Retrospective Analysis and Bayesian Model Averaging of CMIP6 Precipitation in the Nile River Basin

Mohammed Ombadi[1], Phu Nguyen[1], Soroosh Sorooshian[1], and Kuo-lin Hsu[1]

[1]University of California, Irvine

November 26, 2022

Abstract

The Nile river basin is one of the global hotspots vulnerable to climate change impacts due to fast growing population and geopolitical tensions. Previous studies demonstrated that general circulation models (GCMs) frequently show disagreement in the sign of change in annual precipitation projections. Here, we first evaluate the performance of 20 GCMs from the 6 Coupled Model Intercomparison Project (CMIP6) benchmarked against a high spatial resolution precipitation dataset dating back to 1983 from Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR). Next, a Bayesian Model Averaging (BMA) approach is adopted to derive probability distributions of precipitation projections in the Nile basin. Retrospective analysis reveals that most GCMs exhibit considerable (up to 64% of mean annual precipitation) and spatially heterogenous bias in simulating annual precipitation. Moreover, it is shown that all GCMs underestimate interannual variability; thus, the ensemble range is under-dispersive and a poor indicator of uncertainty. The projected changes from the BMA model show that the value and sign of change varies considerably across the Nile basin. Specifically, it is found that projected change in the two headwaters basins, namely Blue Nile and upper White Nile is 0.03% and -1.65% respectively; both statistically insignificant at 0.05. The uncertainty range estimated from the BMA model shows that the probability of a precipitation decrease is much higher in the upper White Nile basin whereas projected change in the Blue Nile is highly uncertain both in magnitude and sign of change.

Retrospective Analysis and Bayesian Model Averaging of CMIP6 Precipitation in the Nile River Basin

Mohammed Ombadi[1], Phu Nguyen[1], Soroosh Sorooshian[1] and Kuo-lin Hsu[1]

[1]Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California Irvine.

Corresponding author: Mohammed Ombadi (mombadi@uci.edu)

Key Points:

- The performance of a 20 CMIP6 GCMs ensemble is evaluated using a high spatial resolution, long-record precipitation dataset.
- GCMs exhibit a considerable bias, and they underestimate interannual variability in simulating annual precipitation.
- Projected change varies across the basin with a mean of 0.03% and -1.65% in the Blue Nile and upper White Nile basins respectively.

Abstract

The Nile river basin is one of the global hotspots vulnerable to climate change impacts due to fast growing population and geopolitical tensions. Previous studies demonstrated that general circulation models (GCMs) frequently show disagreement in the sign of change in annual precipitation projections. Here, we first evaluate the performance of 20 GCMs from the 6th Coupled Model Intercomparison Project (CMIP6) benchmarked against a high spatial resolution precipitation dataset dating back to 1983 from Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR). Next, a Bayesian Model Averaging (BMA) approach is adopted to derive probability distributions of precipitation projections in the Nile basin. Retrospective analysis reveals that most GCMs exhibit considerable (up to 64% of mean annual precipitation) and spatially heterogenous bias in simulating annual precipitation. Moreover, it is shown that all GCMs underestimate interannual variability; thus, the ensemble range is under-dispersive and a poor indicator of uncertainty. The projected changes from the BMA model show that the value and sign of change varies considerably across the Nile basin. Specifically, it is found that projected change in the two headwaters basins, namely Blue Nile and upper White Nile is 0.03% and -1.65% respectively; both statistically insignificant at $\alpha = 0.05$. The uncertainty range estimated from the BMA model shows that the probability of a precipitation decrease is much higher in the upper White Nile basin whereas projected change in the Blue Nile is highly uncertain both in magnitude and sign of change.

1 Introduction

The Nile river basin constitutes approximately 10% of the African continent (Swain, 2008) extending across eleven countries. A total population of 462 million in these countries is growing at an annual growth rate of 2.5%, faster than the average global growth rate estimated at 1.1%. Consequently, the population of these countries is projected to reach 836 million (81% increase) by the year 2050 (The World Bank, 2018; 2020). A key challenge, therefore, that face these countries is to sustain the burgeoning food and energy demand of this growing population. Water lies at the heart of natural resources that play a pivotal role in securing this demand. Therefore, assessment of climate change impacts on precipitation is important due to its direct effect on water availability in headwaters countries as well as its impact on the Nile streamflow yield which is the main source of water for riparian countries, namely Sudan and Egypt.

Several studies have been devoted to the assessment of climate change impacts on precipitation in the Nile River basin and its headwaters basins (e.g. Conway, 1996; Yates & Strzepek, 1996; 1998, Kim & Kaluarachchi, 2009; Elshamy et al, 2009; Bhattacharjee & Zaitchik, 2015; Fenta Mekonnen & Disse, 2018). Earlier studies found that general circulation models (GCMs) frequently show disagreement in the sign of change of annual precipitation projections. For instance, Conway (1996) used 3 GCMs to assess climate change impact on precipitation in the Blue Nile and Lake Victoria sub-basins; results showed that percentage change in precipitation ranges from -1.9% to 7.4% in the two sub-basins. More recently, Kim & Kaluarachchi (2009) showed that mean annual precipitation in the upper Blue Nile sub-basin is projected to increase by 11% based on a weighted average of 6 GCMs outcomes. On the contrary, Elshamy (2009) reported the outcomes of 17 GCMs and showed that projected change in mean annual precipitation in the upper Blue Nile sub-basin ranges from -15% to +14% with more models reporting a decrease in precipitation. These results, among others, emphasize that there is a wide uncertainty and inter-model differences in precipitation projections, and they indicate that a consensus on how climate change will impact water resources in the Nile basin is yet to be reached.

Two different approaches are commonly adopted to treat uncertainty of GCMs. At one end of the spectrum is the ensemble mean which overlooks historical performance of the models and assigns equal weights to all models. At the other end, there is an approach that selects a number of best performing models and discards the remaining ones. The former is less accurate at regional scales and in cases where there is a spread in model projections (Schaller et al, 2011) whereas the latter is highly dependent on the specific metrics used for performance evaluation (Schaller et al, 2011; Bhattacharjee & Zaitchik, 2015). Between these two extremes lies the approach of model averaging in which models are neither weighted equally nor some of them are discarded entirely. Specifically, model averaging methods take advantage of retrospective analysis of GCMs simulations benchmarked against observations, and they assign weights to models according to their performance. A major issue, however, that lessen the effectiveness of such methods is the dearth of quality controlled, dense gauge precipitation observations in the Nile basin. Here, we surmount this issue by resorting to high spatial resolution and long record of historical observations provided from Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR; Ashouri et al., 2015). PERSIANN-CDR is a high spatial resolution satellite-based dataset that is bias adjusted using gauge observations at the monthly scale; thus, providing a unique dataset for retrospective analysis of GCMs.

To this end, the focus of the present study is to first evaluate the performance of 20 GCMs from the 6th Coupled Model Intercomparison Project (CMIP6) against PERSIANN-CDR over the Nile basin. Next, a model averaging approach, namely Bayesian Model Averaging (BMA; Raftery et al., 2005) is implemented to derive probability distributions of precipitation projections in the Nile basin for the future period (2015 – 2100). The remainder of this paper is organized as follows. Section 2 provides a brief description of the data used in this study. Section 3 describes the bias adjustment and Bayesian model averaging approaches used to postprocess CMIP6 GCMs precipitation projections. Section 4 presents the results of retrospective analysis as well as the future projections of precipitation in the Nile basin. Finally, section 5 sums up the findings of the study and draws conclusions.

2 Data and Study Area

2.1 CMIP6

Many climate models participating in CMIP6 have reported their simulations for the different CMIP6 experiments. In the present study, two experiments are of concern: historical and the Shared Socioeconomic Pathways (SSP) SSP5-8.5. The historical experiment provides GCMs simulations for the period (1850-2014), and it is intended to be used for assessment of model performance in simulating historical observations. Here, we only use data from the period 1983 onward to be consistent with the available record of observed precipitation from PERSIANN-CDR. SSP5-8.5 is the future scenario that corresponds to high greenhouse gas emissions, and it is the equivalent to RCP8.5 “business as usual” scenario in CMIP5. Currently, a set of 20 models have reported their simulations for both historical and SSP5-8.5 experiments. These models have been used in this study to examine climate change impact on precipitation in the Nile basin, and they are shown in Table 1. For each model, we only consider the first ensemble member for future projections under SSP5-8.5. Also, we consider dataset at monthly temporal resolution for both historical and SSP5-8.5.

