am. Meteorol. Soc. 91, 95-104.

Hovmöller, E., 1949. The trough- and- ridge diagram. Tellus. 1, 62-66.

IPCC., 2001. Climate Change 2001: The Scientific Basis Cambridge University Press Cambridge. In: UK.

IPCC., 2007. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (Eds.), Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (p. 996). Cambridge: Cambridge University Press.

IPCC., 2013. Climate change 2013: the physical science basis. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds)

Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge .

IPCC., 2014. In: Pachauri, R.K., Meyer, L.A. (Eds.), Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, p. 151.

Jury, M. R., 2015. Statistical evaluation of CMIP5 climate change model simulations for the Ethiopian highlands. Int J Clim. 35, 37-44. https://doi:10.1002/joc.3960.

Kendall, M., 1975. Rank correlation methods (4th edn.) charles griffin. San Francisco, CA, 8.

Kendon, E. J., Roberts, N. M., Senior, C. A., Roberts, M. J., 2012. Realism of rainfall in a very high-resolution regional climate model. J Clim. 25, 5791-5806. https://doi:10.1175/JCLI-D-11-00562.1.

Kisembe, J., Favre, A., Dosio, A., Lennard, C., Sabiiti, G., Nimusiima, A., 2018. Evaluation of rainfall simulations over Uganda in CORDEX regional climate models. Int J Clim. 1-18. https://doi:10.1007/s00704-018-2643-x.

Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol., 424, 264-277. https://doi:10.1016/j.jhydrol.2012.01.011.

Knutti, R., Allen, M. R., Friedlingstein, P., Gregory, J. M., Hegerl, G. C., Meehl, G. A., . . . Raper, S., 2008. A review of uncertainties in global temperature projections over the twenty-first century. J Clim. 21, 2651-2663. https://doi:10.1175/2009JCLI3361.1.

Knutti, R., Masson, D., Gettelman, A., 2013. Climate model genealogy: Generation CMIP5 and how we got there. Geophys Res Lett. 40, 1194-1199.

Kumar, S., Merwade, V., Kinter III, J. L., Niyogi, D., 2013. Evaluation of temperature and rainfall trends and long-term persistence in CMIP5 twentieth-century climate simulations. J Clim. 26, 4168-4185. https://doi:10.1175/JCLI-D-12-00259.1.

Lee, J.-Y., Wang, B., 2014. Future change of global monsoon in the CMIP5. Clim. Dyn. 42, 101-119. https://doi:10.1007/s00382-012-1564-0

Mann, H. B., 1945. Nonparametric tests against trend. Econometrica. 245-259.

Martin, E. R., Thorncroft, C., Booth, B. B., 2014. The multidecadal Atlantic SST—Sahel rainfall teleconnection in CMIP5 simulations. J Clim. 27, 784-806. https://doi:10.1175/JCLI-D-13-00242.1.

Mehran, A., AghaKouchak, A., Phillips, T. J., 2014. Evaluation of CMIP5 continental rainfall simulations relative to satellite- based gauge- adjusted observations. J. Geophys. Res. Atmos. 119, 1695-1707. https://doi:10.1002/2013JD021152.

Müller, C., Waha, K., Bondeau, A., Heinke, J., 2014. Hotspots of climate change impacts in sub- Saharan Africa and implications for adaptation and development. Glob. Change Biol. 20, 2505-2517. https://doi:10.1111/gcb.12586.

Mupangwa, W., Love, D., Twomlow, S., 2006. Soil–water conservation and rainwater harvesting strategies in the semi-arid Mzingwane Catchment, Limpopo Basin, Zimbabwe. Physics and chemistry of the Earth, Parts A/B/C, 31(15-16), 893-900.

Mumo, L., Yu, J., Fang, K., 2018. Assessing Impacts of Seasonal Climate Variability on Maize Yield in Kenya. Int. J. Plant Prod. 12, 297-307. https://doi:10.1007/s42106-018-0027-x.

Mumo, L., Yu, J., Ayugi, B., 2019. Evaluation of spatiotemporal variability of rainfall over Kenya from 1979 to 2017. J. Atmospheric Sol.-Terr. Phys. 105097.

Nicholson, S. E., 2013. The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability. ISRN Meteorology, 2013. https://doi:10.1155/2013/453521

Ntale, H. K., Gan, T. Y., 2004. East African rainfall anomaly patterns in association with El Niño/Southern Oscillation. J. Hydrol. Eng. 9, 257-268. https://doi:10.1061/(ASCE)1084-0699(2004)9:4(257).

Oettli, P., Camberlin, P., 2005. Influence of topography on monthly rainfall distribution over East Africa. Clim. Res. 28, 199-212. https://doi:10.3354/cr028199.

Ogwang, B. A., Chen, H., Li, X., Gao, C., 2014. The influence of topography on East African October to December climate: sensitivity experiments with RegCM4. Adv Meteorol. https://doi:10.1155/2014/143917

Ogwang, B. A., Chen, H., Tan, G., Ongoma, V., Ntwali, D., 2015. Diagnosis of East African climate and the circulation mechanisms associated with extreme wet and dry events: a study based on RegCM4. Arab. J. Geosci. 8, 10255-10265. doi:10.1007/s12517-015-1949-6

Ongoma, V., Chen, H., 2017. Temporal and spatial variability of temperature and rainfall over East Africa from 1951 to 2010. Meteorol. Atmos. Phys. 129, 131-144. https://doi:10.1007/s00703-016-0462-0.

Ongoma, V., Chen, H., Gao, C., 2018. Evaluation of CMIP5 twentieth century rainfall simulation over the equatorial East Africa. Theor.Appl.Climatol. 135. https://doi:10.1007/s00704-018-2392-x.

Otieno, V. O., Anyah, R. O., 2013. CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part 1: contemporary climate. Clim. Dyn. 41, 2081-2097. https://doi:10.1007/s00382-012-1549-z.

Palmer, T., Doblas-Reyes, F., Weisheimer, A., Rodwell, M., 2008. Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. Bull. Am. Meteorol. Soc. 89, 459-470. https://doi:10.1175/BAMS-89-4-459

Raghavan, S. V., Liu, J., Nguyen, N. S., Vu, M. T., & Liong, S.-Y. (2018). Assessment of CMIP5 historical simulations of rainfall over Southeast Asia. *Theor. Appl. Climatol. 132*, 989-1002. doi:https://:doi/10.1007/s00704-017-2111-z

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., . . . Kaplan, A., 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res. ATMOS.108.

Reichler, T., Kim, J., 2008. How well do coupled models simulate today's climate? Bull. Am. Meteorol. Soc. 89, 303-312. https://doi:10.1175/BAMS-89-3-303

Ropelewski, C. F., Halpert, M. S., 1987. Global and regional scale rainfall patterns associated with the El Niño/Southern Oscillation. Mon. Weather Rev.115, 1606-1626. https://doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2

Rowell, D. P., 2013. Simulating SST teleconnection to Africa: What is the state of the art? J Clim. 26, 5397-5418. https://doi:10.1175/JCLI-D-12-00761.1.

Rowell, D. P., Booth, B. B., Nicholson, S. E., Good, P., 2015. Reconciling past and future rainfall trends over East Africa. J Clim. 28, 9768-9788. https://doi:10.1175/JCLID-15-0140.1.

Saji, N., Yamagata, T., 2003. Possible impacts of Indian Ocean dipole mode events on global climate. Clim. Res. 25, 151-169. https://doi:10.3354/cr025151.

