Environ Monit Assess (2023) 195:888
https://doi.org/10.1007/s10661-023-11437-w

RESEARCH

Performance evaluation of multiple regional climate models to simulate rainfall in the Central Rift Valley Lakes Basin of Ethiopia and their selection criteria for the best climate model

Sisay Kebede Balcha · Taye Alemayehu Hulluka · Adane Abebe Awass · Amare Bantider

Received: 8 February 2022 / Accepted: 31 May 2023

Abstract The historical datasets of five regional climate models (RCMs) available in the Coordinated Regional Downscaling Experiment (CORDEX)–Africa database are evaluated against ground-based observed rainfall in the Central Rift Valley Lakes Basin of Ethiopia. The evaluation is aimed at determining how well the RCMs reproduce monthly, seasonal, and annual cycles of rainfall and quantify the uncertainty between the RCMs in downscaling the same global climate model outputs. Root mean square, bias, and correlation coefficient are used to evaluate the ability of the RCM output. The multicriteria decision method of compromise programming was used to choose the best climate models for the climate condition of the Central Rift Valley Lakes subbasin. The Rossby Center Regional Atmospheric Model (RCA4) has downscaled ten global climate models (GCMs) and reproduces the monthly rainfall with a complex spatial distribution of bias and root mean square errors. The monthly bias varies in the range of − 35.8 to 189%. The summer (wet), spring, winter (dry), and annual rainfall varied within the range of 1.44 to 23.66%, − 7.08 to 20.04%, − 7.35 to 57%, and − 3.11 to 16.5%, respectively. To find the source of uncertainty, the same GCMs but downscaled by different RCMs were analyzed. The test results showed that each RCM differently downscaled the same GCM, and there was no single RCM model that consistently simulated the climate conditions over the stations in the study regions. However, the evaluation finds reasonable model skill in representing the temporal cycles of rainfall and suggests the use of RCMs where climate data is scarce after bias correction.

Keywords Seasonality · RCM uncertainty · Multicriteria decision method · Best climate model · CORDEX-Africa · Central Rift Valley Lakes subbasin

S. K. Balcha (✉) · T. A. Hulluka · A. Bantider
Ethiopian Institute of Water Resources, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia
e-mail: kenasisay@gmail.com

S. K. Balcha
College of Agriculture and Environmental Science, Arsi University, P.O. Box 193, Asella, Ethiopia

S. K. Balcha · T. A. Hulluka · A. Bantider
Water and Land Recourse Center, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

A. A. Awass
Institute of Water Technology, Arba Minch University, P.O. Box 21, Arba Minch, Ethiopia

Introduction

Several global climate model (GCM) outputs with different resolutions and emission scenarios are available in the CMIP5 database (http://cmip-pcmdi.llnl.gov/cmip5/terms.html). The GCM outputs are used for climate change impact studies, data filling

Published online: 26 June 2023

for data-scarce areas, and assessments of climate uncertainty for a range of disciplines, including water resources, agriculture, wildfire mitigation, biodiversity conservation, and ecosystem change (Breach et al., 2016; Fordham et al., 2011; Lin & Tung, 2017; Pierce et al., 2009; Shi et al., 2018). Each GCM differently simulates the climate conditions of the same area or region (Kamworapan & Surussavadee, 2019; Pierce et al., 2009; Raju & Kumar, 2017). Numerous questions are raised about the structure of the models and climate system parametrization, initial and boundary conditions, spatial resolution and emission scenarios, and the relevance of GCM outputs (Taylor et al., 2012; Kim et al., 2014; Raju & Kumar 2015; Hidalgo & Alfaro, 2015; Kamworapan & Surussavadee, 2019). Giorgi et al. (2009) highlighted the importance of climate information at regional and local scales. In this regard, GCMs do not provide climate information on a local scale (Giorgi et al., 2009).

It has been more than three decades since the dynamical climate model of regional climate models (RCMs) was developed (Giorgi, 2019) to bridge the gap between GCMs and regional climate responses through downscaling of GCMs (Dibaba et al., 2019; Foley, 2010; Kalognomou et al., 2013; Luhunga et al., 2016; Min et al., 2013). The RCMs have some improvements over the driving GCMs and are used as zooming devices (Dibaba et al., 2019). They are an increasingly important source of climate information and deliver comprehensive climate information on a regional or smaller scale than GCMs (Foley, 2010). However, RCMs are subjected to several uncertainties that are inherited from the boundary conditions of GCMs, the size of domain integration, and the variability between RCM formulations (Giorgi, 2019; Luhunga et al., 2016; Min et al., 2013). For the worst scenario, RCMs unrealistically simulate the frequency, intensity, and spatiotemporal variability of climate variables such as rainfall (Kalognomou et al., 2013).

Numerous studies have been done to simulate the climate conditions in Africa and evaluate the performance of RCMs. Among the researchers who evaluated the performance of CORDEX-RCMs are Luhunga et al. (2016), Okafor et al. (2019), Gyamfi et al. (2021), Ilori and Balogun (2021), and Di Virgilio et al. (2019) for West Africa; Warnatzsch and Reay (2019) for sub-Saharan Africa; and Endris et al. (2013) for East Africa. Most of these studies used gridded or reanalysis data as references to measure the ability of the RCMs to simulate the climate conditions of their regions of interest and concluded that each RCM performed differently over a wide range of the study regions.

Today, reanalysis datasets are widely used as alternative sources of data where the observed data are scarce or not reliable. Particularly in regions of very sparse instrumental coverage, reanalysis products are sometimes used as sources of information. These datasets are produced using the numerical models, various data assimilation techniques, and several observational datasets and multiple variables to reflect the state of the atmosphere at different spatial and temporal scales (Bhattacharya et al., 2020). However, different questions were raised about their performance to simulate the climate conditions of a specific regions (Hoffman et al., 2017) and some researchers warn that reanalysis data should not be equated with "real" observations and measurements (Bosilovich et al., 2013). Furthermore, many researchers advise evaluating their performance first before using them in other applications, including to GCM and RCM performance evaluation (Hassler & Lauer, 2021; Nkiaka et al., 2017). The study reveals that the performance of individual reanalysis products depends on the assimilation of different portions of input observations, model physics, observing techniques, data assimilation schemes, and available observations and resolutions (Bhattacharya et al., 2020). Due to this other reasons, the reanalysis data has poorly simulated the climate conditions in tropical oceans, complex mountain regions, northern Africa, and high latitudes (Bhattacharya et al., 2020). Additionally, there are significant discrepancies between annual and seasonal estimates between reanalyses and observations (Hassler & Lauer, 2021; Sun et al., 2018). For example, Goodarzi et al. (2022) have evaluated the performance of TRMM, CDR, CCS-CDR, and CFSR and concluded that the CDR and CFSR satellite data are not helpful for the Haraz-Gharehsoo Basin, as the corresponding data suffered a significant underestimation of daily rainfall. Even some researchers identified the potential bias of reanalysis data and advised applying a bias correction before this dataset was used for analysis (Garibay et al., 2021). Generally, since multiple

factors can affect the results of reanalysis products, such as topographic and climatic conditions, it has been suggested by many authors to evaluate the accuracy of satellite-based products to obtain more robust results (Bhattacharya et al., 2020; Goodarzi et al., 2022; Hassler & Lauer, 2021; Nkiaka et al., 2017). The reliability of reanalysis datasets is limited by large differences in annual and seasonal estimates in complex mountain areas and high-latitude regions (Hassler & Lauer, 2021; Sun et al., 2018). Therefore, care should be given while using the reanalysis products as observed and evaluating the performance of GCM/RCM outputs.

There are some studies on the performance evaluation of RCMs in Ethiopia (Alemseged & Tom, 2015; Dibaba et al., 2019; Endris et al., 2013; Jury, 2015; Kim et al., 2014; Tesfaye et al., 2019; Worku et al., 2018). All of these studies were mainly focused on the spatial and temporal distribution of precipitation (Alemseged & Tom, 2015; Worku et al., 2018), the impact of large-scale climate circulation patterns or teleconnection on precipitation (Endris et al., 2013), and the spatial variability of annual precipitation (Kim et al., 2014). Apart from these, almost all studies focused on the Upper Nile Basin and used the reanalysis data as observation with all weakness that it has (Alemseged & Tom, 2015; Dibaba et al., 2019; Jury, 2015; Kim et al., 2014; Worku et al., 2018).

It is well known that all RCMs do not perform equally well across regions (Endris et al., 2013). Lack of enough study on RCM evaluation against ground-based observed data, inconsistency of model evaluation techniques, and lack of a clear criterion or guideline to choose the premium RCM are some of their gaps. Therefore, quantifying the performance of RCMs for diverse climate conditions like the Central Rift Valley Lakes subbasin (hereafter referred to as the "CRV Lakes subbasin") is very important for data gap filling, which is a key to any development in the subbasin. Keeping this in mind, the aims of the study are (i) to evaluate the performance of five RCMs, archived by CORDEX-Africa, in reproducing monthly, seasonal, and annual rainfall over the CRV; (ii) to quantify the uncertainty of regional climate models to downscale the same GCM output; and (iii) to select the best RCM representing the climate condition of the CRV Lakes subbasin using a multicriteria decision-making compromise programming method.

Materials and methods

Description of study area

The study was conducted in the Katar and Meki watersheds, which are located in the central part of Ethiopia's Rift Valley Lakes Basin. Geographically, the Katar watershed is located between 38.88 to 39.41° E longitudes and 7.36 to 8.18° N latitudes, with an altitudinal range of 1630 to 4188 m a.m.s.l. Similarly, the Meki watershed is located between 38.22 to 39.00° E longitude and 7.83 to 8.46° N latitude, with an altitudinal range of 1686 to 3614 m a.m.s.l. (Fig. 1). Based on rainfall patterns, Ethiopia's climate is classified into three seasons: summer (from June to September), spring (from October to January), and winter (from February to May) (Ademe et al., 2020; Seleshi & Zanke, 2004). The climate condition in the study area is influenced by altitude, wind direction, and complex topographic characteristics. The mean annual precipitation ranges between 749 and 1276 mm and 712 and 1150 mm in the Katar and Meki watersheds, respectively. The average monthly minimum and maximum temperatures range between 13.4 and 14.2 and 27.5 and 28.7 °C in the low-lying areas and 24 and 27 and 27.5 and 30 °C in the high-lying areas of the Katar and Meki watersheds, respectively.