2.2 PERSIANN-CDR

PERSIANN-CDR (Ashouri et al., 2015; see also Nguyen et al., 2018) is a satellite-based precipitation dataset based on infrared (IR) imagery. It has near-global coverage (60°S - 60°N) with a spatial resolution of 0.25° x 0.25° and a daily temporal resolution. PERSIANN-CDR is suitable for climatic studies because of its long record of +37 years (1983 – delayed present). It is particularly advantageous because it is bias adjusted using Global Precipitation Climatology Project (GPCP) monthly 2.5° x 2.5° precipitation data (Adler et al., 2018). Therefore, it maintains monthly precipitation at 2.5° x 2.5° that is consistent with GPCP while capturing spatial rainfall variability at higher spatial resolution. This last point emphasizes that PERSIANN-CDR has sufficient credibility to be used as a baseline dataset for evaluation of CMIP6 GCMs. PERSIANN-CDR is widely used for a range of hydrologic and hydroclimatic studies (e.g. Ombadi et al., 2018; Nguyen et al., 2020), and it has previously been used for evaluation of GCMs (Nguyen et al., 2017). Here, we use PERSIANN-CDR at monthly temporal accumulations.

2.3 Study Area

In this study, we consider the entire Nile basin for our analysis (shown in Figure 1; gray lines). The analysis is performed at the grid scale (1° x 1°) due to the wide variability of climate and precipitation regimes in the Nile basin. This variability is clearly shown in Figure 1 with the south-to-north gradient in precipitation which represents the variability in climate from tropical humid in the south to hyper arid in the north. Throughout this study, we carry out the analysis at the grid scale and then aggregate the results at the entire Nile basin as well as its headwaters basins, namely the Blue Nile and upper White Nile basins (purple dashed lines in Figure 1). We focus on these two basins due to their significant contribution to the Nile streamflow yield.

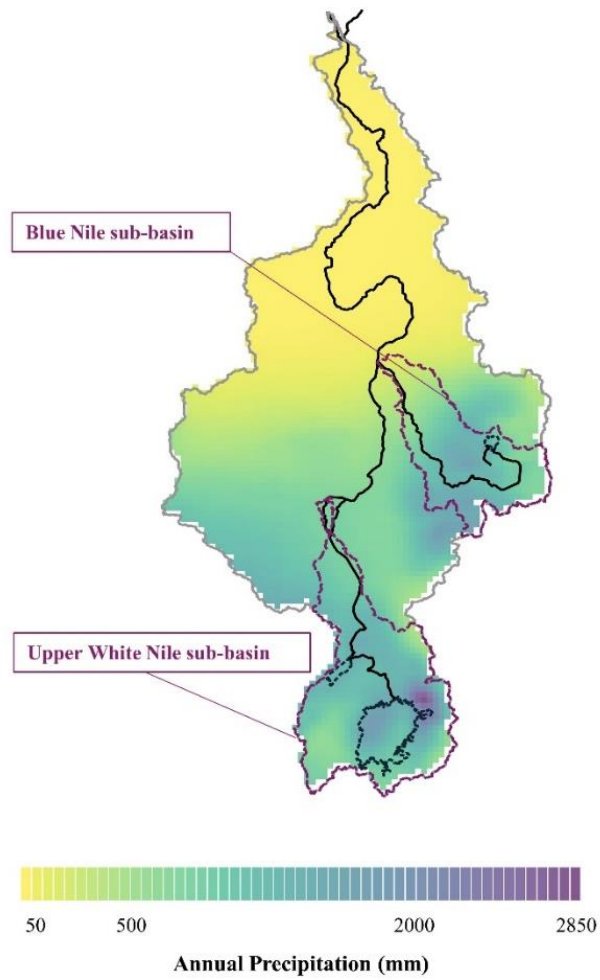

Figure 1. The Nile river basin (gray line) and its headwaters basins, namely the Blue Nile and upper White Nile sub-basins (dashed purple line). The Nile river and its tributaries are shown in solid balck line. Mean annual precipitation is computed from PERSIANN-CDR for the period (1983-2014).

Table 1. CMIP6 models used in the present study and their spatial resolution.

Model	Institute	Resolution (Lat° x Lon°)
Earth3	EC-Earth-Consortium, Europe	0.702 x 0.703 *
Earth3-Veg	EC-Earth-Consortium, Europe	0.702 x 0.703 *
MPI-ESM1-2-HR	Deutsches Klimarechenzentrum, Germany	0.935 x 0.9375 *
CESM2-WACCM	National Center for Atmospheric Research (NCAR), USA	0.942 x 1.25
FIO-ESM-2-0	First Institute of Oceanography-Qingdao National Laboratory for Marine Science and Technology (FIO-QLNM), China	0.942 x 1.25
NorESM2-MM	NorESM Climate modeling Consortium (NCC), Norway	0.942 x 1.25
FGOALS-f3-L	Chinese Academy of Sciences, China	1 x 1.25
BCC-CSM2-MR	Beijing Climate Center, China	1.121 x 1.125 *
MIROC6	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, The University of Tokyo, National Institute for Environmental Studies, and RIKEN Center for Computational Science (MIROC), Japan	1.4 x 1.406 *
ACCESS-CM2	Commonwealth Scientific and Industrial Research Organisation-Australian Research Council Centre of Excellence for Climate System Science (CSIRO-ARCCSS), Australia	1.25 x 1.875
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia	1.25 x 1.875
KAGE-1-0-G	National Institute of Meteorological Sciences/Korea Meteorological Administration (NIMS-KMA), Republic of Korea	1.25 x 1.875
INM-CM4-8	Institute for Numerical Mathematics, Russia	1.5 x 2
INM-CM5-0	Institute for Numerical Mathematics, Russia	1.5 x 2
IPSL-CM6A-LR	Institut Pierre Simon Laplace, France	1.268 x 2.5
MPI-ESM1-2-LR	Max Planck Institute for Meteorology, Germany	1.865 x 1.875 *
NESM3	Nanjing University of Information Science and Technology, China	1.865 x 1.875 *
FGOALS-g3	Chinese Academy of Sciences, China	2.279 x 2 *
NorESM2-LM	NorESM Climate modeling Consortium (NCC), Norway	1.895 x 2.5
CanESM5	Canadian Centre for Climate Modelling and Analysis, Canada	2.789 x 2.813 *

***** Approximate resolution since the native resolution is not in regular grids.

3 Methods

3.1 Bias Adjustment

CMIP6 model simulations and PERSIANN-CDR data were first re-gridded to a common spatial resolution of (1° x 1°) using bilinear interpolation. Bias adjustment coefficients were then calculated for each grid from the historical simulations (1983-2014) according to the following linear model:

$$\boldsymbol{y^H} = a + b * \boldsymbol{f}_{\boldsymbol{k}}^{H} \tag{1}$$

Where $\boldsymbol{y^H}$ is PERSIANN-CDR annual (or monthly) precipitation time series at a given grid for the period (1983-2014), and $\boldsymbol{f_k^H}$ is the corresponding annual precipitation from the $\boldsymbol{k}^{\text{th}}$ GCM model; the superscript $\boldsymbol{H}$ refers to "historical" whereas a and b are the bias adjustment coefficients.