Scherrer, S. C., 2011. Present- day interannual variability of surface climate in CMIP3 models and its relation to future warming. Int J Clim. 31, 1518-1529. https://doi:10.1002/joc.2170.

Scherrer, S. C., Appenzeller, C., Liniger, M. A., Schär, C., 2005. European temperature distribution changes in observations and climate change scenarios. Geophys. Res. Lett. 32. https://doi:10.1029/2005GL024108.

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B., 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor Appl Climatol. 115,15–40. https://doi.org/10.1007/s00704-013-0860-x

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese B ., 2016. GPCC full data reanalysis version 7.0: monthly land-surface precipitation from rain gauges built on GTS based and historic data. Research data archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D600007

Sen, P. K., 1968. Estimates of the regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379-1389.

Shongwe ME, van Oldenborgh GJ, van den Hurk B, van Aalst M., 2011. Projected changes in mean and extreme precipitation in Africa under global warming. Part II: East Africa. J Clim.24, 3718–3733. https://doi.org/10.1175/2010JCLI2883.1

Sperber, K., Palmer, T., 1996. Interannual tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project. J Clim. 9, 2727-2750. https://doi:10.1175/1520 0442(1996)009<2727:ITRVIG>2.0.CO;2.

Taylor, K., Stouffer, R., Meehl, G., 2012. An overview of CMIP5 and the experiment design, B. Am. Meteorol. Soc. 93, 485–498.

Taylor, K. E., 2001. Summarizing multiple aspects of model performance in a single diagram. . J. Geophys. Res. Atmos. 106, 7183-7192. https://doi:10.1029/2000JD900719.

Tebaldi, C., Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Philos. Trans. Royal Soc. A: mathematical, physical and engineering sciences. 365, 2053-2075. https://doi:10.1098/rsta.2007.2076.

Tierney, J. E., Ummenhofer, C. C., deMenocal, P. B., 2015. Past and future rainfall in the Horn of Africa. Sci. Adv. 1, e1500682. https://doi:10.1126/sciadv.1500682.

Ullah, S., You, Q., Ullah, W., Ali, A., 2018. Observed changes in rainfall in China-Pakistan economic corridor during 1980–2016. Atmospheric Res. 210, 1-14. https://doi:10.1016/j.atmosres.2018.04.007.

Wakachala, F.M., Shilenje, Z.W., Nguyo, J., Shaka, S., Apondo, W., 2015. Statistical patterns of rainfall variability in the Great Rift valley of Kenya. J. Environ. Agric. Sci. 5, 17–26

Wilby, R. L., Dawson, C. W., 2007. SDSM 4.2-A decision support tool for the assessment of regional climate change impacts. *User manual, 94*. https://doi:10.1016/S1364-8152(01)00060-3.

Williams, A. P., Funk, C., 2011. A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. Clim. Dyn. 37, 2417-2435. https://doi:10.1007/s00382-010-0884-y.

World Bank, 2012. Doing business in the East African economies.IFC/World Bank Rep., 116

Woldemeskel, F., Sharma, A., Sivakumar, B., Mehrotra, R., 2016. Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. J. Geophys. Res. Atmos. 121, 3-17. https://doi:10.1002/2015JD023719

Wood, A. W., Leung, L. R., Sridhar, V., Lettenmaier, D., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim. Change. 62, 189-216. https://doi:10.1023/B:CLIM.0000013685.99609.9e.

Yamagata, T., Behera, S. K., Luo, J.-J., Masson, S., Jury, M. R., Rao, S. A., 2004. Coupled ocean-atmosphere variability in the tropical Indian Ocean. Earth's Climate: The Ocean–Atmosphere Interaction, Geophys. Monogr.147, 189-212. https://doi:10.1029/147GM12

Yang W, Seager R, Cane MA, Lyon B., 2014. The east African long rains in observations and models. J Clim. 27, 7185–7202. https://doi. org/10.1175/JCLI-D-13-00447.1

Yang, W., Seager, R., Cane, M. A., Lyon, B.,, 2015. The rainfall annual cycle bias over East Africa in CMIP5 coupled climate models. J Clim. 28, 9789-9802. https://doi:10.1175/JCLI-D-15-0323.1

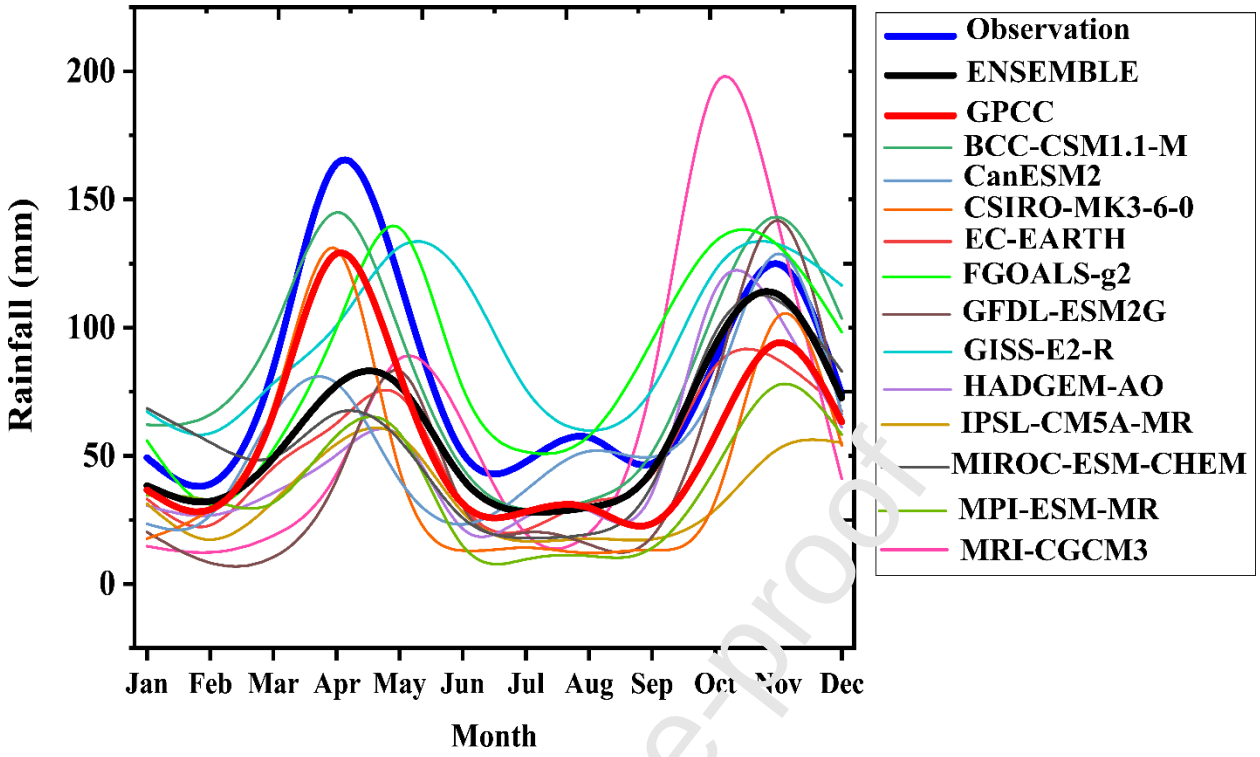

Fig. 1. Mean annual rainfall cycle comparison between CMIP5 models, MME model in thick black color, GPCC in thick red color, and station data (thick blue color).