Input data

The Coordinated Regional Downscaling Experiment (CORDEX) program was launched by the World Climate Research Program (WCRP) to advance regional climate research and application through downscaling to a higher resolution of climate information for fourteen (14) regions of the world (Giorgi & Gutowski, 2016; Giorgi, 2019). There are 10 RCMs in the CORDEX archive for the Africa domain, developed by different institutions (https://esgf-node.llnl.gov/search/esgf-llnl/). Historical monthly rainfall datasets of twenty-two CMIP5 GCMs have been downscaled by five RCMs (CCLM4-8–17, CRCM5, RACMO22T, RCA4, and REMO2009) downloaded from https://esgf-node.llnl.gov/search/esgf-llnl/ for the period 1983–2005. The data is available with a horizontal resolution of 0.44° (50 km), and the projections are in a rotated grid format. The projections of the data were remapped bilinearly into latitude/longitude, and

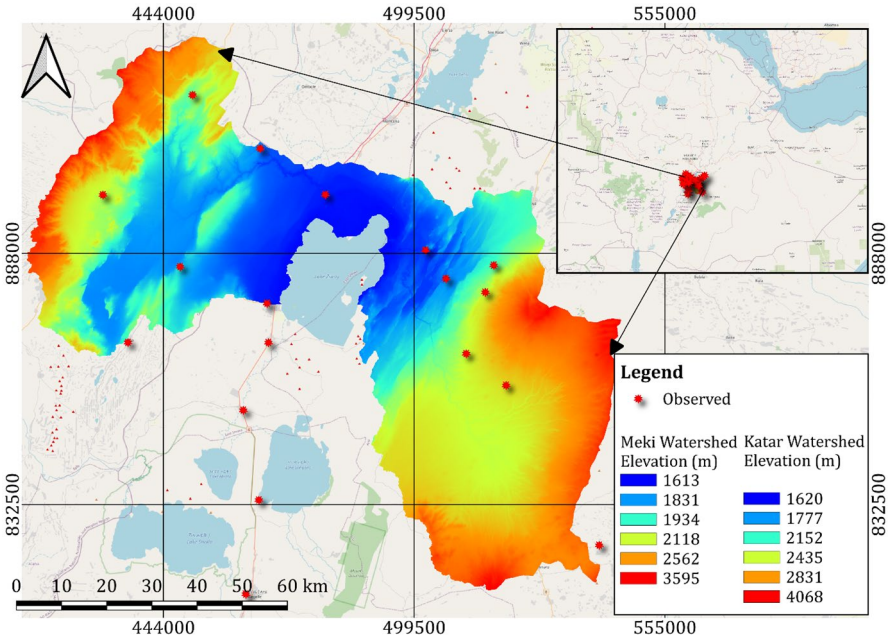

Fig. 1 The study area digital elevation model (DEM), CORDEX grid points, and observed stations in the CRV Lakes subbasin

rainfall flux was converted to total monthly rainfall (mm/month) using the Climate Data Operator (CDO) on Ubuntu operating system.

Monthly observed rainfall data for 22 stations over the study region was collected from Ethiopia's National Meteorological Agency. The data were carefully checked for missing values, and the homogeneity was checked using the standard normal homogeneity test (SNHT) (Alexandersson, 1986), Buishand's range test (BRT) (Buishand, 1982), the Pettitt test (Pettitt, 1979), and the Von Neumann ratio test (VNRT) (Von, Neumann, 1941). Finally, the test results have been summarized following the Wijngaard et al. (2003) procedure, and 15 stations were selected for the performance evaluation of RCMs.

Data analysis methods

Interpolation of gridded climate datasets

To have a fair comparison between observed and simulated datasets, the spatial grids of all RCMs were aggregated into observed stations using the nearest four grids around the observed stations using the inverse distance weighted method (IDWM), which is represented by Eq. 1 (Dibaba et al., 2019; Hartkamp et al., 1999; Luhunga et al., 2016; Ly et al., 2011; Teegavarapu, 2009).

$$P_{mi} = \frac{\sum_{i=1}^{n} \boldsymbol{d_{mi}^{-k}} * \boldsymbol{q_{j(i)}}}{\sum_{i=1}^{n} \boldsymbol{d_{mi}^{-k}}} \quad (1)$$

where P_{mi} is the estimated value at the target station m, n is the total number of grid points, d_{mi} is the distance from grid points i to station m, and $q_{j(i)}$ is the value at grid point i. k is the coefficient—the most commonly suggested value is 2 (Teegavarapu & Chandramouli, 2005; Vieux, 2004). The distance between the grid point and the target station was calculated using the Haversine formula (Eq. 2) for explicit longitude and latitude (Ingole & Nichat, 2015).

$$d_{m(i)} = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\Phi_2 - \Phi_1}{2}\right)\right) + \cos\Phi_1\cos\Phi_2\sin^2\left(\frac{\psi_2 - \psi_1}{2}\right)}\right) \quad (2)$$

where $d_{m(i)}$ is the distance in a kilometer between two points in longitude (Φ) and latitude (ψ) and r is the radius of the earth.

Performance indicators

Several statistical metrics were available to measure the accuracy of the RCM simulation (Raju & Kumar, 2014; Fordham et al., 2011; Raju et al., 2016; Raju & Kumar, 2018). The most commonly used is the correlation coefficient (CC),

normalized root mean square error (NRMSE), and percent of bias (PBIAS) (Moriasi et al., 2007; Raju & Kumar, 2014; Fordham et al., 2011; Raju et al., 2016; Raju & Kumar, 2018). To describe the temporal and spatial similarity between the observation and the simulation, the correlation coefficient was used (Su et al., 2013). It measures the mean difference between the observed and simulated variables. The correlation coefficient can have values ranging from − 1 (negative correlation) to 1 (positive correlation). Sometimes, the CC might include a value of zero indicating that there is no relationship between the two variables and is calculated by Eq. 3 (Raju & Kumar, 2018).

$$\mathrm{CC} = \frac{\sum_{i=1}^{n} \left(X_{iobs} - X_{mean}\right)\left(Y_{i\mathrm{sim}} - Y_{\mathrm{mean}}\right)}{(n-1)\sigma_{\mathrm{obs}}\sigma_{\mathrm{sim}}} \tag{3}$$

where $X_{i\mathrm{obs}}$ and $Y_{i\mathrm{sim}}$ are the monthly observed and simulated values i months, X_{mean} and Y_{mean} are the average values of observed and simulated values, and σ_{obs} and σ_{sim} are the standard deviations of observed and simulated values.

Normalized root mean square error (NRMSE) was used to describe the ratio of root mean square error between observed and simulated variables to the corresponding standard deviation of the observed (Su et al., 2013). The smaller the value near zero is, the better the model performance and is calculated using Eq. 4 (Fu et al., 2013).

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{i\mathrm{obs}} - Y_{i\mathrm{sim}})^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_{i\mathrm{obs}} - X_m)^2}} \tag{4}$$

The percent of bias (PBIAS) is used to measure the average propensity of the simulated data to be larger or smaller than its observed counterparts and is calculated by Eq. 5 (Barrios et al., 2018; Moriasi et al., 2007).

$$\mathrm{PBIAS} = 100 * \left(\sum_{i=1}^{n}(Y_{i\mathrm{obs}} - Y_{i\mathrm{sim}}) / \sum_{i=1}^{n}(Y_{i\mathrm{obs}})\right) \tag{5}$$

The optimum value of PBIAS is zero for perfect model simulation and a positive or negative value that shows model underestimation or overestimation of rainfall variables (Moriasi et al., 2007).

Multicriteria decision method (MCDM) of compromise programing

In the fields of engineering and social science, several decision-making methods are available to solve different types of problems (Majmder, 2015; Mardani et al., 2015; Raju & Kumar, 2018; Sabaei et al., 2015). Among the several methods, compromise programming (CP) takes on the challenge of determining decision-makers' preferences realistically without relying on the standard utility theory's faulty assumptions. The primary goal of CP is to find an ideal solution, and every decision-maker pursues a solution that is as close to the ideal point as possible. A distance function is used in the analysis to obtain this proximity. The distance ($L_{p(a)}$) between the ideal values is calculated using Eq. 6 (Raju et al., 2016; Beula & Prasad, 2012).

$$L_{p(a)} = \left[\sum_{j=1}^{J} w_j^p \left(f_j^* - f_{j(a)}\right)^p\right]^{\frac{1}{p}} \tag{6}$$

where indicators $j = 1,2\ldots J$; $L_{p(a)} = L_p$ metric for imputation method for the chosen value of parameter p; $f_{j(a)}$ = normalized value of indicator j; w_j = weight of performance indicator j; and p is a parameter (1 for linear, 2 for Euclidean distance measure) and for this study, we adopted the p-value of 2. An entropy method was used to normalize the measured values and calculate the weight of various criteria from the given payoff matrix (P_{ij}). It helps to make a decision independently of the views of the decision-maker and reduces subjectivity (Pomerol & Romero, 2000; Raju & Kumar, 2014; Zhu et al., 2020). The first step of the entropy method is standardizing the payoff matrix P_{ij} using Eq. 7 (Beula & Prasad, 2012; Zhu et al., 2020).

$$P_{ij} = \frac{P_{ij}}{\sum P_{ij}} \tag{7}$$

For a given normalized payoff matrix p_{ij}, an entropy E_i is calculated using Eq. 8 for a set of alternative criteria j (Beula & Prasad, 2012; Zhu et al., 2020).

$$E_j = \frac{1}{\ln(N)}\sum_{i=1}^{N} p_{ij}\ln\left(p_{ij}\right) \mathrm{for} j = 1, \ldots J \tag{8}$$

where N is the number of imputation techniques and j is the number of indicators and degree of diversification (D_j) of the information provided by the outcomes of criterion j and is calculated using Eq. 9 (Beula & Prasad, 2012; Zhu et al., 2020).

$$D_j = 1 - E_j \text{ for } j = 1, ..., J \tag{9}$$

Then, finally, the normalized weights of indicators are calculated using Eq. 10 (Beula & Prasad, 2012; Zhu et al., 2020) and all the calculations are done on Microsoft Excel 2016.

$$w_j = \frac{D_j}{\sum_{i=1}^{J} D_j} \tag{10}$$

Results

Spatiotemporal distribution of biases root mean square error and correlation coefficient

The evaluations have been conducted following guidelines for the performance assessment of the public weather service of the World Meteorological Organization (WMO) as reported by Gordon and Shaykewich (2000). An overview of the results is presented for each climate model as follows.

The monthly cycle of rainfall

The monthly cycle of average rainfall is presented in Fig. 2 for stations in the study area. The figures display the temporal variations of rainfall between the GCMs and RCMs. All GCMs preserved the temporal cycle of the observed rainfall with different magnitudes of over or underestimation.