3.2 Bayesian Model Averaging (BMA)

Bayesian model averaging (BMA; Raftery et al., 2005; see also Duan et al., 2007; Ajami et al., 2007) aims to reduce multi-model uncertainty by assigning weights to all models, with the weights representing posterior probabilities of the models given historical observations. BMA has previously been used to derive probability distributions of continental precipitation and temperature projections from a CMIP3 multi-model ensemble (Duan & Philips, 2010). The BMA predictive distribution is a weighted sum of conditional probability distributions of individual models. Let's consider the same notations used earlier and denote by $\boldsymbol{f_k}$ annual (or monthly) precipitation projections of the $\boldsymbol{k}^{\text{th}}$ model. BMA yields the following predictive model:

$$p(\boldsymbol{y}|\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_K}) = \sum_{k=1}^{K} w_k \, p_k(\boldsymbol{y}|\boldsymbol{f_k}) \tag{2}$$

Where $\boldsymbol{y}$ is the sought-after precipitation projections. The left-hand side represents the probability density function (pdf) of the BMA model which is equal to a weighted sum of the individual conditional pdfs of models 1, 2, …, K. As noted earlier, the weights w_k represent posterior probabilities of models conditioned on historical observations; thus, they sum to 1. The pdfs p_k for k= 1, 2, …, K are commonly assumed to be normal distributions. The weights w_k are estimated by maximizing the log-likelihood function of the pdf in the left-hand side using historical observations. Put simply, $\boldsymbol{y^H}$ and $\boldsymbol{f_k^H}$ are substituted for $\boldsymbol{y}$ and $\boldsymbol{f_k}$ respectively in equation 2 in order to estimate w_k. Several techniques such as the expectation-maximization (EM) algorithm (Dempster et al., 1977) can be used to converge to a local maximum of the log-likelihood function. Here, we use a Differential Evolution – Markov Chain (DE-MC) algorithm (Ter Braak, 2006) to find the optimum values of w_k.

4 Results and Discussion

4.1 Evaluation of CMIP6 GCMs for the recent past (1983-2014)

4.1.1 Bias in annual precipitation

We first examine the performance of the different GCMs in simulating the mean value of annual precipitation for the baseline period (1983 – 2014). Figure 2 shows the bias in spatially averaged annual precipitation over the Nile, Blue Nile and upper White Nile basins for each GCM as well as the ensemble mean with respect to PERSIANN-CDR. There is a clear spread between the models with a bias range of (-430 – 389 mm), (-619 – 661 mm) and (-738 – 791 mm) in the Nile, Blue Nile and upper White Nile basins respectively; see Table 2. These biases are significant in terms of mean annual precipitation as they represent up to 64%, 61% and 64% in the three basins respectively. Although the ensemble mean reduces the biases, it fails to outperform the best performing model in the three basins.

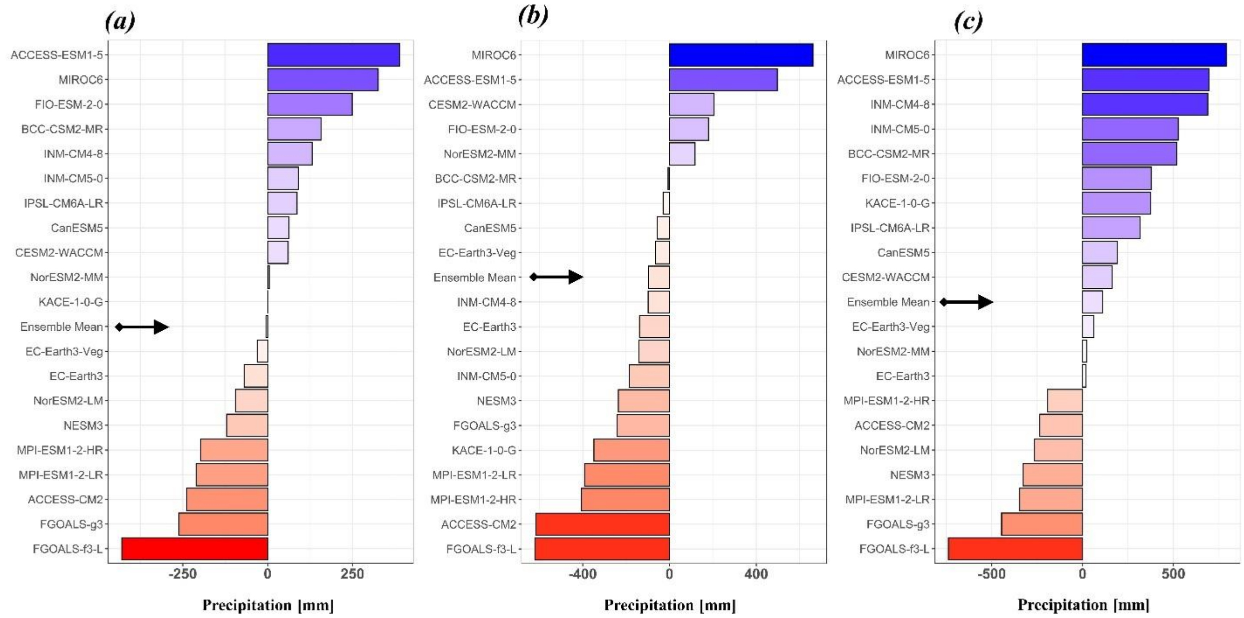

Figure 2. Barchart shows the bias in annual precipitation of the 20 CMIP6 GCMs and the ensemble mean with respect to PERSIANN-CDR in the period (1983 – 2014). Annual precipitation is spatially averaged over **(a)** the entire Nile basin, **(b)** the Blue Nile basin and **(c)** the upper White Nile basin. Black arrows point to ensemble mean.

Figure 3 shows the biases proportional to mean annual precipitation (i.e. relative bias) of the 20 GCMs in addition to the ensemble mean for each grid (1° x 1°) in the Nile basin. Apart from inter-model differences in bias, Figure 3 shows that there is a considerable spatial variability in bias within individual models. The values of relative bias over large areas of the basin exceed $\pm$ 0.3 (stapled grids in Figure 3) which underscore the importance of bias adjustment of GCM outputs prior to evaluation of future projections. In addition to examining the ability of GCMs in simulating the amount of total precipitation in the basins, it is important to investigate their accuracy in simulating the spatial patterns of precipitation. Table 2 shows the spatial correlation coefficient of the 20 GCMs and the ensemble mean against PERSIANN-CDR. This reflects how well each model represents the spatial variability of annual precipitation within the Nile basin and its two headwaters basins. Clearly, all the models fairly represent the spatial variability of annual precipitation within the Nile basin as evidenced by correlation coefficients greater than 0.8. Furthermore, the ability of the models to represent spatial variability within the Blue Nile basin is quite reasonable with a minimum correlation coefficient of 0.58. However, the correlation of spatial variability within the upper White Nile basin is drastically lower, with many models showing a negative correlation, and a maximum correlation coefficient of only 0.49. This highlights that while the GCMs performance in terms of bias is comparable in the Nile basin and its headwaters basins, the GCMs specifically underperform in the upper White Nile basin with regard to representation of spatial variability.