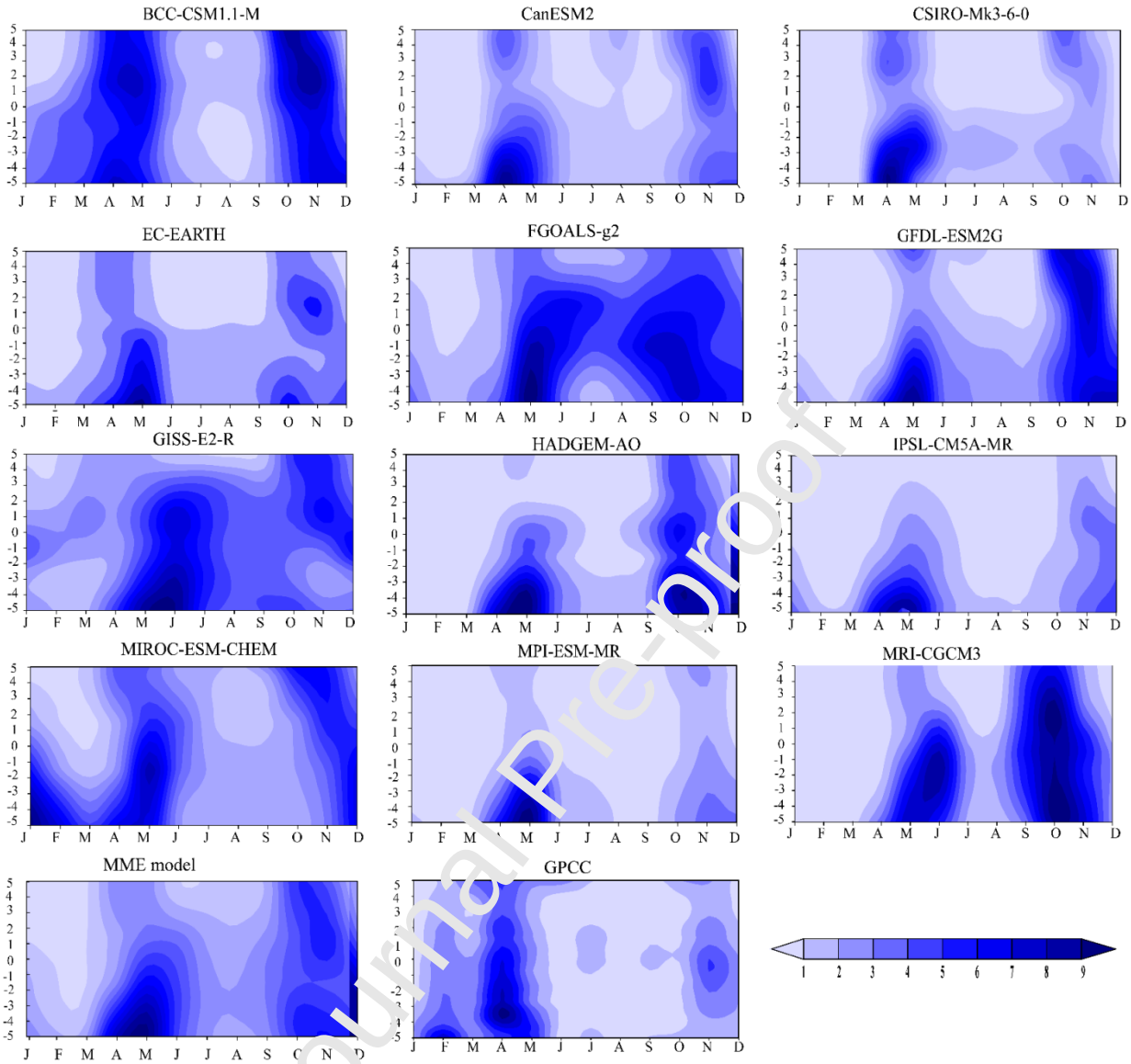

Fig. 2. Hovmöller diagram depicting the monthly average rainfall along latitude (5^0 S - 5^0 N) in each month. J - D denotes January - December respectively.

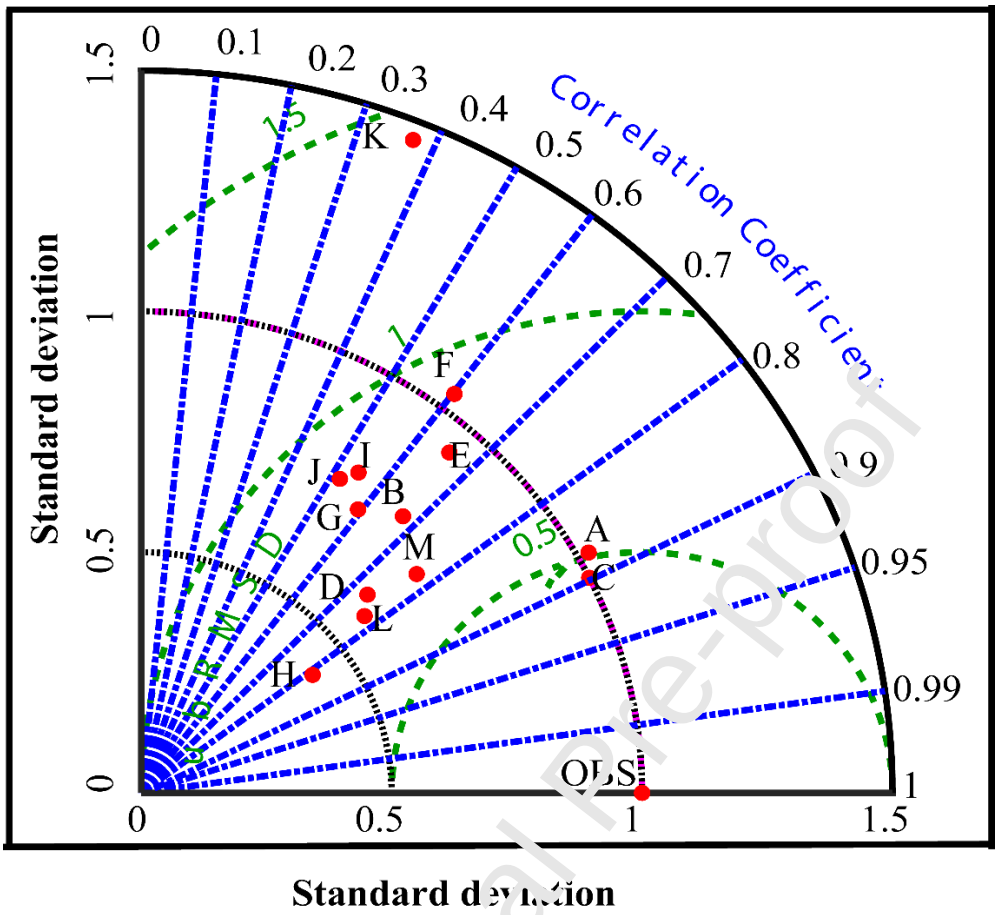

Fig. 3. Zonal average rainfall of CMIP5 models in comparison with the OBS (station data). The alphabetical letters represent the models; **A**- BCC-CSM1.1-M, **B**- CanESM2, **C**- CSIRO-Mk3-6-0, **D**- EC-EARTH, **E**- FGOALS-g2, **F**- GFDL-ESM2G, **G** - GISS-E2-R, **H** - HADGEM2-AO, **I**- IPSL-CM5A-MR, **J** - MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L**-MRI-CGCM3, and **M**- MME model.