The biases are calculated using Eq. 5 to assess the performance of RCMs to simulate long-term (1983–2005) average monthly rainfall at different stations in the CRV Lakes subbasin (Fig. 3a–e). Each RCM differently simulates the monthly observed rainfall amount over each station. From this figure, CCLM4-8–17 was derived by CMIP5 GCMs (HadGEM2-ES, MPI-ESM-LR and CNRM-CM5). HadGEM2-ES and MPI-ESM-LR are underestimating monthly rainfall of all stations in the CRV Lakes subbasin; with a magnitude bias in the range of − 60.1 to − 25.41 and − 52.8 to − 0.35%, respectively. However, CNRM-CM5 overestimated the monthly rainfall of Adami-Tulu, Arata, Bui, Ejerse-Lele, Katar-genet,

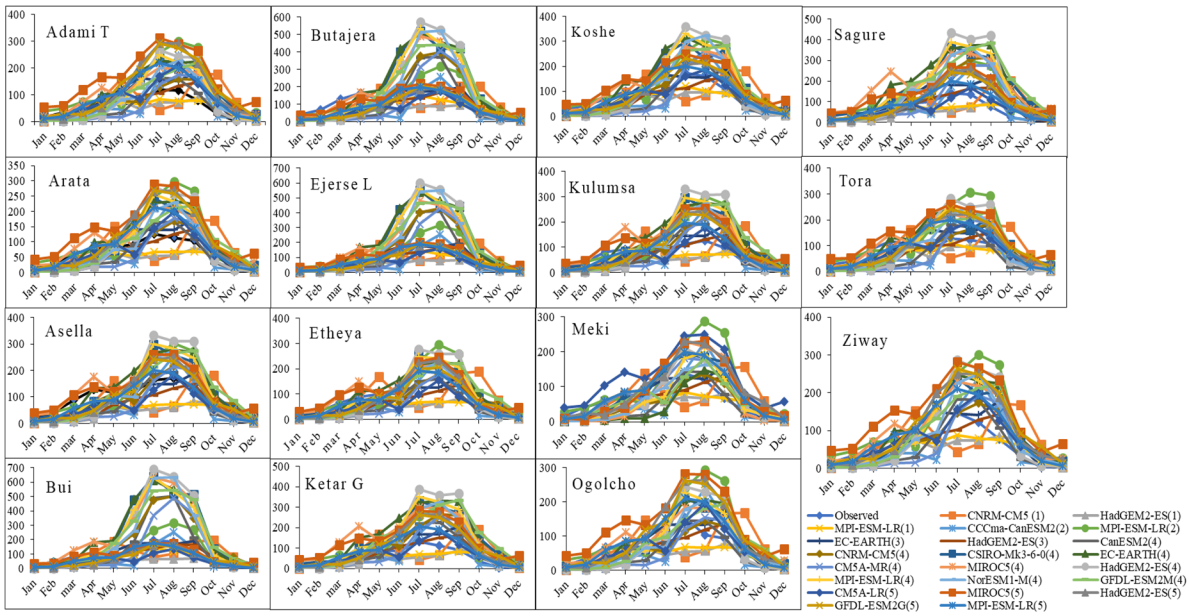

Fig. 2 Average monthly observed and simulated rainfall in millimeter. The number in the bracket stands for (1) for CCLM4-8–17, (2) for CRCM5, (3) for RACMO22T, (4) for RCA4, and (5) for REMO2009 regional climate models (RCMs)

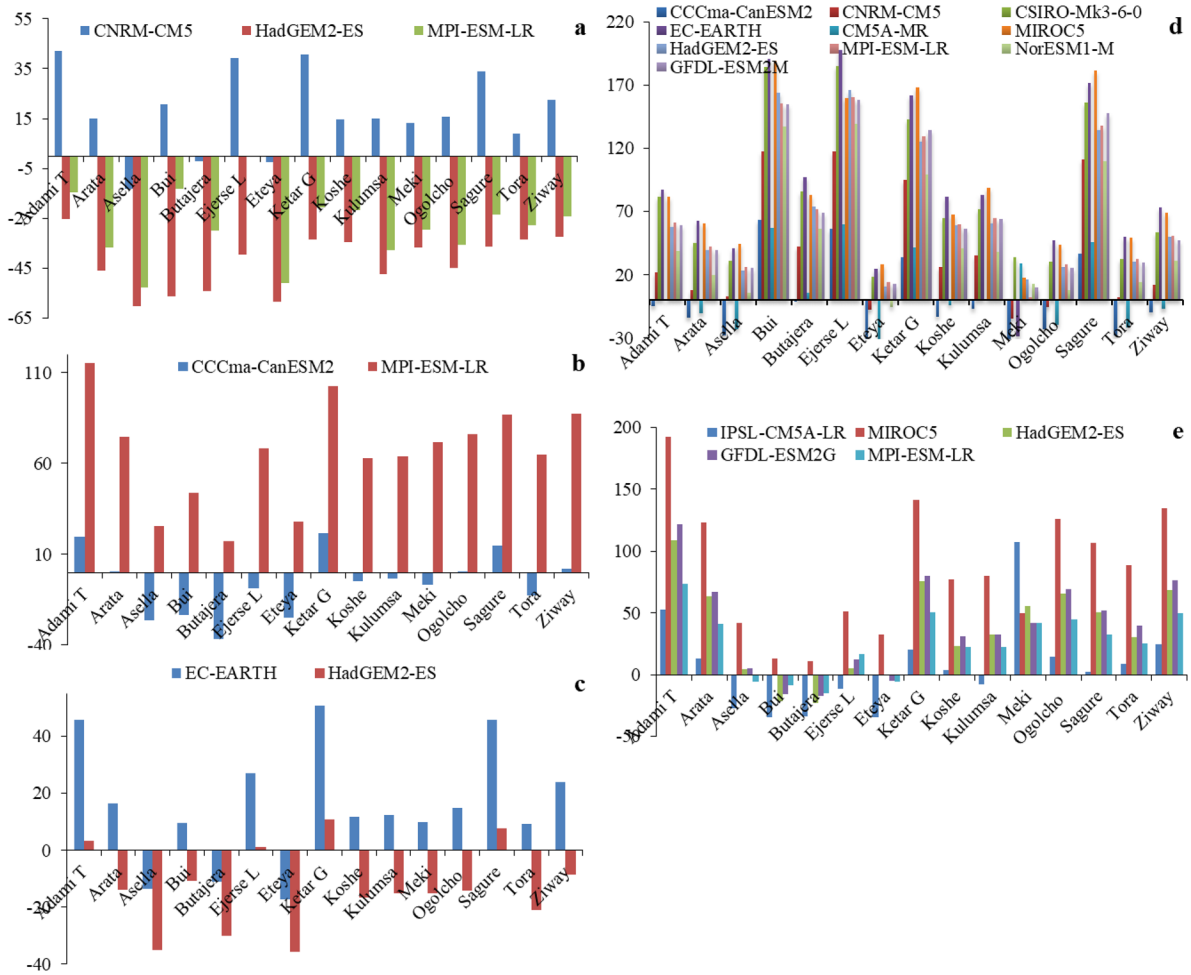

Fig. 3 Spatial distribution of monthly bias (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

Koshe, Kulumsa, Meki, Ogolcho, Sagure, and Ziway with biases in the range of 13.3 to 42% and underestimated the monthly rainfall of Asella, Butajera, and Etheya with biases in the range of − 2.12 to − 13.14% (Fig. 2a).

CRCM5 was derived from two global climate models, MPI-ESM-LR and CanESM2. The MPI-ESM-LR model overestimated the monthly rainfall of all stations with biases in the range of 17.4 to 115.48%. The CanESM2 model overestimated the monthly rainfall of Adami-Tulu, Katar-genet, Ogolcho, Sagure, and Ziway stations by a magnitude in the range of 0.08 to 19.95%. However, CCCma-CanESM2 underestimated precipitation in Asella, Bui, Butajera, Ejerse-Lele, Koshe, Kulumsa, and Meki, within the biases in the range of − 3.51 to − 36.37%. The performance evaluation of CRCM5 based on bias is presented in Fig. 3b. From this figure, the CCCma-CanESM2 model exhibited the lowest bias among the MPI-M-MPI-ESM-LR over all stations.

RACMO22T was derived by two GCMs, EC-EARTH and HadGEM2-ES. EC-EARTH underestimated the monthly rainfall of Asella, Eteya, and Butajera stations with biases in the range of 11 to 17.3% and overestimated the monthly rainfall over most stations with biases in the range of 9 to 45%. The MOHC-HadGEM2-ES model overestimated the monthly rainfall of Adami-Tulu, Ejerse-Lele, Katar-genet, and Sagure by up to 10% and underestimated the monthly rainfall of most stations with a bias in the

range of 8 to 35%. In comparison to these two models, MOHC-HadGEM2-ES has lower biases than EC-EARTH (Fig. 3c).

RCA4 was derived by ten (10) CMIP5 GCMs, which is the maximum number of GCMs in the CORDEX archive. CSIRO-Mk3-6–0, EC-EARTH, MIROC5, HADGEM2-ES, MPI-ESM-LR, and GFDL-ESM2M overestimated monthly rainfall over all stations with biases in the range of 1.96 to 197.62%. CM5A-MR and CanESM2 underestimated monthly rainfall over most stations with biases in the range of 0.62 to 37% and overestimated rainfall over a few stations with biases in the range of 0.42 to 63%. On the other hand, CNRM-CM5 and NorESM1-M overestimated the monthly rainfall over most stations with biases in the range of 2.5 to 139.16% and underestimated the monthly rainfall of very few stations with biases in the range of 5.38 to 14.69%. The percentage of bias errors is very high for most climate models at stations named Bui, Ejerse-Lele, Katargenet, and Sagure. As presented in Fig. 3d, Etheya, Meki, Tora, Asella, Arata, and Adami-Tulu stations performed well with the lower value of bias percentages. Among the ten RCA4 climate model outputs, CanESM2 better simulates the monthly rainfall over all stations.

REMO2009 was derived by five GCMs. MIROC-MIROC5 is the only GCM that overestimates monthly rainfall over all stations with biases in the range of 11.12 to 192.28%. On the other hand, HadGEM2-ES and GFDL-ESM2G underestimated the monthly rainfall amounts of Bui, Butajera, and Eteya and overestimated most stations (Fig. 3e). CM5A-LR and MPI-ESM-LR are climate models that over- and undersimulate rainfall amounts with biases of magnitude ranging between 3.3 and 107.5% and − 5.41 and − 34.8%, respectively.