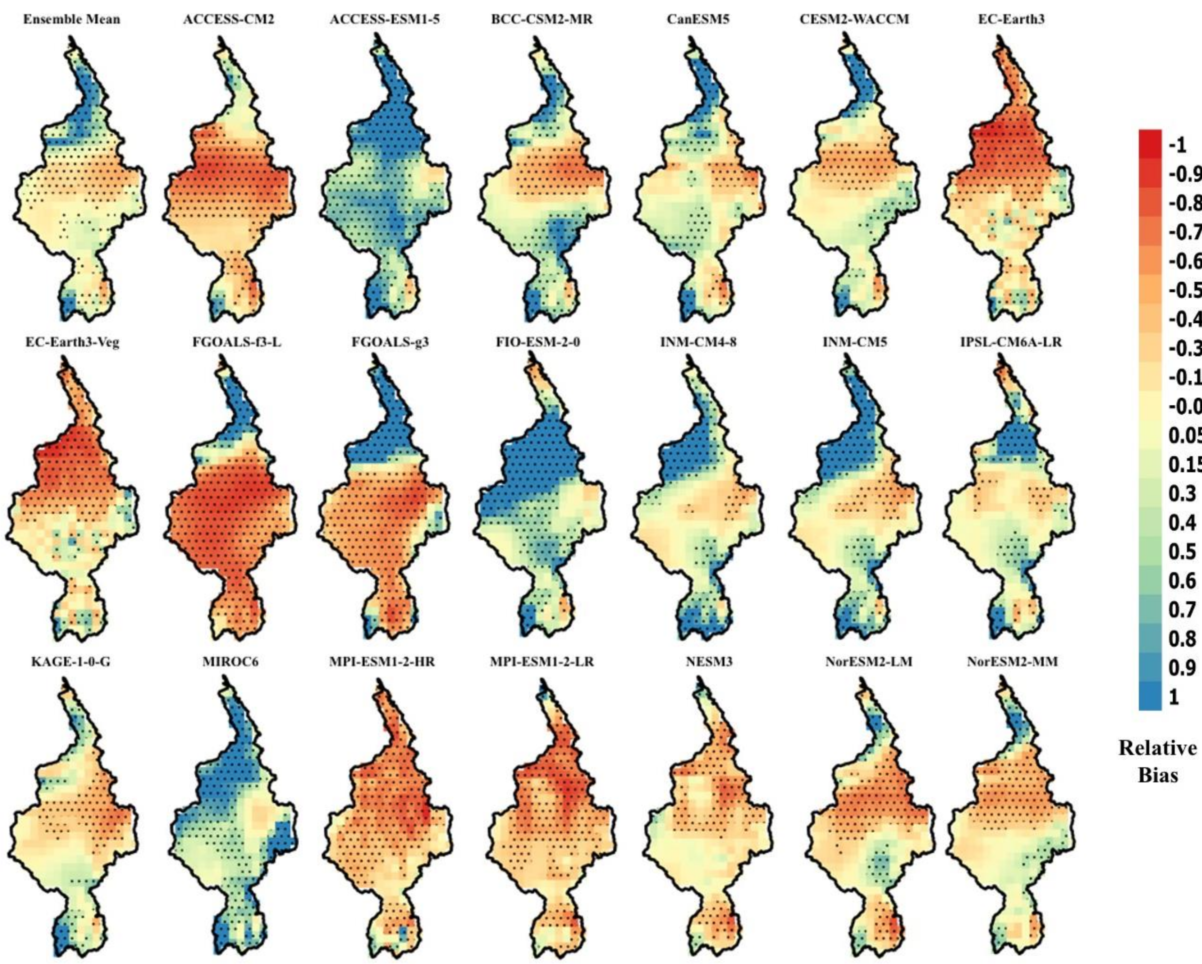

Figure 3. Maps show the relative bias of annual precipitaion during the baseline period (1983 – 2014) for each model of the 20 CMIP6 GCMs and the ensemble mean benchmarked against PERSIANN-CDR. Relative bias is calculated as the absolute bias (annual precipitation $_{\mathrm{GCM}}$ – annual precipitation $_{\mathrm{PERSIANN\text{-}CDR}}$) normalized by annual precipitation $_{\mathrm{PERSIANN\text{-}CDR}}$. Blue and red colors show overestimation and underestimation bias respectively. Stapled grids indicate values of relative bias > 0.3 or < -0.3.

4.1.2 Interannual variability and uncertainty

Figure 4a shows the annual precipitation coefficient of variation (ratio of standard deviation to mean) for the 20 GCMs and PERSIANN-CDR. Clearly, all models severely underestimate the interannual variability in the Nile basin and its headwaters basins. Specifically, the average coefficient of variation for the 20 GCMs is 4 to 7 times less than that of PERSIANN-CDR. Consequently, the bias adjusted ensemble of GCMs is under-dispersive which entails that the ensemble does not represent the true uncertainty in annual precipitation. This is demonstrated in Figure 4b which shows the rank histogram of PERSIANN-CDR with respect to the bias-adjusted GCMs ensemble for the period (1983 – 2014). If the ensemble truly captures the variability of annual precipitation, the rank histogram in the bins (2 - 19) should contain 19/21, or 90.5%, of PERSIANN-CDR values. Instead, the ensemble only contains 25% of PERSIANN-CDR observations. These results highlight that using bias adjusted GCMs will lead to underestimation in the uncertainty of precipitation projections. It will be shown later how this issue can be resolved using the Bayesian model averaging approach.

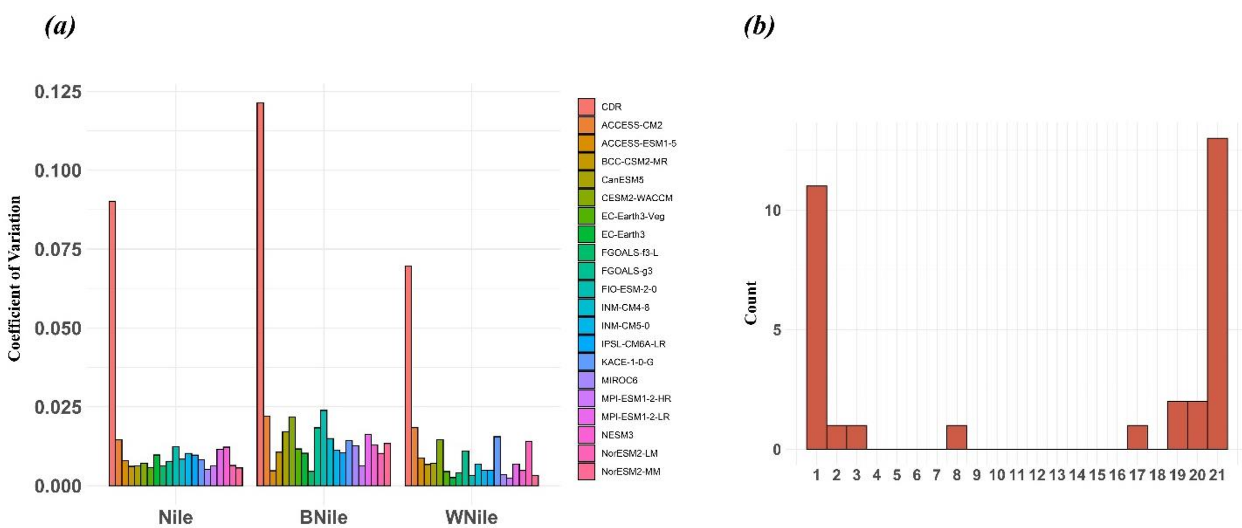

Figure 4. (a) Coefficient of variation of PERSIANN-CDR and the 20 GCMs annual precipitation for the period (1983-2014). **(b)** Rank histogram of PERSIANN-CDR annual precipitation for the period (1983-2014) with respect to the 20 GCMs.

4.1.3 Seasonal cycle

Here we evaluate the performance of the GCMs in capturing the seasonal cycle of precipitation. This is particularly important from the standpoint of assessing the hydrological impacts of climate projections which have consequences on Nile river flow and reservoir operations. Figure 5 shows the observed climatology monthly precipitation (red line) as well as the simulations of the 20 GCMs (black dashed lines) and their ensemble mean (solid black line). The two headwaters basins are characterized by distinct precipitation regimes; see Figures 5b and 5c. Specifically, precipitation in the Blue Nile sub-basin is monsoonal with pronounced seasonality (July – September) meanwhile upper White Nile sub-basin experiences two rainy seasons (March – May, October – December) (Conway, 2005). The seasonal cycle over the entire Nile basin, thereby, is a reflection of the cycles at the two headwaters basins; specifically, there is a major peak in (July – September) and a less pronounced one around (April – May). Despite overestimation and underestimation bias, the GCMs adequately capture the seasonal variability in precipitation. This is particularly apparent in the Nile and Blue Nile basins with the ensemble mean showing a correlation coefficient of 0.96 and 0.99 respectively in capturing the seasonal cycle; see Table 2. On the contrary, the GCMs are less capable of capturing the seasonal cycle over the upper White Nile basin; correlation coefficient of ensemble mean is equal to 0.71. Specifically, the ensemble mean overestimates the second rainy season (October – December).

Overall, there are numerous observations to be drawn from the retrospective analysis of GCMs simulations; however, two key findings are particularly worthy of consideration. First, the notion of a best performing model is very sensitive to the specific metric used for evaluation as well as the spatial domain of analysis. Table 2 shows the best performing model with respect to each metric (in bold font and an asterisk). Clearly, a different "best performing" model can be selected according to each metric and spatial domain. For instance, KAGE-1-0-G is the best performing model in terms of bias in annual precipitation over the entire Nile basin (bias= 1 mm) whereas NorESM2-MM is the best performing model in capturing the seasonal cycle of precipitation in the three basins. Second, the ensemble mean, although it provides adequate performance, does not outperform all individual models; this is clearly shown in Table 2. This pinpoints that the ensemble mean is sensitive to ensemble members at the end of the performance spectrum. It also underlines that analysis of future projections can benefit from advanced model averaging schemes that take into account retrospective model performance to provide an estimate that outperforms individual models.