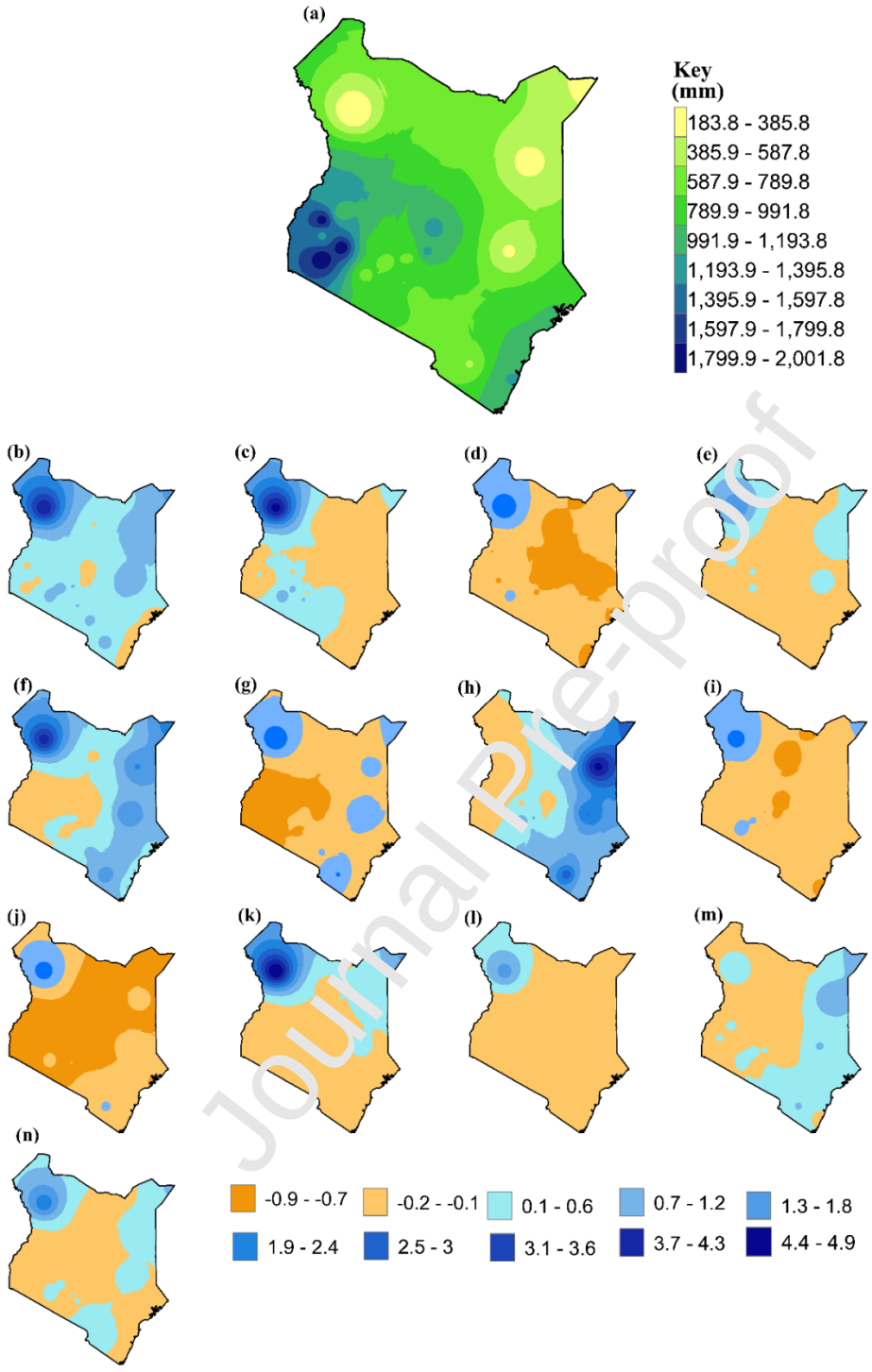

Fig. 4. Spatial distribution of mean annual rainfall based on station data (a), and Relative bias of CMIP5 models in relation to station data. The models are as follows; BCC-CSM-1.1-M(b), CanESM2 (c), CSIRO-Mk3-6-0 (d), EC-EARTH (e), FGOALS-g2 (f), GFDL-ESM2G (g), GISS-E2-R (h), HADGEM2-AO (i), IPSL-CM5A-MR (j), MIROC-ESM-CHEM (k), MPI-ESM-MR (l), MRI-CGCM3 (m), and MME model (n).

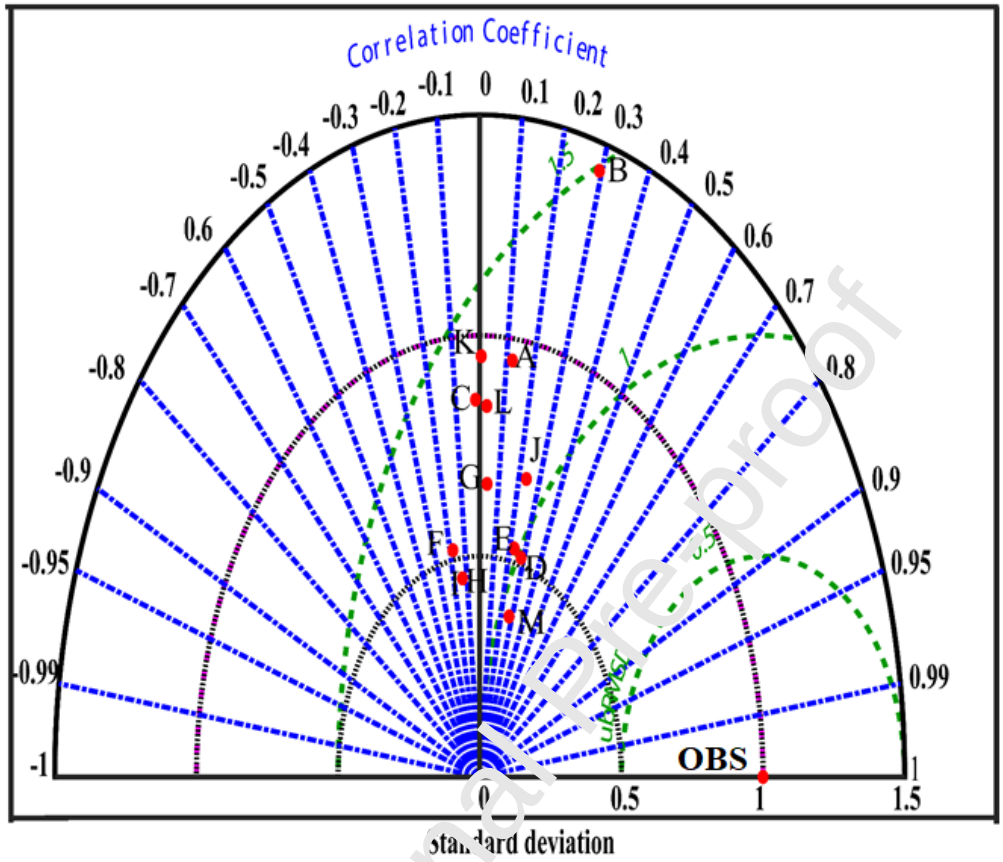

Fig. 5. Taylor diagram showing performance of CMIP5 models against station data at annual scale. The alphabetical letters represent the models; **A**- BCC-CSM1.1-M, **B** - CanESM2, **C**- CSIRO-Mk3-6-0, **D** - EC-EARTH, **E**- FGOALS-g2, **F**- GFDL-ESM2G, **G** - GISS-E2-R, **H**- HADGEM2-AO, **I** - IPSL-CM5A-MR, **J**- MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L**- MRI-CGCM3, and **M** - MME model.