A lower value of bias alone does not exhibit the models' performance. NRMSE has been used to identify the best model based on the assumption that an excellent model must have a smaller bias and NRMSE. Figure 4a–e presents the results of the performance evaluation of each RCM based on NRMSE. The accuracy of the CCLM4-8–17 to downscale historical monthly rainfall over most stations varied, with errors in the range of 41.38 to 123.22% (Fig. 4a). CNRM-CM5 and HadGEM2-ES simulated monthly rainfall over all stations with the NRMSE in the range of 70 to 123% and 51 to 80%, respectively. Compared to these three models, the MPI-ESM-LR better performed over most stations.

CRCM5 was derived by the CCCma-CanESM2 and MPI-M-MPI-ESM-LR GCMs and the result of their performance based on NRMSE are presented in Fig. 4b. CCCma-CanESM2 and MPI-M-MPI-ESM-LR simulated monthly rainfall over all stations with NRMSE in the range of 48.64 to 91.75% and 61.45 to 169.19%, respectively. Compared to these two models, CCCma-CanESM2 is better performed in simulating monthly rainfall. RACMO22T was derived by ICHEC-EC-EARTH and MOHC-HadGEM2-ES models. Figure 4c presents the NRMSE results of both models to simulate monthly rainfall over all stations with an NRMSE in the range of 37.33 to 89.98% and 45.69–59.91% respectively. However, MOHC-HadGEM2-ES is relatively better performed over most stations.

Figure 4d presents the performances of RCA4 based on NRMSE to downscale ten GCMs. All climate models simulated the monthly rainfall over each station with an error in the range of 36.25 to 309.96%. The largest error observed at Bui station with NRME was in the range of 171.95 to 309.96%. Among these climate models, there is no single model that consistently simulates the monthly rainfall of the stations. However, the NRMSE on Bui, Butajera, and Ejerse-Lele is higher compared to that on other stations. The REMO2009 was derived by five GCMs (CM5A-LR, MIROC5, HadGEM2-ES, GFDL-ESM2G, and MPI-ESM-LR) and simulated the monthly rainfall over the stations in the study region with an error in the range of 15.89 to 219.59% (Fig. 4e).

The third statistical measure of performance is the correlation coefficient (CC). It measures the relationship between the simulated and observed monthly rainfall. The spatial distribution of the CC between simulation and observed data is shown in Table 1. Except for CCLM4-8–17 driven by CNRM-CM5, all RCMs driven by each GCM have positively correlated with the observed rainfall with a value exceeding 0.5 over each station. Among those RCMs and their driving GCMs, REMO2009 driven by MPI-ESM-LR is the only climate model that better simulates monthly rainfall over all stations compared to other RCMs in respect of the driving GCMs. This might be resulting in the numerical model structure and formulation of RCM. In Di Luca et al.'s (2015) study, there are two

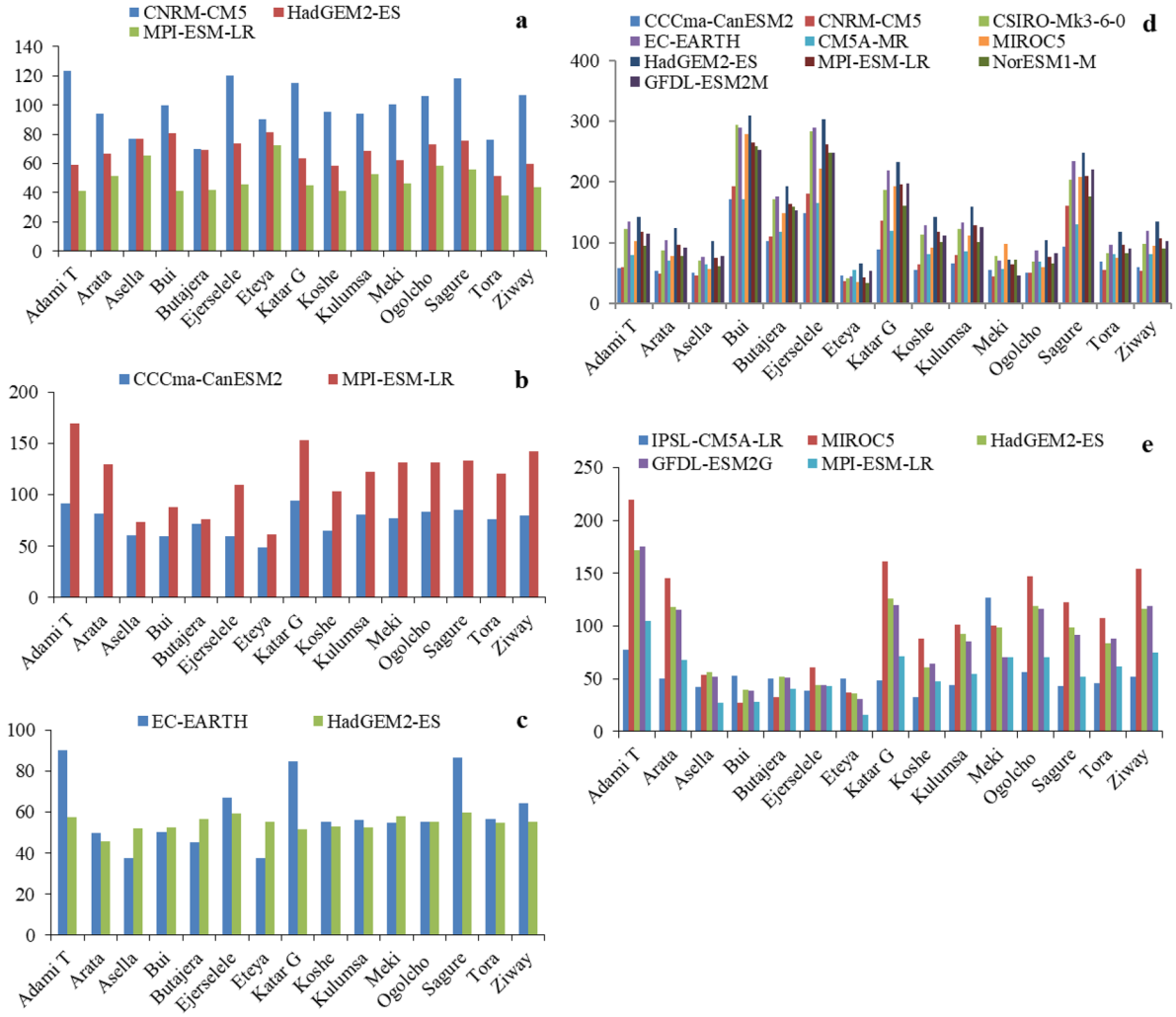

Fig. 4 Spatial distribution of monthly normalized root mean square error (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RAC-MO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

categories of factors influencing the RCMs: those that are related to numerical experimental design and those that are related to the particular climate statistics to be evaluated. Another issue is related to the scale difference between observed data (station data) and gridded data (RCM), the lack of enough spatial and temporal observed data, and the presence of orographic features (Rummukainen et al., 2015). More specifically, in the study region, there are strong effects of topography and direction of the wind on the rainfall amount and distribution over the CRV Lakes subbasin that the RCMs are never able to capture.

The seasonal cycle of rainfall

Summer, winter, and spring are the seasons considered in this evaluation (Ademe et al., 2020; Seleshi & Zanke, 2004). Figures 5 and 6 have presented the seasonal cycle of rainfall over the Katar and Meki watersheds respectively. GCMs downscaled by the CCLM4-8–17 regional climate model underestimated the summer season rainfall over all stations in Katar and Meki watersheds except for Adami-T. However, GCMs downscaled by CRCM5, RAC-MO22T, RCA4, and REMO2009 climate models overestimated the summer season rainfall over all

Table 1 Pearson correlation of monthly rainfall between observed and simulated

RCMs	GCMs	Adami-T	Arata	Asella	Bui	Butajera	Ejerse-L	Iteya	Ketar G	Koshe	Kulumsa	Meki	Ogolcho	Sagure	Tora	Ziway
CCLM4-8–17	CNRM-CM5	0.24	0.37	0.35	0.25	0.39	0.23	0.30	0.38	0.27	0.31	0.19	0.22	0.21	0.45	0.25
	HadGEM2-ES	0.66	0.68	0.63	0.66	0.65	0.69	0.69	0.69	0.74	0.57	0.66	0.60	0.51	0.76	0.71
	MPI-ESM-LR	0.83	0.85	0.85	0.85	0.85	0.84	0.81	0.85	0.87	0.78	0.84	0.77	0.73	0.89	0.87
CRCM5	CanESM2	0.83	0.77	0.79	0.82	0.63	0.86	0.89	0.79	0.84	0.75	0.77	0.74	0.77	0.67	0.78
	MPI-ESM-LR	0.88	0.85	0.82	0.87	0.72	0.88	0.92	0.84	0.87	0.80	0.80	0.83	0.80	0.73	0.83
RACMO22T	EC-EARTH	0.84	0.91	0.88	0.84	0.81	0.81	0.90	0.92	0.82	0.84	0.80	0.84	0.84	0.82	0.84
	HadGEM2-ES	0.80	0.83	0.80	0.77	0.67	0.77	0.84	0.88	0.76	0.74	0.69	0.74	0.79	0.71	0.76
RCA4	CanESM2	0.89	0.86	0.85	0.91	0.68	0.93	0.94	0.89	0.89	0.80	0.86	0.86	0.85	0.69	0.88
	CNRM-CM5	0.82	0.83	0.82	0.91	0.72	0.91	0.89	0.85	0.85	0.76	0.83	0.77	0.76	0.74	0.83
	CSIRO-Mk3-6–0	0.85	0.87	0.84	0.85	0.68	0.85	0.94	0.89	0.81	0.77	0.86	0.87	0.80	0.69	0.85
	EC-EARTH	0.89	0.91	0.91	0.89	0.76	0.89	0.95	0.93	0.89	0.85	0.68	0.86	0.85	0.82	0.89
	CM5A-MR	0.75	0.72	0.70	0.79	0.59	0.82	0.82	0.73	0.79	0.64	0.92	0.69	0.63	0.61	0.72
	MIROC5	0.95	0.96	0.96	0.94	0.81	0.95	0.97	0.94	0.95	0.93	0.83	0.93	0.90	0.89	0.93
	HadGEM2-ES	0.86	0.85	0.82	0.89	0.70	0.89	0.93	0.88	0.86	0.77	0.87	0.85	0.80	0.72	0.86
	MPI-ESM-LR	0.90	0.89	0.88	0.90	0.73	0.91	0.96	0.92	0.89	0.83	0.85	0.90	0.86	0.76	0.89
	NorESM1-M	0.88	0.83	0.83	0.92	0.71	0.93	0.93	0.88	0.90	0.78	0.72	0.81	0.82	0.75	0.87
	GFDL-SM2M	0.74	0.76	0.74	0.85	0.66	0.85	0.83	0.79	0.80	0.67	0.85	0.69	0.67	0.68	0.76
REMO2009	CM5A-LR	0.89	0.85	0.85	0.89	0.76	0.91	0.90	0.87	0.90	0.81	0.93	0.80	0.83	0.80	0.88
	MIROC5	0.95	0.96	0.95	0.95	0.88	0.92	0.98	0.96	0.93	0.92	0.85	0.94	0.93	0.87	0.95
	HadGEM2-ES	0.88	0.88	0.85	0.90	0.72	0.88	0.95	0.90	0.85	0.81	0.84	0.87	0.85	0.72	0.87
	GFDL-ESM2G	0.86	0.87	0.84	0.88	0.70	0.88	0.93	0.90	0.85	0.78	0.93	0.85	0.83	0.71	0.87
	MPI-ESM-LR	0.94	0.96	0.94	0.94	0.84	0.91	0.98	0.96	0.91	0.91	0.93	0.94	0.94	0.84	0.95