Table 2. Evaluation of CMIP6 GCMs precipitation against PERSIANN-CDR over the entire Nile, Blue Nile (B Nile) and upper White Nile (W Nile) basins.

Model	Bias (mm)			Spatial correlation			Seasonality correlation		
	Nile	B Nile	W Nile	Nile	B Nile	W Nile	Nile	B Nile	W Nile
ACCESS-CM2	-239	-615	-238	0.81	0.82	†	0.94	0.97	0.55
ACCESS-ESM1-5	389	497	696	0.89	0.9	†	0.91	0.94	0.55
BCC-CSM2-MR	157	**-6 ***	517	0.87	0.88	†	0.96	**0.99 ***	0.75
CanESM5	63	-56	190	0.83	0.91	†	0.93	0.96	0.55
CESM2-WACCM	61	205	163	0.91	0.9	0.11	0.91	0.96	0.68
Earth3	-31	-65	61	0.9	0.79	**0.49 ***	0.88	**0.99 ***	0.5
Earth3-Veg	-68	-138	**18 ***	0.89	0.78	0.47	0.89	**0.99 ***	0.54
FGOALS-f3-L	-430	-619	-738	0.85	0.87	0.09	0.87	0.97	0.5
FGOALS-g3	-261	-242	-446	0.68	0.58	†	0.81	0.95	0.54
FIO-ESM-2-0	249	179	377	0.9	0.93	0	0.81	0.72	0.59
INM-CM4-8	131	-98	688	0.86	0.83	0.27	0.88	0.95	0.72
INM-CM5-0	90	-184	525	0.87	0.86	0.12	0.87	0.9	0.77
IPSL-CM6A-LR	87	-30	315	0.86	0.85	†	0.9	0.94	0.63
KAGE-1-0-G	**1***	-348	374	0.85	0.8	0.08	0.95	**0.99 ***	0.52
MIROC6	326	661	791	0.87	0.73	0.18	**0.98 ***	0.98	0.71
MPI-ESM1-2-HR	-198	-405	-193	0.85	0.84	0.22	0.9	0.93	0.76
MPI-ESM1-2-LR	-210	-391	-349	0.91	**0.95 ***	0.19	0.95	0.96	0.7
NESM3	-121	-235	-328	0.89	**0.95 ***	†	0.96	0.97	0.76
NorESM2-LM	-94	-140	-266	0.84	0.9	†	0.95	0.97	0.68
NorESM2-MM	5	118	21	0.92	0.9	0.27	**0.98 ***	**0.99 ***	**0.93 ***
Ensemble Mean	-5	-96	109	**0.92 ***	0.89	0.08	0.96	**0.99 ***	0.71

***** The best performing model according to the metric under consideration. † correlation coefficient is negative.

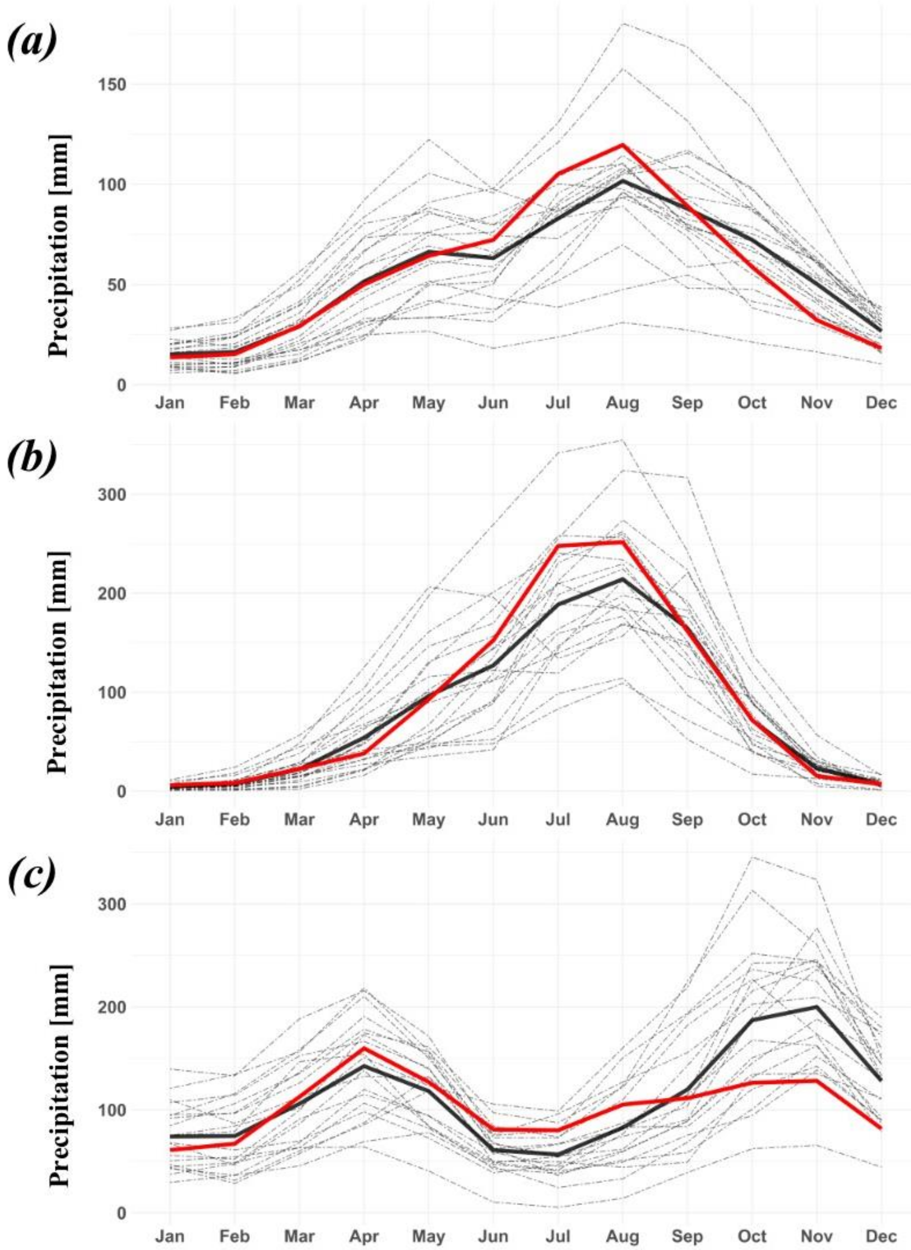

Figure 5. Climatology of monthly precipitation for the period (1983 – 2014) spatially averaged over: **(a)** the entire Nile basin, **(b)** the Blue Nile basin, and **(c)** the upper White Nile basin. The 20 CMIP6 GCMs are shown in thin black dashed lines. The ensemble mean is shown in solid black line whereas the observed precipitation from PERSIANN-CDR is shown in red.

4.2 Precipitation Projections for the period (2015-2100)

4.2.1 BMA mean precipitation projections

Here, we analyze mean precipitation projections obtained from the BMA model for the period (2015 - 2100) with respect to PERSIANN-CDR for the baseline period (1983 – 2014). Annual precipitation series of the 20 GCMs for the period (2015 – 2100) were first bias adjusted using the coefficients estimated from equation 1. Next, the BMA weights and their corresponding

BMA precipitation projections were computed from equation 2. These calculations were performed at the grid scale as opposed to the entire spatial domain due to the wide climatic variability and the different precipitation regimes in the Nile basin.

Figure 6a shows the percentage change in mean annual precipitation projected from BMA for the period (2015 – 2100) with respect to the baseline period (1983 – 2014). Clearly, there is spatial variability both in the sign and magnitude of change. A slight decrease in precipitation is observed in southern regions (the upper White Nile sub-basin) whereas the eastern regions (Blue Nile sub-basin) show both an increase and a decrease in precipitation. The statistically significant changes in precipitation ($\alpha = 0.05$) are observed over the riparian arid regions (stapled grids in Figure 6a) which have almost no impact on Nile streamflow. Specifically, there is a significant increase in precipitation in Northern Sudan, and a precipitation decrease to the northward. Figure 6b shows the projected changes in mean annual precipitation spatially averaged over the entire Nile basin from the 20 GCMs, ensemble mean and BMA. There is a spread in model projections with 14 models indicating an increase in mean annual precipitation and 6 models showing a decrease. Overall, percentage change in mean annual precipitation ranges from -1.7 % to 3.2 %. The BMA shows a statistically insignificant increase of 1.34 % (p-value = 0.2) (see Figure 6b and Table 3) compared to 0.82% from the ensemble mean.