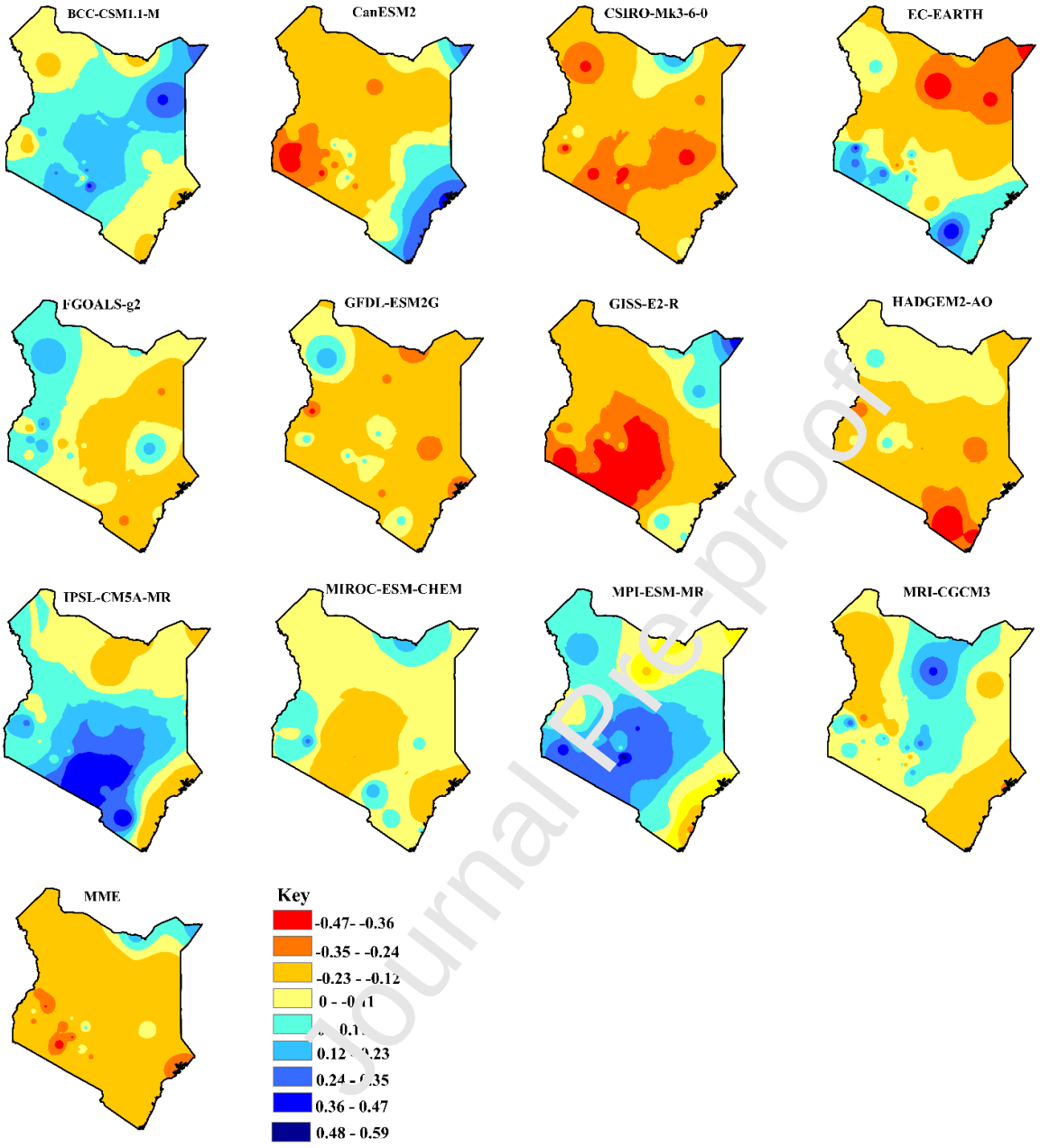

Fig. 6. Spatial distribution of temporal agreement between observed MAM rains and CMIP5 models at each station from 1979-2005 over Kenya.

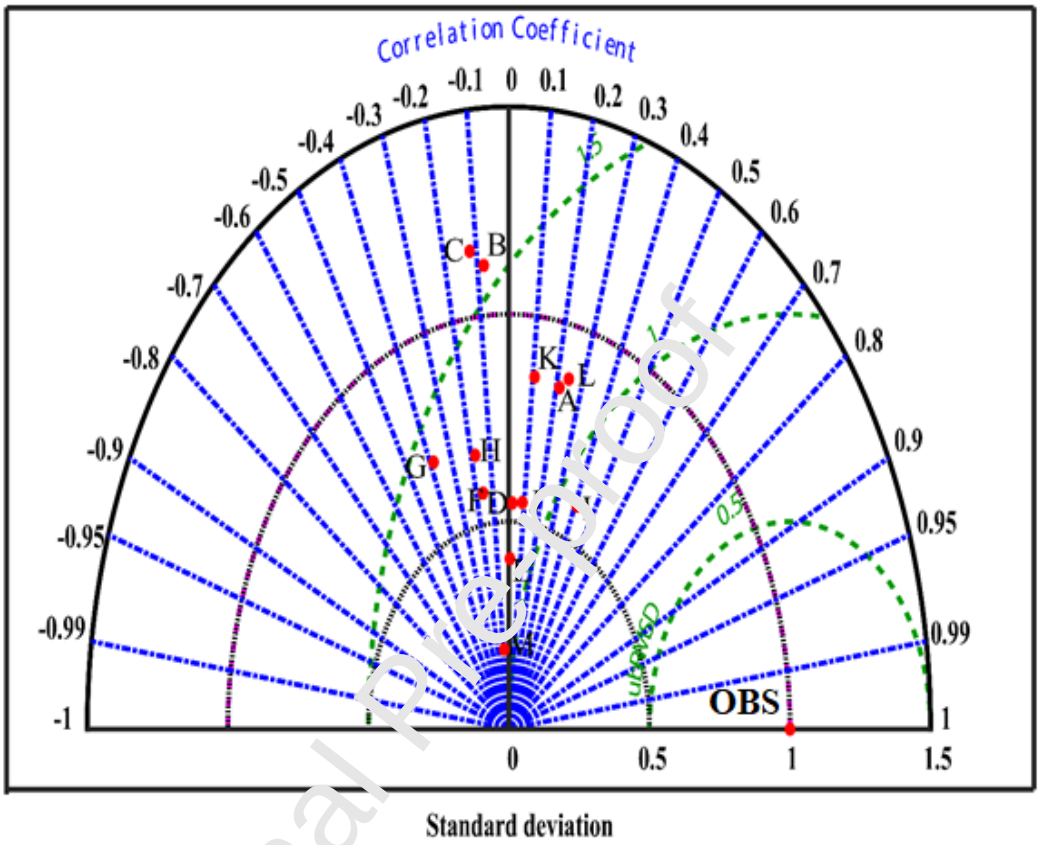

Fig.7. Fig. caption same as Fig. 5, but during MAM rainy season.