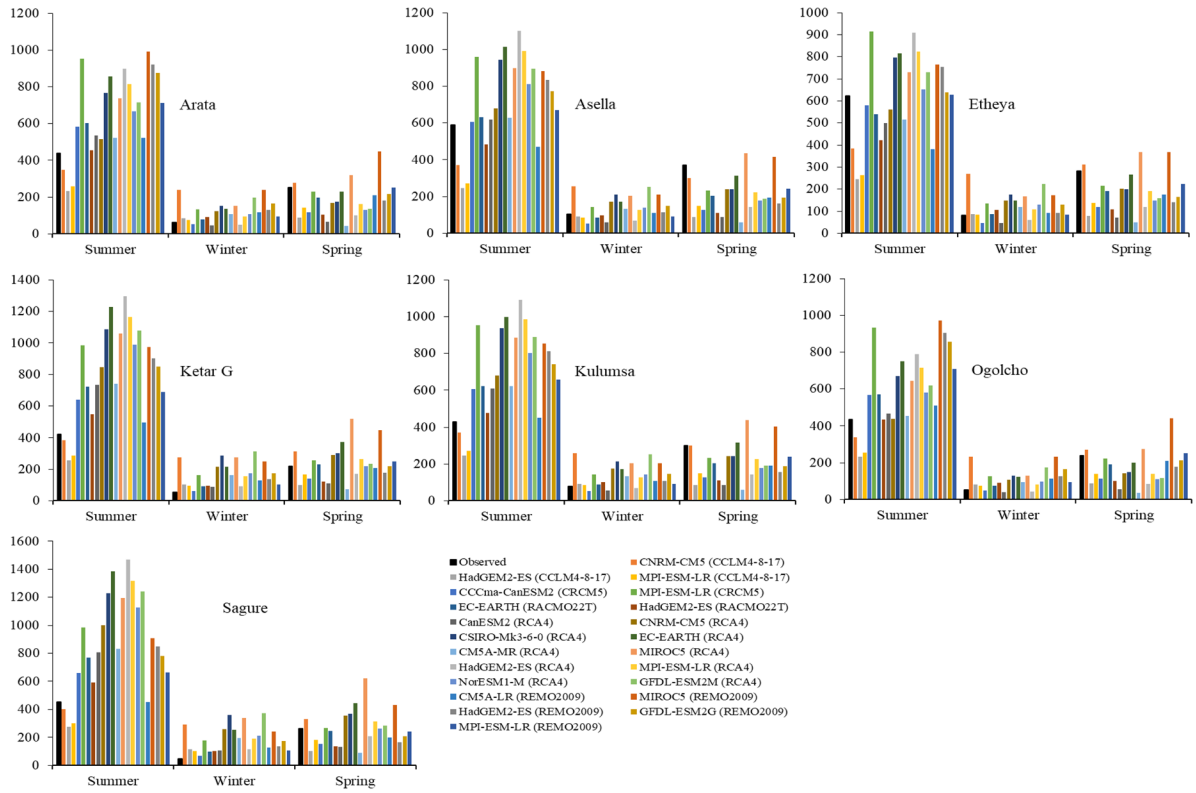

Fig. 5 Seasonal observed and simulated rainfall (mm) of stations in the Katar watershed. Text in the bracket stands for downscaling regional climate model (RCMs) such as CCLM4-8–17, CRCM5, RACMO22T, RCA4, and REMO2009

stations considered for this study. All climate models (RCM/GCMs) overestimated the winter season rainfall at Ketar G, Sagure, and Ejerse-L stations and underestimate only at Adami-T station (Fig. 6). However, over the rest of the stations, the RCMs/GCMs show over- and underestimation characteristics in both watersheds. Spring is the second rainy season in the study region and most climate models underestimated the spring season rainfall compared with the observed rainfall at each station. Generally, the global climate models (GCMs) which were downscaled by different RCMs overestimate the summer rainfall except for CCLM4-8–17 and underestimate winter and spring season rainfall. This shows that the global/regional climate model will not be able to capture the seasonal cycle of rainfall where topography is the major determinant factor for rainfall condition for a region like CRV. This shows that before applying any climate model for impact assessment, it needs to quantify the degree of discrepancy and bias correction is mandatory.

Based on the temporal cycle rainfall over the CRV Lakes subbasin, the ability of each RCM was evaluated for seasonal variation. Figure 7a–e presents the performance evaluation of all RCMs based on their biases. From this figure, CCLM4-8–17 underestimated the seasonal rainfall over most stations with an error in the range of 0.6 to 5.48% (Fig. 7a). However, CNRM-CM5 and MPI-ESM-LR overestimated the summer rainfall at Adami-Tulu by a magnitude of 2.24 and 0.4% respectively. CRCM5 overestimated the summer rainfall over most stations, with a bias in the range of 1.24 to 19.55%. However, CanESM2 underestimated the summer rainfall at Asella, Bui Butajera, and Etheya with a bias in the range of 0.17 to 0.9% (Fig. 7b).

RACMO22T over- and underestimated the summer season rainfall over the stations with bias error in the range of 0.22 to 3.55% and 4.76 to 9.86% respectively. As presented in Fig. 7c, compared to these two GCMs of RACMO22T, HadGEM2-ES better simulates the summer rainfall over most stations. The

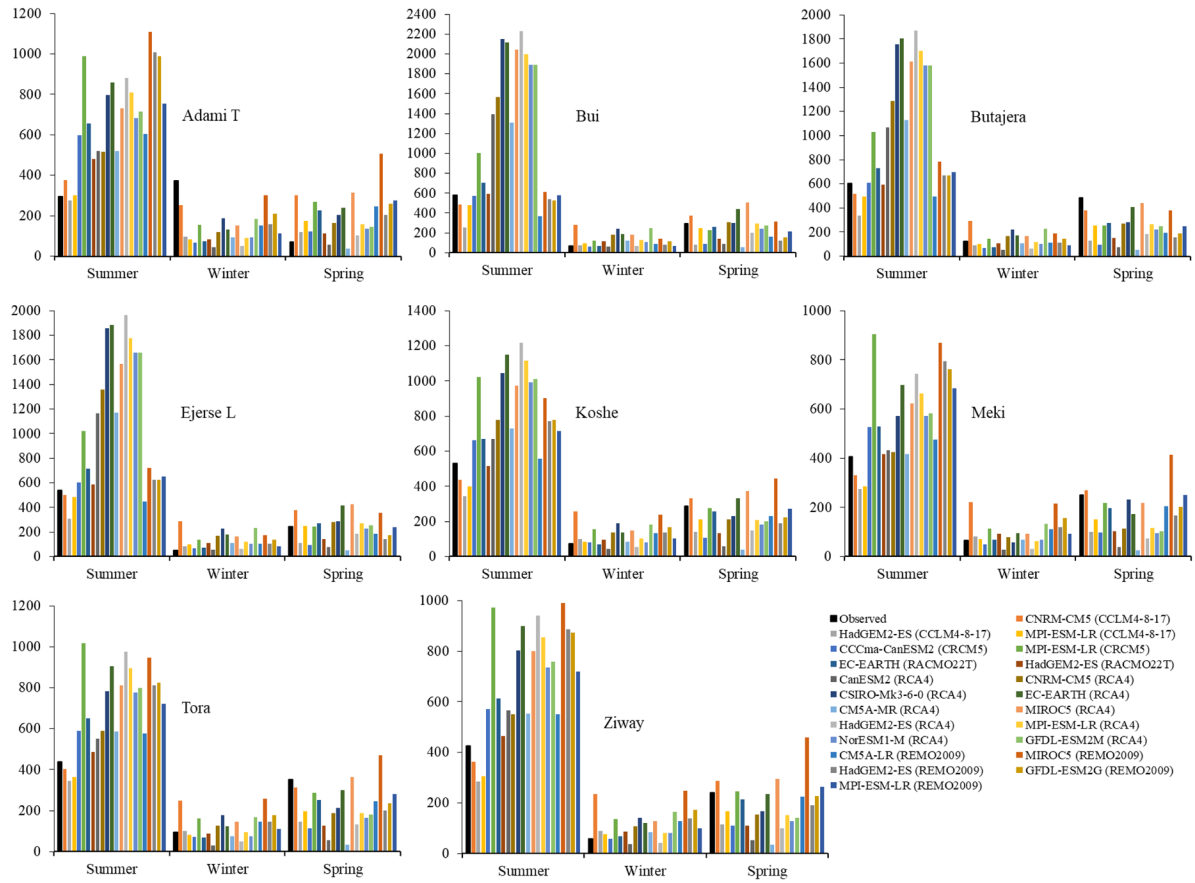

Fig. 6 Seasonal observed and simulated rainfall (mm) of stations in the Meki watershed. Text in the bracket stands for downscaling regional climate model (RCMs) such as CCLM4-8–17, CRCM5, RACMO22T, RCA4, and REMO2009

bias error of RCA4 is presented in Fig. 7d, and each GCM overestimated the summer rainfall over most stations with an error in the range of 0.38 to 21.17%. Among these ten GCMs, CM5A-MR presented the lowest bias ranging from 2.21 to 7.67%. The rest of the GCMs have poorly simulated the summer rainfall over most stations, with a maximum value bias of 23.66%. The REMO2009 is presented in Fig. 7e and the downscaled GCM results overestimated the summer season rainfall over most stations with an error in the range of 3.34 to 24.62%. However, REMO2009 underestimated the summer season rainfall over Asella, Bui, Butajera, Koshe, and Sagure stations (Fig. 7e).