Table 3. Projected changes in mean annual precipitation in the Nile, Blue Nile and Upper White Nile basins. In parentheses are the p-values of the projected changes.

Basin	***Ensemble mean***	***Best 3 models***	***BMA***	***BMA 90% Confidence Interval***	
				Lower (%)	***Upper (%)***
Nile	0.82 % (**0.3**)	-0.19% (**0.45**)	1.34 % (**0.2**)	-1.9	5.5
Blue Nile	0.43% (**0.42**)	-0.92% (**0.33**)	0.03% (**0.49**)	-6.8	7.2
upper White Nile	-0.45% (**0.36**)	0.17% (**0.44**)	-1.65% (**0.09**)	-6.9	1.9

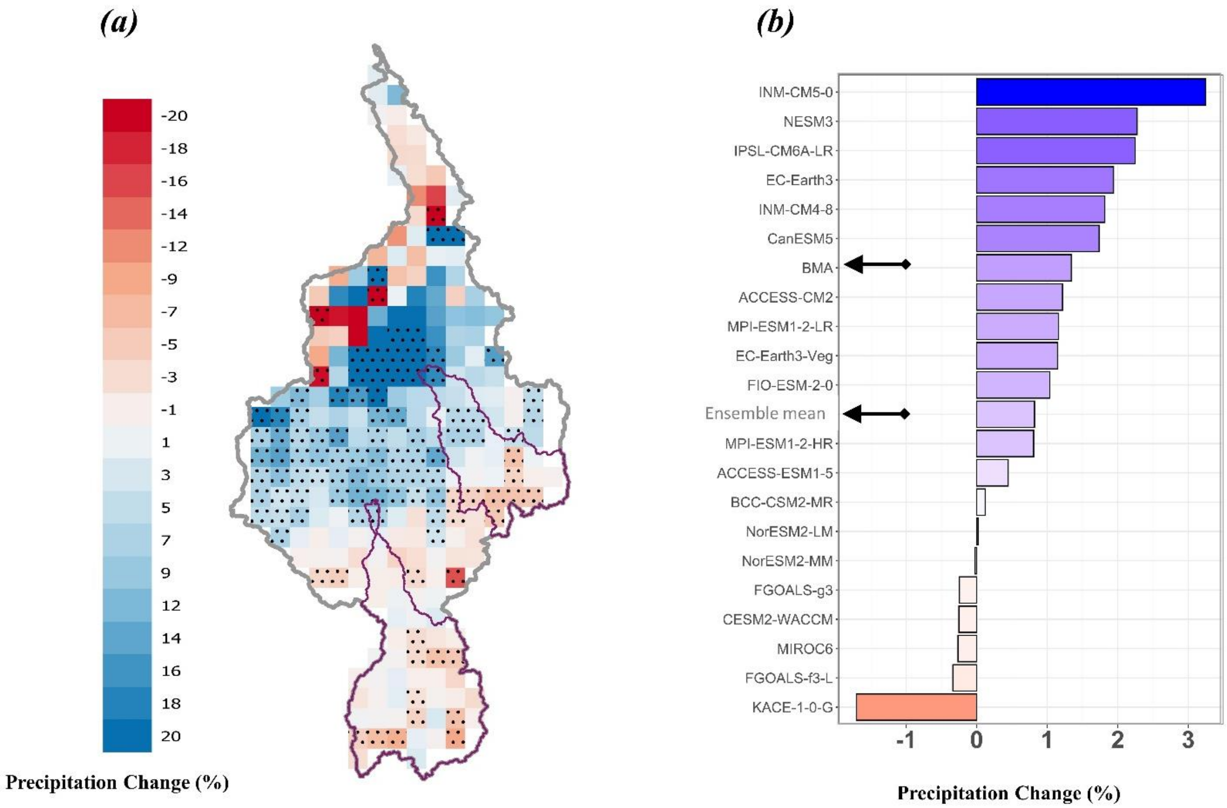

Figure 6. ***(a)*** Percentage change of annual precipitation projected from the BMA model for the period (2015 – 2100) with respect to the baseline period (1983 – 2014) at spatial grids of 1° x 1°. Stapled grids indicate a statistically significant change at α= 0.05. ***(b)*** Percentage change of spatially averaged annual precipitation projected from 20 bias adjusted GCMs, ensemble mean and BMA model. Spatial averaging is carried out over the entire Nile river basin. Black arrows point to ensemble mean and BMA.

It is important to narrow the analysis down to regional scales of unified precipitation regimes. Here we focus on headwaters basins, namely the Blue Nile and upper White Nile sub-basins (see Figure 1). These basins are characterized by distinct precipitation regimes as shown in Figure 5. Figure 7a shows the decadal moving average of percentage change in projected annual precipitation at the Blue Nile sub-basin. Inter-model differences are clearly present with a range of -5% to 5% (dashed thin black lines). BMA and ensemble mean are nearly equivalent, and they show no noticeable change in precipitation. Precisely, BMA shows a change of 0.03%, not statistically significant with p-value of 0.49 (see Table 3). At the upper White Nile sub-basin (Figure 7b), BMA deviates from ensemble mean, and it indicates a decrease of -1.65% in mean annual precipitation, p-value of 0.09 (see Table 3).

In addition to precipitation projections of the BMA and ensemble mean, Table 3 shows the projected change in precipitation in each basin from a selected subset of 3 models. The selection criterion is to identify the 3 models with the least bias in the historical period (1983 – 2014); see Figure 1 and Table 2. In each basin, a subset of 3 models is selected, and its mean is calculated.

Table 3 shows that the estimate of the best 3 models is consistently opposite in sign to the estimate of BMA and ensemble mean. However, all its projected changes are small (< 1%) and statistically insignificant at $\alpha = 0.05$. We also examined precipitation projections for the rainy seasons in the Nile headwaters basins due to their impact on the variability of the Nile streamflow. The results are shown in Table 4, and they don't show a statistically significant trend, whether decreasing or increasing. Of particular importance is the (June – August) rainy season in the Blue Nile basin since it contributes 60% of the annual Nile flow. Table 4 shows that the projected change is statistically insignificant with a decrease of -0.09% (p-value = 0.49).