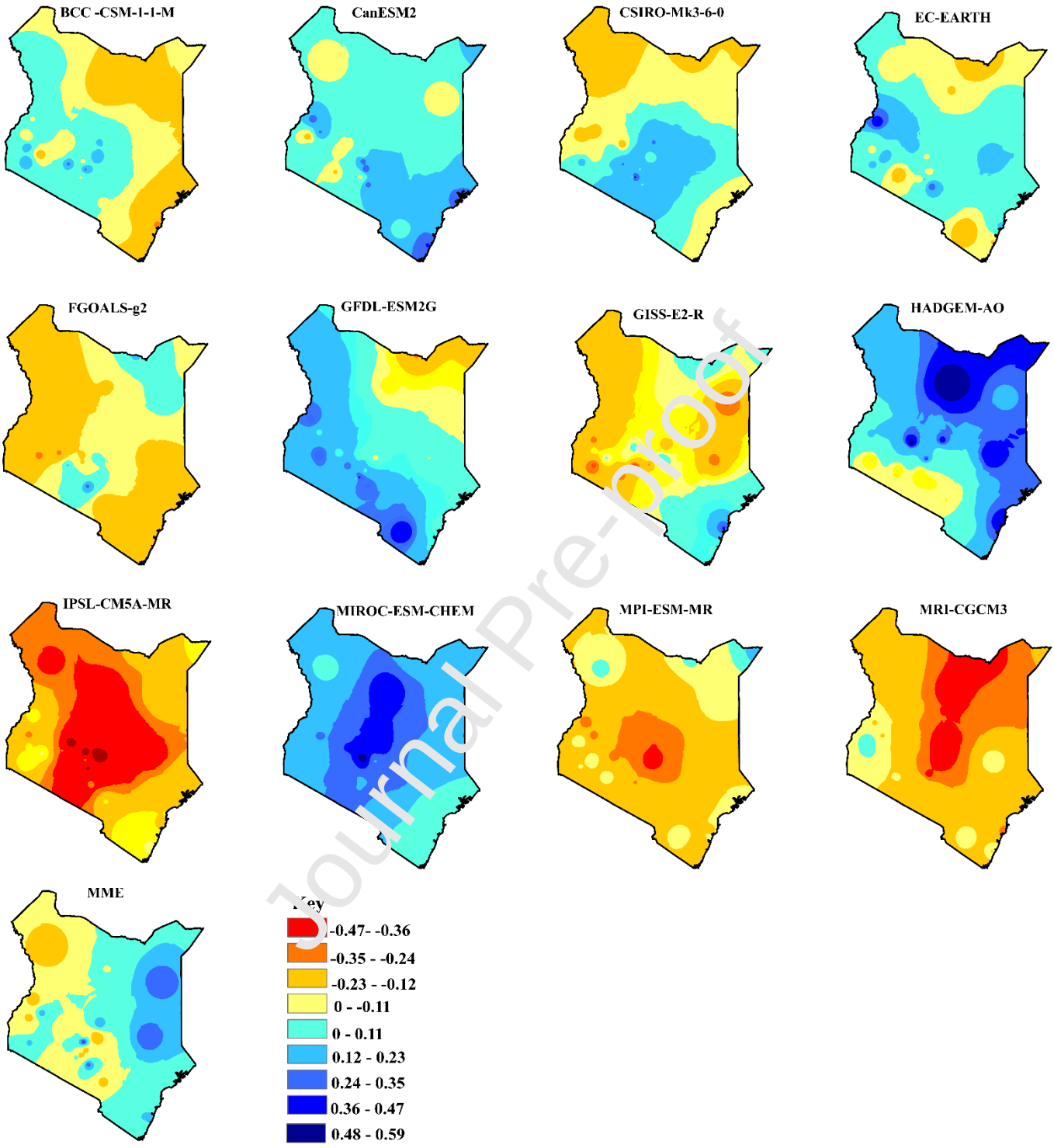

Fig. 8. Spatial distribution of temporal agreement between observed OND rains and CMIP5 models at each station from 1979-2005 over Kenya.

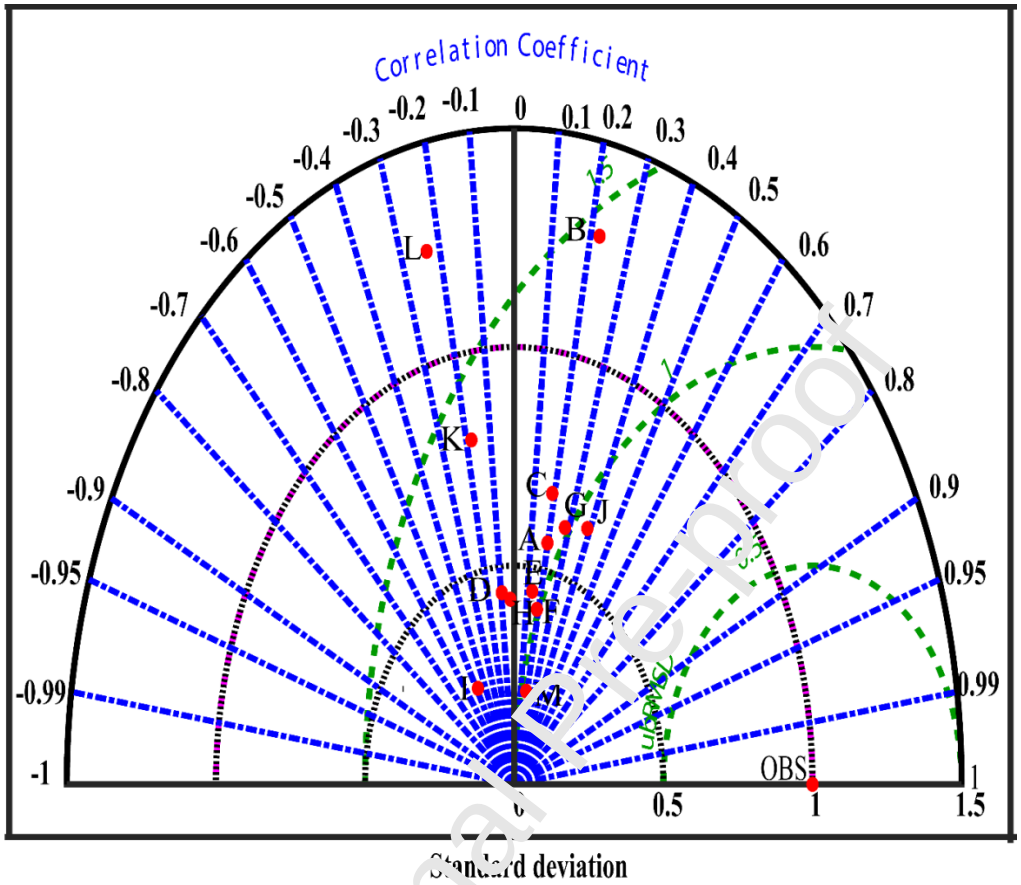

Fig. 9. Fig. caption same as Fig. 5, but during OND rainy season.