The NRMSE of the summer season for CCLM4-8–17 driven by CNRM-CM5, HadGEM2-ES, and MPI-ESM-LR varied between 8.22 and 22.23% over all stations (Fig. 8a). The CRCM5 model, which is driven by CanESM2 and MPI-ESM-LR, has simulated the summer rainfall of each station with an error in the range of 5.21 to 68.63%. However, the CanESM2 model has better performance than the MPI-ESM-LR at simulating summer rainfall over most stations (Fig. 8b).

The NRMSE of the RACMO22T model, which is driven by EC-EARTH and HadGEM2-ES, varied in the range of 5.61 to 35.38% and 8.1 to 21.09% respectively over all stations (Fig. 8c). Figure 8d shows the NRMSE for RCA4, which has been driven by ten different GCMs and has varied within the range of 7.56 to 74.12%. The accuracy of the RCA4 model to simulate summer rainfall over all stations in the Gurage highland (Bui, Butajera, and Ejerse-Lele) and the lowland of Meki (e.g., Adami-Tulu) and the midland of Katar watershed (Sagure) is very poor. Similarly, to the above models, the ability of REMO2009

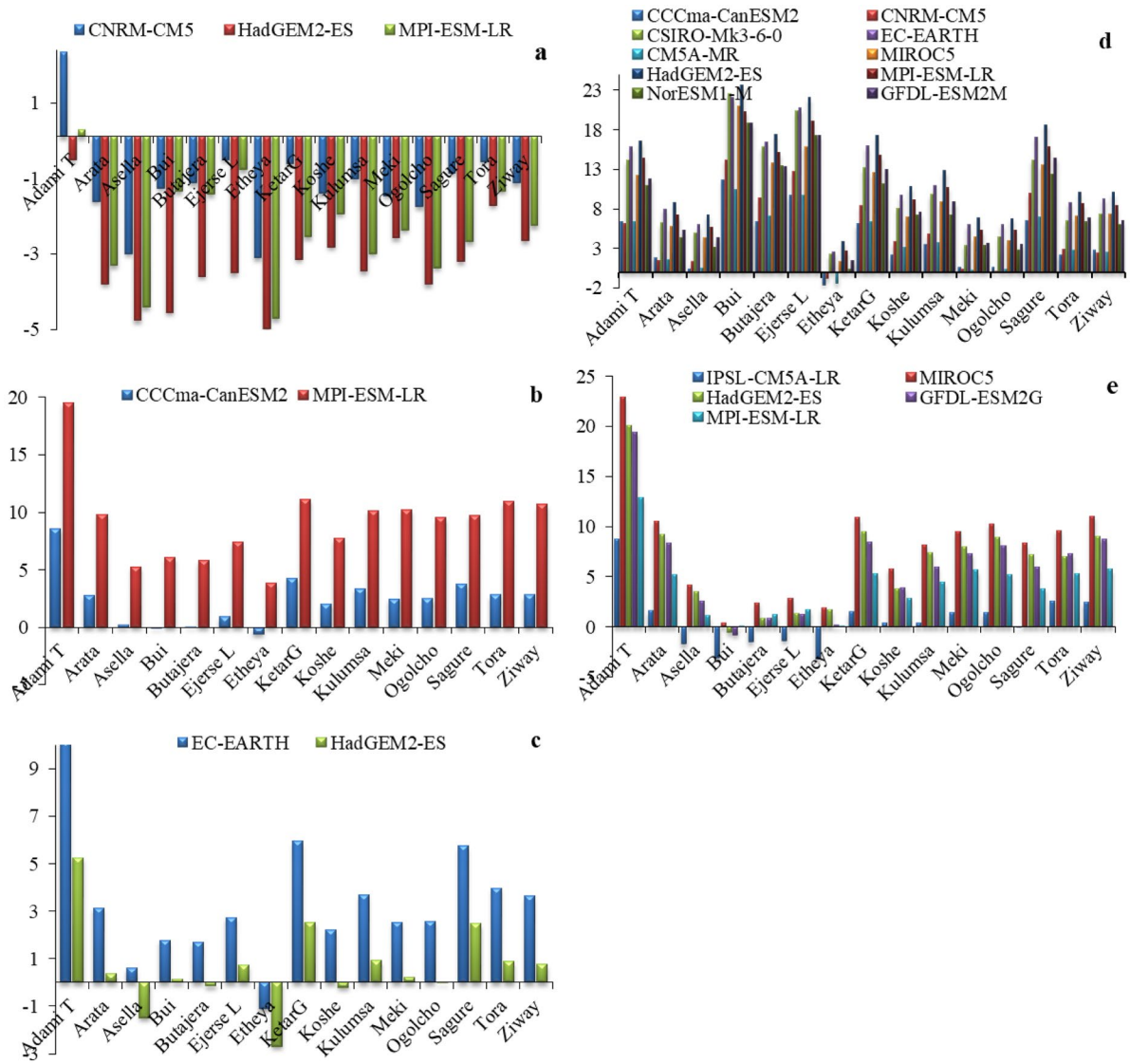

Fig. 7 Summer biases (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

to simulate summer rainfall is presented in Fig. 8e. The NRMSE of all models varies in the range of 7.71to 86.58%. The worst conditions have been seen at Adami-Tulu station, with an error in the range of 35.92 to 86.58%.

The ability of RCMs and their respective GCMs to be assessed for spring rainfall is evaluated by comparing the observed dataset of selected stations. As Fig. 9a–e shows, the normalized bias for CCLM4-8–17 driven by GCMs (CNRM-CM5, HadGEM2-ES, and MPI-ESM-LR) has varied within the range of − 6.37 to 27.32% (Fig. 9a). Except for Adami-Tulu, all models have performed very well with a bias in the range of less than 10% over most stations. CRCM5 driven by CanESM2 and MPI-ESM-LR has simulated the spring rainfall with a bias error in the range of − 6.71 to 6.07% and − 3.98 to 23.16% respectively (Fig. 9b). The bias error of RACMO22T to downscale EC-EARTH and HadGEM2-ES models' outputs for spring season rainfall over most stations is varied within the range of − 3.74 to 18.36 and − 5.84 to 5% respectively (Fig. 9c). The HadGEM2-ES

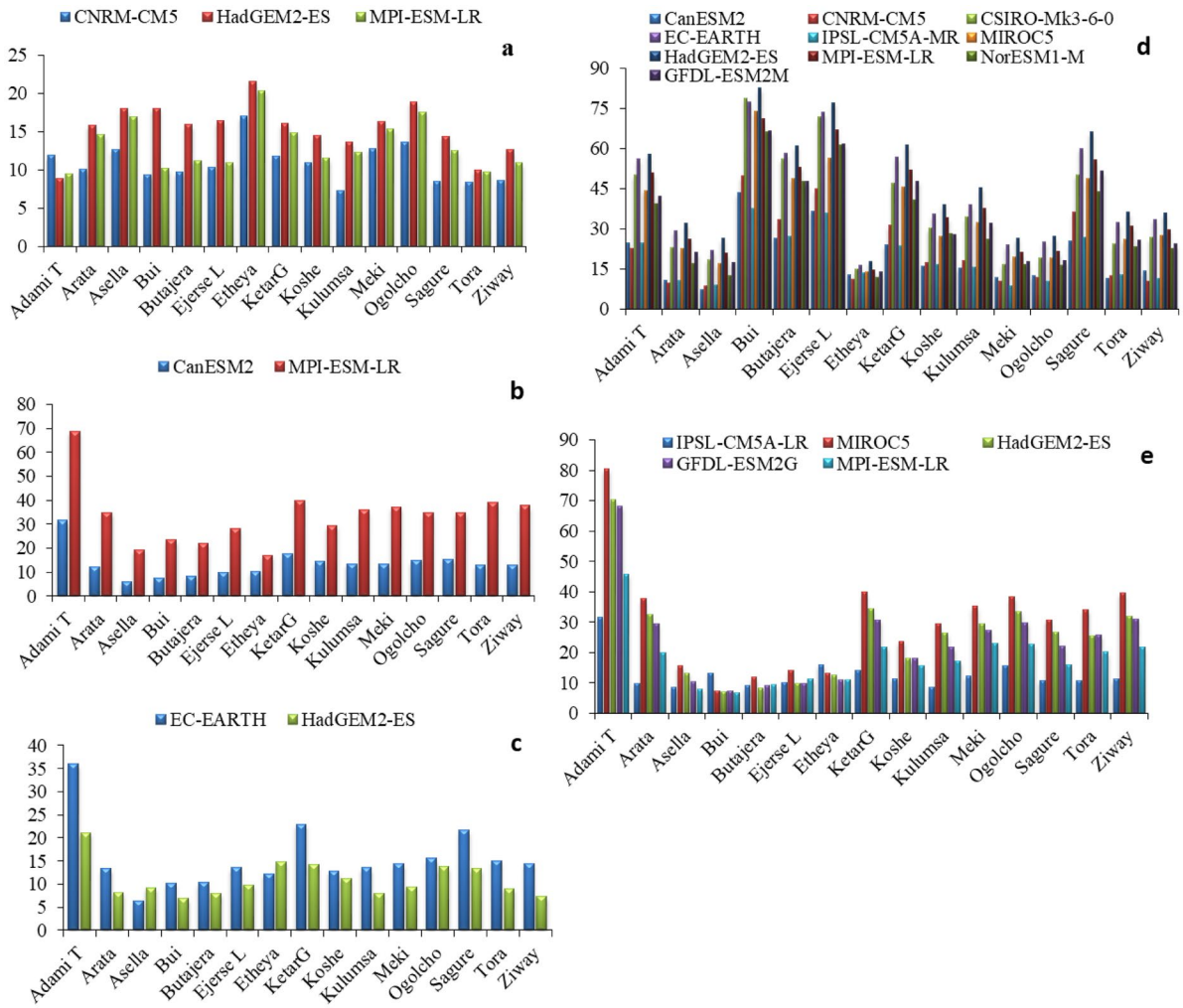

Fig. 8 Summer normalized root mean square error (%) of RCMs (**a**) CCLM4-8–17,**b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

climate model can simulate the spring seasonal rainfall better than EC-EARTH.

RCA4 downscaled the output of ten different GCMs for the spring season over most stations with a bias error in the range of 7.08 to 28.64%. The reliability of most models is very good for most stations except for the Adami-Tulu station (Fig. 9d). Exceptionally, spring rainfall is more than summer rainfall at Adami-Tulu station and the climate model cannot capture this microclimate variability. The GCMs downscaled by REMO2009 exhibited inconsistently over- or underestimated spring rainfall over all stations with an error in the range of − 5.81 to 51.49% (Fig. 9e).