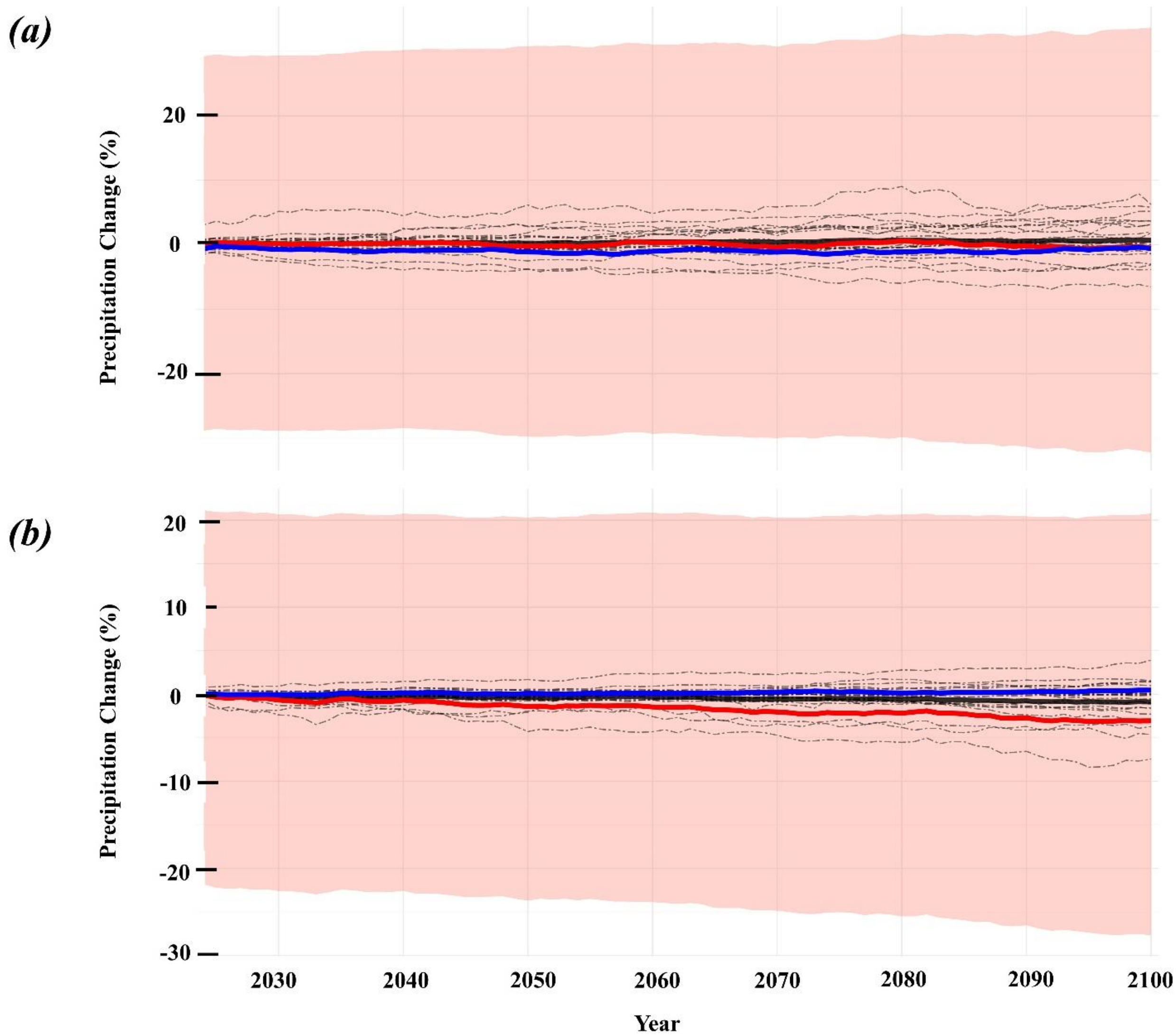

Figure 7. 10-years moving averages of percentage change in projected annual precipitaiton for the period (2015 – 2100) with respect to the baseline period (1983 – 2014). The horizontal axis shows the year at the end of the 10-years time window. Dashed thin black lines, thick balck, red and blue lines indicate projections of the 20 GCMs, ensemble mean, BMA and "best 3 models" respectively. The pink shaded area represents 90% uncertainty bounds of the BMA model. **(a)** Spatially averged over the Blue Nile sub-basin. **(b)** Spatillay averaged over the upper White Nile sub-basin.

Table 4. BMA projected changes in seasonal precipitation in the Nile, Blue Nile and Upper White Nile basins.

Basin	*June - August*		*October - December*		*March - May*	
	Change in mean (%)	*p-value*	*Change in mean (%)*	*p-value*	*Change in mean (%)*	*p-value*
Blue Nile	-0.09	0.49	-0.83	0.45	0.18	0.49
upper White Nile	0.37	0.48	0.47	0.45	-0.09	0.49

4.2.2 Uncertainty in BMA precipitation projections

As discussed earlier, the bias adjusted GCMs ensemble is under-dispersive; thus, it underestimates the uncertainty of precipitation. The BMA approach provides a remedy to this problem because it accounts for two types of variability. Specifically, the BMA total variability is decomposed into two components: between and within variability (Raftery et al., 2005). The former considers the spread of ensemble members whereas the latter accounts for the variability within the individual members. This is clearly shown in Figure 7 which shows the BMA 90% confidence interval (shaded pink area). While the spread of models (black dashed lines) is limited to a range of approximately (-5% - 5%) in the two basins, the BMA 90% confidence interval extends beyond $\pm$20%. This extended uncertainty is the result of the BMA approach consideration of the within variability that is not accounted for in the multi-model ensemble.

Figure 8 shows the distributions of the BMA precipitation projections for the period (2015 – 2100) expressed as a percentage change with respect to the baseline period (1983 – 2014). The distributions show also the means (black dashed lines) and the 90% confidence interval limits (red dashed lines). The mean values are the same as those shown in Table 3. Figure 8a shows the distribution for the Nile basin; the 90% interval range is (-1.9% - 5.5%) with a width of 7.4%. This shows that the probability of an increase in precipitation is higher than that of a decrease. On the contrary, Figure 8c shows that the probability of a decrease in rainfall at the upper White Nile basin is higher with a 90% confidence interval range of (-6.9% - 1.9%) with a width of 8.8%. As for the Blue Nile basin, the uncertainty range is wider; specifically, (-6.8% - 7.2%) with a width of 14%. Besides the wide range of uncertainty in the Blue Nile basin, Figure 8c shows that the distribution is more centered around 0%; thus, there is also increased uncertainty in the sign of change of precipitation projections.

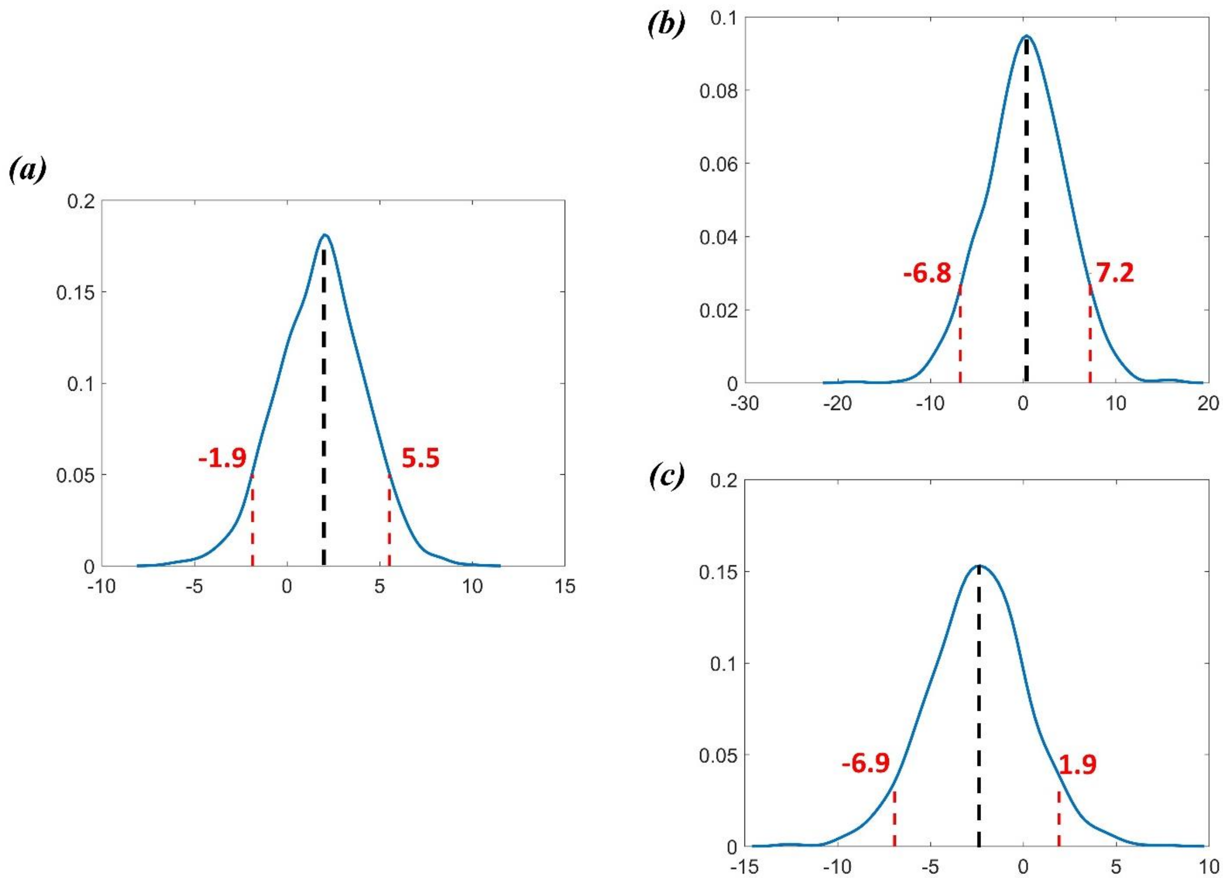

Figure 8. The distirbution of BMA precipitation projections expressed as the mean percentage change with respect to baseline period (1983 – 2014). The distribution mean and 90% confidence bounds are shown in black and red dashed lines respectively. Precipitation is spatially averaged over: **(a)** the Nile basin, **(b)** the Blue Nile basin and **(c)** the upper White Nile basin.