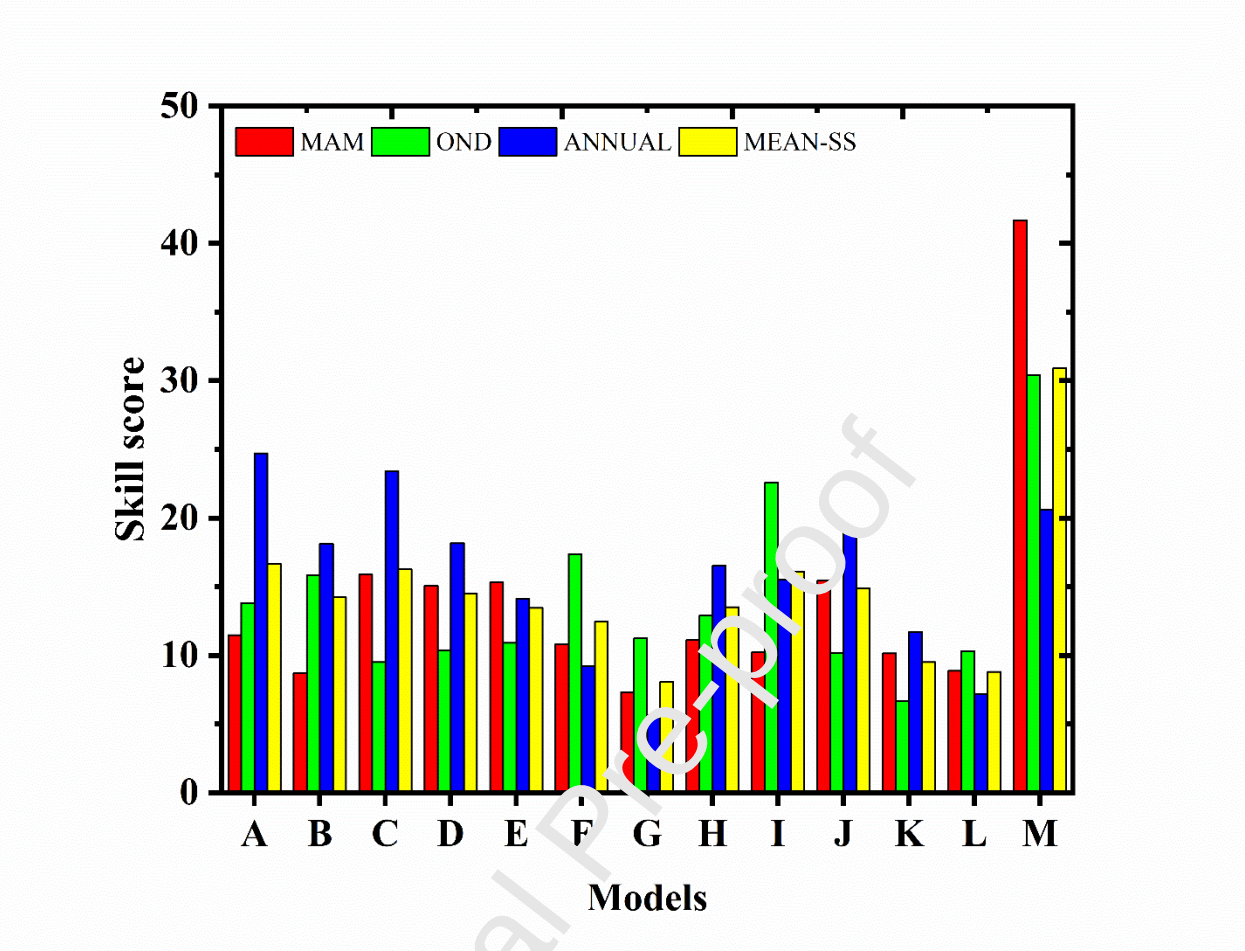

Fig. 10. Skill score of CMIP5 models during MAM, OND, and annual rainfall. Mean skill (Mean-SS) is in yellow color. The alphabetical letters represent the models; **A** - BCC-CSM1.1-M, **B** - CanESM2, **C**- CSIRO Mk3-6-0, **D** - EC-EARTH, **E**- FGOALS-g2, **F**- GFDL-ESM2G, **G**- GISS-E2-R, **H**- HADGEM2-AO, **I**- IPSL-CM5A-MR, **J**- MIROC-ESM-CHEM, **K**- MPI-ESM-MR, **L**-MRI-CGCM3, and **M**- MME model.

Table 1. The CMIP5 models used, their greenhouse gas composition, and the modeling center

Model Abbreviation	Resolution	Composition Elements	Modelling Centre
BCC-CSM1-1-M	$2.8125^0 \times 2.8125^0$	Nat Ant GHG SD Oz Sl Vl SS Ds BC OC	Beijing Climate Center (BCC), China Meteorological Administration, China
CanESM2	$2.8^0 \times 2.8^0$	GHG, Oz, SA, BC, OC, LU, Sl, Vl (GHG includes CO2, CH4, N2O, CFC11, effective CFC12)	Canadian Centre for Climate Modeling and Analysis (Canada)
CSIRO-Mk3-6-0	$1.875^0 \times 1.875^0$	Ant, Nat (all forcing)	Australian Commonwealth Scientific and Industrial Research Organization (CSIRO) Marine and Atmospheric Research
EC-EARTH	$1.125^0 \times 1.12^0$	Nat, Ant	EC-Earth (European Earth System Model)
FGOALS-g2	$2.8125^0 \times 2.8125^0$	GHG, Oz, SA, BC, Ds, OC, SS, Sl, Vl (GHG includes CO2, CH4, N2O, CFC11, effective CFC12. Aerosol also includes sulfate	IAP (Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China) and THU (Tsinghua University)
GFDL-ESM2G	$2.5^0 \times 2.0^0$	GHG, SD, Oz, LU, Sl, Vl, SS, BC, MD, OC (GHG includes CO2, CH4, N2O, CFC11, CFC12, HCFC22, CFC113	NOAA GFDL (201 Forrestal Rd, Princeton, NJ, 08540)
GISS-E2-R	$2.5^0 \times 2^0$	GHG, LU, Sl, Vl, BC, OC, SA, Oz (also includes orbital change - BC on snow - Nitrate aerosols	NASA/GISS (Goddard Institute for Space Studies) New York, NY
HadGEM2-AO	$1.25^0 \times 1.9^0$	Nat, Ant, GHG, SA, Oz, LU, Sl, Vl, SS, Ds, BC, MD, OC	NIMR (National Institute of Meteorological Research, Seoul, South Korea)
IPSL-CM5A-MR	$2.5^0 \times 1.258^0$	Nat, Ant, GHG, SA, Oz, LU, SS, Ds, BC, MD, OC, AA	IPSL (Institut Pierre Simon Laplace, Paris, France)
MIROC-ESM-CHEM	$2.8125^0 \times 2.8125^0$	GHG, SA, Oz, LU, Sl, Vl, MD, BC, OC (Ozone is predicted)	AMSTEC (Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan), AORI (Atmosphere and Ocean Research Institute, The University of Tokyo,

				Chiba, Japan), and NIES (National Institute for Environmental Studies, Ibaraki, Japan
MPI-ESM-MR	1.875^0 1.875^0	×	GHG, Oz, SD, Sl, Vl, LU	Max Planck Institute for Meteorology
MRI-CGCM3	1.125^0 2.250^0	×	GHG, SA, Oz, LU, Sl, Vl, BC, OC (GHG includes CO2, CH4, N2O, CFC-11, CFC-12, and HCFC-22)	MRI (Meteorological Research Institute, Tsukuba, Japan)

Table 2. Annual statistics for CMIP5 models and observation.