Similarly, to the bias for the spring season, the accuracy of RCMs in simulating spring rainfall has been assessed by NRMSE and the results are presented in Fig. 10a–e. The performance assessment of the CCLM4-8–17 model, to downscale outputs of CNRM-CM5, HadGEM2-ES, and MPI-ESM-LR, is varied spatially over the study area with the NRMSE in the range of 12.79 to 101.77% (Fig. 10a). CRCM5 driven by CanESM2 and MPI-ESM-LR has simulated the spring rainfall over most stations with an error in the range of 11.92 to 87.37% (Fig. 10b). RACMO22T has downscaled the outputs of EC-EARTH and HadGEM2-ES to a regional scale. Figure 10c has presented the accuracy of the RACMO22T to

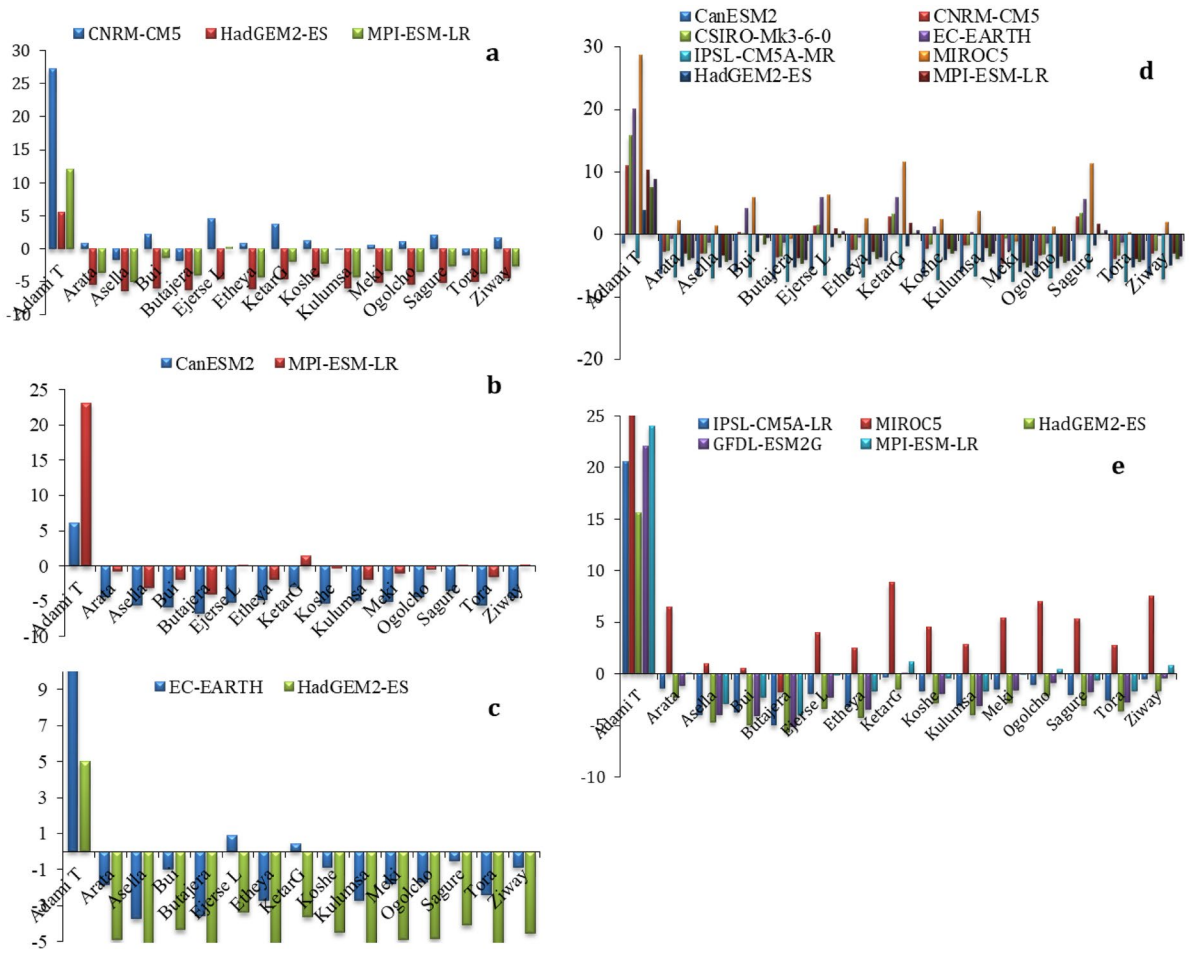

Fig. 9 Spatial distribution of spring biases (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

simulate spring rainfall spatially varied over most stations, with an error in the range of 10.89 to 76.77% to 41.71% and 17.63 to 34.76% respectively. Similarly, the performance assessment results for the outputs of RCA4 and REMO2009 regional climate models are displayed in Fig. 10d and e. Each RCM and their respective GCMs are simulated at a different level of accuracy. In general, the ability of each RCM and their respective GCMs to simulate spring rainfall differently over each station was evaluated, and there is no single RCM or GCM that best estimates the observed rainfall in spring over the stations.

Even if the dry (winter) season rainfall is insignificant compared to the summer and spring seasons, there is a small amount of rainfall from October to January. Figure 11a–e has shown that the bias error values of CCLM4-8–17 climate models of CNRM-CM5, HadGEM2-ES, and MPI-ESM-LR only underestimate the winter season rainfall over Adami-Tulu station with a bias error in the range of 2.68 to 6.44% (Fig. 11a). Compared to these three driving GCMs, CNRM-CM5 poorly simulates the winter season rainfall over all stations. CRCM5 driven by CanESM2 underestimated the winter season rainfall over some stations, with a bias in the range of 1.15 to 6.85%. However, except for the Adami-Tulu station, MPI-ESM-LR overestimated the winter season rainfall over all stations, which had a bias in the range of 3.75 to 22.76% (Fig. 11b).

The capability of RACMO22T to downscale EC-EARTH and HadGEM2-ES climate model outputs to a regional scale is presented in Fig. 11c. Both climate

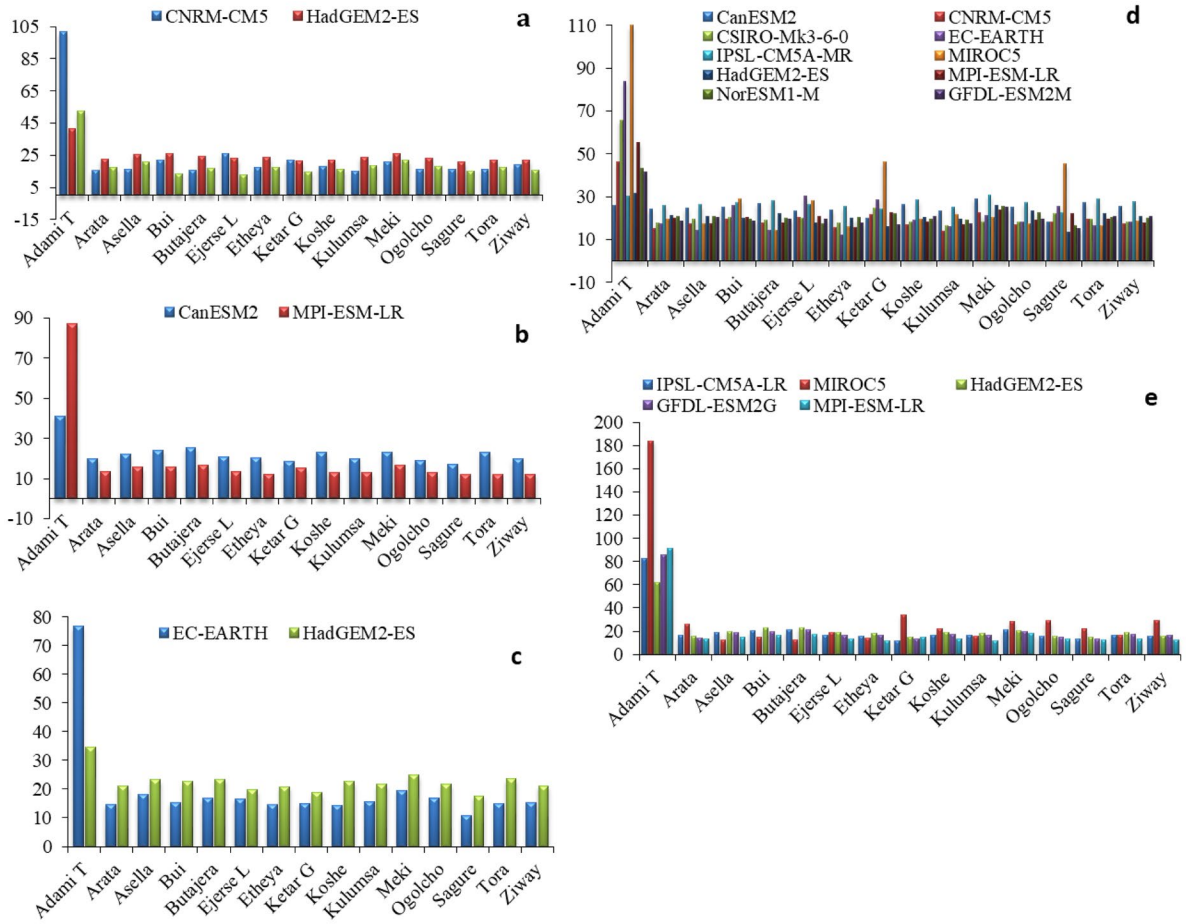

Fig. 10 Spring normalized root mean square error (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

models under- and overestimated the winter season rainfall over all stations, with biases in the range of − 6.7 to 9.5%. There is no significant difference in the bias errors between these two models over most stations. RCA4 downscaled the outputs of CMIP5 GCMs for winter season rainfall and is presented in Fig. 11d. Few GCMs underestimated winter rainfall over some stations, with biases in the range of 4.13 to 6.29%. However, the majority of GCMs overestimate the winter season rainfall with bias errors in the range of 10.69 to 57.06%. There is no single best model among these ten GCMs' output by the RCA4 regional climate model for simulating winter rainfall in the study area (Fig. 8d). REMO2009, which was derived from five different GCMs, is presented in Fig. 11e. Its performance in simulating winter season rainfall is presented in Fig. 8e. Except for Adami-Tulu stations, all GCMs overestimate the winter season rainfall over all stations with biases in the range of 10.11 to 34.22%.