5 Conclusions

This study examined the performance of 20 CMIP6 GCMs in simulating precipitation for the period (1983 – 2014) over the Nile basin, and then used a Bayesian model averaging scheme to derive precipitation projections for the period (2015 – 2100). The main findings of retrospective analysis are as follows:

- The bias in most GCMs simulations is significant (up to 64% of mean annual precipitation) which consequently pinpoints the importance of bias adjustment prior to analysis of precipitation projections. In addition, the spatial patterns of bias vary considerably within individual models both in the sign and value.
- Although all models fairly represent spatial patterns and seasonal cycle of precipitation over most regions in the Nile basin, the results show that the performance of models is less accurate at the upper White Nile basin.

- Selection of a "best performing model" is highly dependent on the specific metric chosen as a criterion. Moreover, the results show that the ensemble mean usually does not outperform all individual models.
- All models severely underestimate the interannual variability as represented by the coefficient of variation. As a result, the ensemble range underestimates the uncertainty of precipitation.

Bayesian model averaging show that projected changes in precipitation varies spatially across the 4Nile basin with clear regional patterns; in particular, a mild decrease of -1.65% in the upper White Nile sub-basin, almost no change (0.03%) in the Blue Nile sub-basin, and significant changes (both increasing and decreasing) in the arid riparian Nile basin. Regarding the Blue Nile sub-basin, our results are similar to those reported by Elshamy et al. (2009) which showed no change in annual precipitation based on 17 CMIP3 GCMs. However, they are at odds with results in Kim and Kaluarachchi (2009), and Fenta Mekonnen and Disse (2018) which showed an increase of 11% and (2.1% - 43.8%) respectively. Generally, it is not possible to make a conclusive judgement on which study, among previous studies and including the present one, has more credibility because they differ significantly in the models, climate scenarios, future time period and geographical regions. Nonetheless, we argue that a strict and more cautious approach compared to previous ones has been adopted in this study.

Lastly, the BMA probability distributions show that the probability of a decrease in annual precipitation is more likely in the upper White Nile basin. Moreover, the uncertainty in annual precipitation projections over the Blue Nile basin is higher both in terms of values and sign of change. Precisely, the 90% confidence interval of BMA has a range of (-6.8% - 7.2%) centered approximately at 0%.

Acknowledgments and Data

This research was partially supported by the Department of Energy (DoE prime award DE-IA0000018), California Energy Commission (CEC award 300-15-005), University of California (#4600010378 TO#15 Am 22).

CMIP6 data used in this study are available online at (https://esgf-node.llnl.gov/search/cmip6/). PERSIANN-CDR dataset are publicly available from the Center for Hydrometeorology and Remote Sensing (CHRS) Data Portal (https://chrsdata.eng.uci.edu/).

Author contributions: M. Ombadi designed and implemented the study. M. Ombadi wrote the manuscript. P.Nguyen, S. Sorooshian and K. Hsu provided insightful feedback on the manuscript.

References

Adler, R. F., Sapiano, M. R., Huffman, G. J., Wang, J. J., Gu, G., Bolvin, D., ... & Xie, P. (2018). The Global Precipitation Climatology Project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation. *Atmosphere*, **9**(4), 138. https://doi.org/10.3390/atmos9040138

Ajami, N. K., Duan, Q., & Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water resources research*, **43**(1). https://doi.org/10.1029/2005WR004745

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., ... & Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, **96**(1), 69-83. https://doi.org/10.1175/BAMS-D-13-00068.1

Bhattacharjee, P. S., & Zaitchik, B. F. (2015). Perspectives on CMIP5 model performance in the Nile River headwaters regions. *International Journal of Climatology*, **35**(14), 4262-4275. https://doi.org/10.1002/joc.4284

Conway, D. (1996). The impacts of climate variability and future climate change in the Nile Basin on water resources in Egypt. *International Journal of Water Resources Development*, **12**(3), 277-296. https://doi.org/10.1080/07900629650178

Conway, D. (2005). From headwater tributaries to international river: Observing and adapting to climate variability and change in the Nile basin. *Global Environmental Change*, **15**(2), 99-114. https://doi.org/10.1016/j.gloenvcha.2005.01.003

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 1-22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Duan, Q., & Phillips, T. J. (2010). Bayesian estimation of local signal and noise in multimodel simulations of climate change. Journal of Geophysical Research: Atmospheres, 115(D18). https://doi.org/10.1029/2009JD013654

Duan, Q., Ajami, N. K., Gao, X., & Sorooshian, S. (2007). Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, **30**(5), 1371-1386. https://doi.org/10.1016/j.advwatres.2006.11.014

Elshamy, M., Seierstad, I. A., & Sorteberg, A. (2009). Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios. *Hydrology and Earth System Sciences*, **13**, 551–565. https://doi.org/10.5194/hess-13-551-2009

Fenta Mekonnen, D., & Disse, M. (2018). Analyzing the future climate change of Upper Blue Nile River basin using statistical downscaling techniques. *Hydrology and Earth System Sciences*, **22**(4), 2391-2408. https://doi.org/10.5194/hess-22-2391-2018

Kim, U., & Kaluarachchi, J. J. (2009). Climate Change Impacts on Water Resources in the Upper Blue Nile River Basin, Ethiopia 1. *JAWRA Journal of the American Water Resources Association*, **45**(6), 1361-1378. https://doi.org/10.1111/j.1752-1688.2009.00369.x

Nguyen, P., Ashouri, H., Ombadi, M., Hayatbini, N., Hsu, K. L., & Sorooshian, S. (2020). PERSIANN-CDR for Hydrology and Hydro-climatic Applications. In *Satellite Precipitation Measurement* (pp. 993-1012). Springer, Cham.

Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., ... & Thorstensen, A. R. (2018). The PERSIANN family of global satellite precipitation data: a review and evaluation of products. *Hydrology and Earth System Sciences*, **22**(11), 5801-5816. https://doi.org/10.5194/hess-22-5801-2018

Nguyen, P., Thorstensen, A., Sorooshian, S., Zhu, Q., Tran, H., Ashouri, H., ... & Gao, X. (2017). Evaluation of CMIP5 model precipitation using PERSIANN-CDR. *Journal of Hydrometeorology*, **18**(9), 2313-2330. https://journals.ametsoc.org/doi/full/10.1175/JHM-D-16-0201.1

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. L. (2018). Developing Intensity-Duration-Frequency (IDF) Curves From Satellite-Based Precipitation: Methodology and Evaluation. *Water Resources Research*, **54**(10), 7752-7766. https://doi.org/10.1029/2018WR022929

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly weather review*, **133**(5), 1155-1174. https://doi.org/10.1175/MWR2906.1

Schaller, N., Mahlstein, I., Cermak, J., & Knutti, R. (2011). Analyzing precipitation projections: A comparison of different approaches to climate model evaluation. *Journal of Geophysical Research: Atmospheres*, **116**(D10). https://doi.org/10.1029/2010JD014963

Swain, A. (2008). Mission not yet accomplished: managing water resources in the Nile River basin. *Journal of International Affairs*, 201-214.

Ter Braak, C. J. (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing*, **16**(3), 239-249. https://doi.org/10.1007/s11222-006-8769-1

The World Bank, World Development Indicators (2018). *Population growth (annual %)*. Retrieved from. https://data.worldbank.org/indicator/SP.POP.GROW

The World Bank, World Development Indicators (2020). *Population, total*. Retrieved from. https://data.worldbank.org/indicator/SP.POP.TOTL

Yates, D. N., & Strzepek, K. M. (1996). Modeling economy-wide climate change impacts on Egypt: A case for an integrated approach. *Environmental Modeling & Assessment*, **1**(3), 119-135. https://doi.org/10.1007/BF01874900

Yates, D. N., & Strzepek, K. M. (1998). Modeling the Nile Basin under climatic change. *Journal of Hydrologic Engineering*, **3**(2), 98-108. https://doi.org/10.1061/(ASCE)1084-0699(1998)3:2(98)