Models	μ	σ	Med	MK	SSE	Skw	Krts	KGE`	Corr	%bias	nRMSE	VI
Observation	953	153	928	-0.04	-0.9	0.7	0.4	1.0	1.0	0.00	0.0	0.0
Ensemble	701	58	692	0.24	2.0	0.6	-0.5	-1.6	0.3	-26.5	0.3	5.1
Bcc-csm1.1-m	1097	145	1095	0.28	7.9	0.3	-0.5	0.1	0.1	15.1	0.2	0.0
CanESM2	833	219	833	0.19	5.8	0.4	-1.0	0.2	0.3	-12.6	0.3	0.5
Csiro-mk3-6-0	581	130	560	0.29	7.2	0.2	-1.1	-0.1	-0.02	-39.1	0.5	0.1
Ec-earth	678	79	668	0.23	3.4	0.2	-1.1	0.1	0.3	-28.9	0.4	2.0
Fgoals-g2	974	81	961	0.16	2.1	0.2	-1.3	0.1	0.2	2.2	0.2	1.8
Gfdl-esm2g	521	80	540	-0.11	-0.6	-0.7	-0.1	-0.4	-0.2	-45.4	0.5	1.9
Giss-e2-r	1181	101	1178	-0.04	-0.6	0.7	0.1	-0.04	0.04	23.9	0.3	0.7
Hadgem2-ao	368	69	350	0.15	1.26	0.1	0.0	-0.4	-0.1	-60.4	0.4	3.1
Ipsl-cm5a-mr	368	69	350	0.14	1.25	0.1	0.0	-0.4	-0.1	-62.9	0.6	3.1
Miroc-esm-chem	594	106	583	0.24	4.9	0.2	-1.0	0.1	0.3	-37.6	0.3	0.6
Mpi-esm-mr	526	128	439	-0.03	-0.03	-0.1	-1.4	0.001	0.0	-9.1	0.3	0.0
Mri-cgcm3	867	145	833	0.06	2.51	0.0	-1.1	-0.1	0.0	-44.9	0.5	0.1

Table 3. MAM seasonal rainfall statistics for CMIP5 models and observation. The values in bold depicts significant correlation values at $\alpha \leq 0.1$.

Models	μ	σ	Med	Mk	SSE	Skw	Krts	KGE`	Corr	%bias	nRMSE	VI
Observation	390	80.3	394.2	-0.17	-2.4	-0.4	-0.3	1.0	1.0	0.0	0.0	0.0
Ensemble	184	15.5	182.1	0.27	0.67	0.5	-0.4	-0.5	-0.08	-52.7	0.6	25
Bcc-csm1.1-m	385	67.7	384.5	0.24	2.8	0.5	-0.1	0.2	0.2	-1.1	0.2	0.12
CanESM2	235	90.1	207.5	-0.09	-0.5	0.6	-0.8	-0.2	-0.1	-39.1	0.6	0.05
Csiro-mk3-6-0	269	93.2	279.7	0.19	2.0	0.1	-0.6	-0.2	-0.1	-31.0	0.6	0.09
Ec-earth	203	43.8	203.4	0.12	0.8	0.2	0.0	-0.2	0.001	-47.8	0.6	1.66
Fgoals-g2	282	33.0	284.3	0.05	0.4	-0.4	0.2	-0.2	0.001	-27.5	0.3	4.32
Gfdl-esm2g	124	46.3	122.6	0.15	0.2	0.3	-0.6	-0.4	-0.2	-68.0	0.7	1.34
Giss-er-2	342	56.0	333.2	-0.13	-1.4	0.3	-0.2	-0.4	**-0.4**	-12.2	0.3	0.54
Hadgem2-ao	172	53.9	163.5	0.11	0.93	0.6	-0.2	-0.3	-0.2	-55.7	0.7	0.67
Ipsl-cm5a-mr	135	47.4	124.8	0.1	0.2	1.3	1.7	0.03	**0.4**	-65.1	0.7	1.22
Miroc-esm-chem	139	44.0	134.8	0.21	2.64	0.6	-0.3	-0.2	0.1	-64.1	0.6	1.63
Mpi-esm-mr	171	69.9	166.8	0.04	0.6	0.5	0.3	0.1	0.3	-56.1	0.7	0.08
Mri-cgcm3	178	68.6	169.5	0.16	2.14	0.3	-0.7	-0.1	0.1	-54.3	0.7	0.10

Table 4. OND seasonal rainfall statistics for the models and observation. The values in bold signifies significant correlation values at $\alpha \leq 0.1$.

Models	μ	σ	Med	MK	SS	Skw	Krts	KGE	Corr	%bias	nRMSE	VI
Observation	268	120	241	0.06	0.95	2.1	5.5	1.0	1.0			
Ensemble	315	26	306	0.35	1.7	0.3	-0.9	-0.1	0.2	38.8	0.4	19.3
Bcc-csm1.1-m	404	67	401	0.29	2.8	-0.1	-0.8	0.0	0.2	50.7	0.6	1.5
CanESM2	383	154	352	0.29	6.5	0.5	-1.1	0.1	0.2	42.9	0.5	0.3
Csiro-mk3-6-0	226	81	227	0.32	1.0	0.3	-0.3	0.1	0.2	-15.7	0.5	0.6
Ec-earth	280	53	278	0.11	1.5	0.2	-0.4	0.03	0.1	4.7	0.5	3.2
Fgoals-g2	372	53	367	0.13	1.5	0.6	-0.5	-0.3	-0.1	10.5	0.6	3.4
Gfdl-esm2g	296	49	296	-0.23	-1.8	-0.2	-0.6	0.0	0.2	48.1	0.5	4.2
Giss-er-2	397	51	390	0.13	1.2	0.5	-0.7	-0.3	-0.03	23.3	0.6	3.8
Hadgem2-ao	330	73	327	-0.16	-2.2	0.3	-0.9	0.2	0.3	-52.9	0.4	1.1
Ipsl-cm5a-mr	126	30	122	0.09	0.6	0.4	0.3	-0.7	**-0.5**	5.3	0.7	13.9
Miroc-esm-mr	282	76	281	0.23	3.0	0.2	-0.8	0.3	**0.4**	67.7	0.4	0.9
Mpi-esm-mr	234	96	212	-0.01	1.1	0.8	-0.6	-0.2	-0.2	-12.7	0.6	0.2
Mri-cgcm3	449	150	417	0.14	2.2	1.2	1.2	-0.4	-0.2	-1.1	0.8	0.2

Table 5. Statistical links between observed rainfall, and CMIP5 models with ENSO/IOD indices. The values in bold show significant correlation values $\alpha \leq 0.1$.

	ENSO			IOD		
Models	MAM	OND	Annual	MAM	OND	Annual
Observation	-0.04	**0.5**	0.3	0.05	**0.77**	**0.48**
Ensemble	-0.06	0.2	0.15	**0.4**	0.16	0.31
Bcc-csm1.1-m	-0.06	0.02	-0.07	0.27	0.22	0.3
CanESM2	-0.05	0.13	0.2	0.21	0.19	0.3
Csiro-mk3-6-0	-0.02	-0.09	-0.11	**0.43**	0.09	0.22
Ec-earth	0.2	-0.2	0.11	**0.3o**	0.06	0.28
Fgoals-g2	0.02	0.25	0.21	-0.13	0.03	-0.14
Gfdl-esm2g	0.29	0.04	0.14	0.0[illegible]	-0.01	-0.14
Giss-e2-r	0.16	-0.02	0.08	0.16	0.11	0.24
Hadgem2-ao	-0.22	0.13	-0.21	-0.04	0.1	**-0.35**
Ipsl-cm5a-mr	0.02	-0.23	-0.21	0.14	-0.32	**-0.35**
Miroc-esm-mr	-0.31	0.22	0.0[illegible]	0.27	**0.34**	0.27
Mpi-esm-mr	0.1	0.15	0.0[illegible]	-0.12	-0.15	0.01
Mri-cgcm3	-0.23	0.01	-0.13	0.02	-0.02	0.1

Highlights

- Various ranking metrics on CMIP5 models leads to different types of models' selection.
- Models performance differ from one locale to another and even within different climate variables.
- Evaluation of CMIP5 models should be improved to identification of links with teleconnection patterns rather than concentrating on reproducing mean rainfall statistics only.
- Failure of a model should be based on the intended use of the models.
- Caution should be exercised when using CMIP5 models' projection over Kenya.