Figure 12a shows the accuracy of the climate model of CCLM4-8–17 to simulate winter season rainfall over the stations varied with NRMSE in the range of 1245 to 157.78%. CNRM-CM5 poorly simulated the winter season rainfall at Bui, Ejerse-Lele, Ogolcho, and Sagure stations with an NRMSE of between 98.53 and 157.78. The spatial distribution of NRMSE between the driving HadGEM2-ES and MPI-ESM-LR varies between 22.18 and 62.17% over all stations. The performance of CRCM5 to simulate winter rainfall over most stations varied with an error in the range of 24.28 to 36.48% and 18.22 to

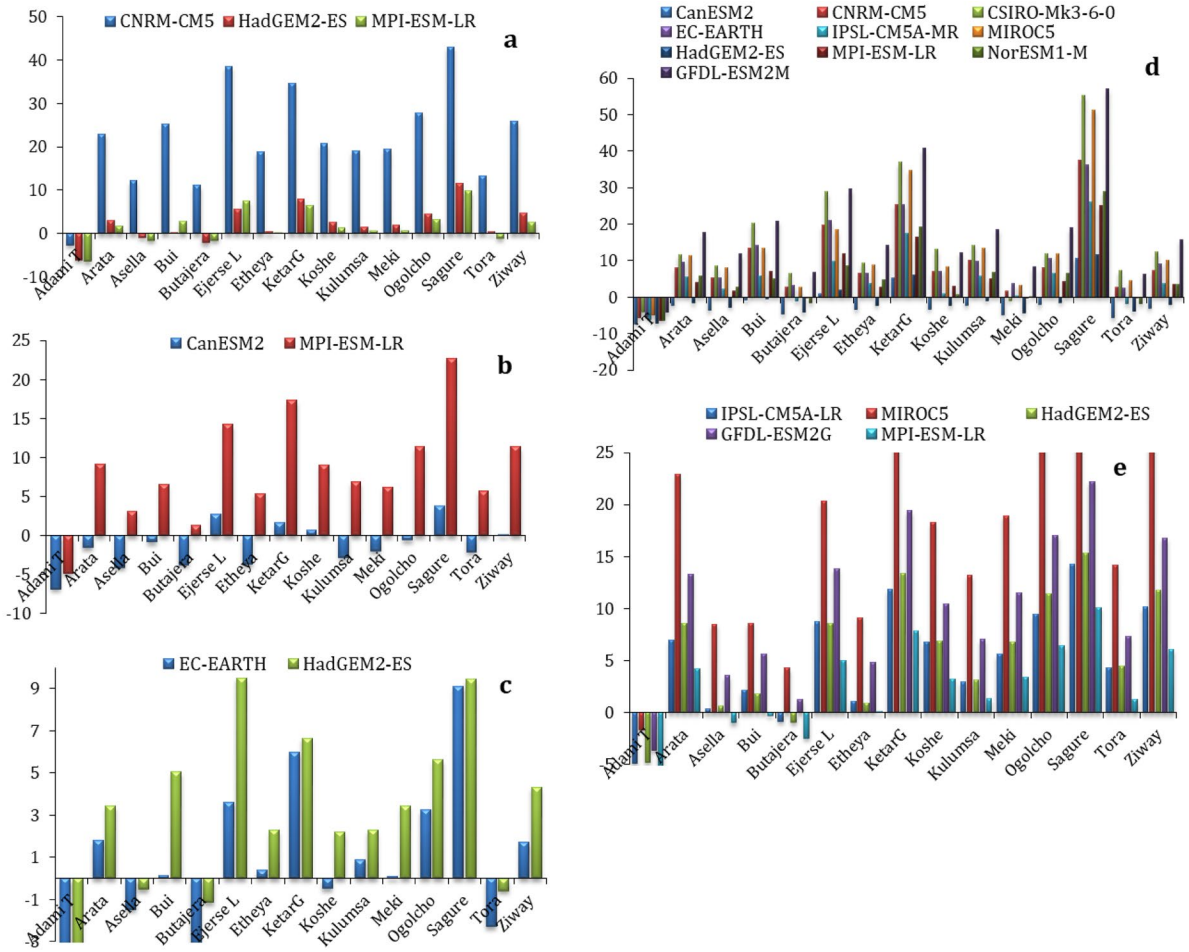

Fig. 11 Winter biases (%) of RCMs in the CRV Lakes subbasin (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009

97.78% respectively (Fig. 12b). Similarly, for RACMO22T driven by EC-EARTH and HadGEM2-ES, the NRME spatially varied over most stations in the range of 22.33 to 41.71% and 23.36 to 58.02% respectively (Fig. 12c). The accuracy of CRA4 to simulate winter rainfall over each station has varied with an error in the range of 19.02 to 235.02% (Fig. 12d). Exceptionally, all the driving GCMs of RCA4 poorly simulate winter season rainfall at Bui, Ejerse-Lele, Katar-genet, Ogolcho, and Sagure with an NRMSE in the range of 42.52 to 235.02%. The performance evaluation result for REMO2009 is shown in Fig. 12e and the NRMSE has varied from 9.07 to 127.28% for the study region. Relatively, MIROC5 is the climate model that poorly simulates the winter season rainfall over most stations, with the NRMSE in the range between 49 and 127%. However, the MPI-ESM-LR performed well with an NRMSE of less than 50% over most stations compared to the rest. In general, the performance of each RCM and their driving GCMs is evaluated against the observed winter (dry) season rainfall over all stations. All RCMs and their driving GCMs poorly simulated the winter season rainfall over most stations.

Annual cycle of rainfall

Figure 13a–e presents the performance evaluation of the annual cycle of rainfall for stations. The evaluations were done for each RCM against observed time series data from 1983 to 2005. The relative bias of CCLM4-8–17 driven by GCMs (CNRM-CM5,

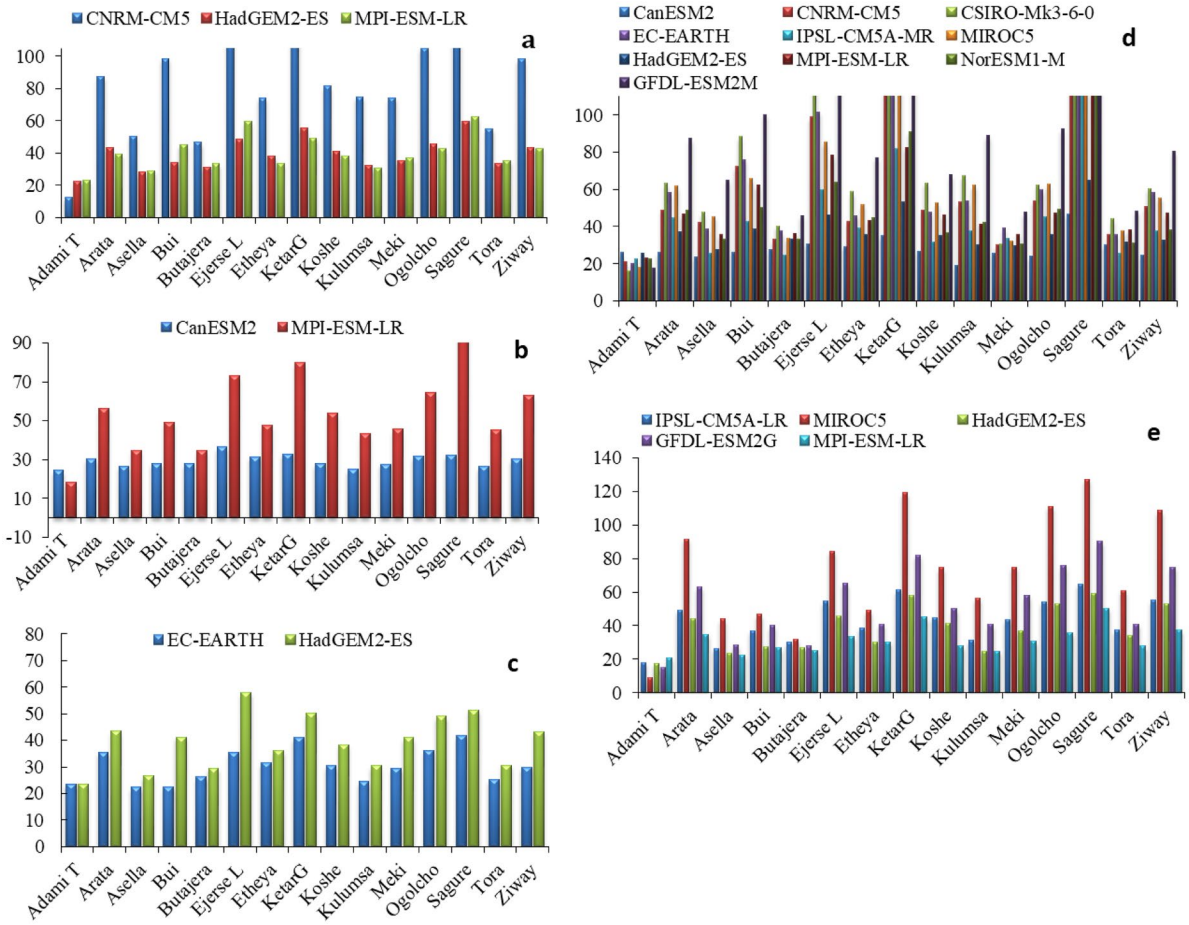

Fig. 12 Winter normalized root mean square error (%) of historical regional climate models (RCMs) (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

HadGEM2-ES, and MPI-ESM-LR) is presented in Fig. 13a. Except for Arata and Ejerse-Lele stations, CNRM-CM5 overestimated the annual rainfall with a bias in the range of 0.15 to 25.82% over most stations. HadGEM2-ES and MPI-ESM-LR underestimated the annual rainfall over most stations with an error in the range of − 0.03 to − 5.0% and over the estimate for Ziway station with an error in the range of 2.65 to 4.67%. CRCM5 driven by CanESM2 underestimated the annual rainfall over Arata, Asella, Butajera, Katar-genet, Koshe, Kulumsa, and Sagure with a bias value in the range of − 0.72 to − 2.81% stations. However, for the rest of the stations, the output of the CRCM5 model overestimated the annual rainfall by 0.01 to 1.81% (Fig. 13b). The MPI-ESM-LR climate model has overestimated the annual rainfall over all stations with a relative bias in the range of 1.84 to 11.36%.

RACMO22T driven by EC-EARTH and HadGEM2-ES over- and undersimulated the annual rainfall with a bias in the range of − 2.9 to 27.9% (Fig. 13c). The RCA4 model, driven by CSIRO-Mk3-6–0, MIOCO5, EC-EARTH, NorESM1-M, and GFDL-ESM2M, overestimated the annual rainfall with a bias range of 1 to 16.46% (Fig. 13d). However, the CanESM2, CNRM-CM5, CM5A-MR, and HadGEM2-ES models are simulating the annual rainfall with complex spatial configuration (under and overestimation) in the study region. Similarly, REMO2009 over- and undersimulated the annual rainfall of the study region with a bias in the range of − 2.9 to 27.9% (Fig. 13e).

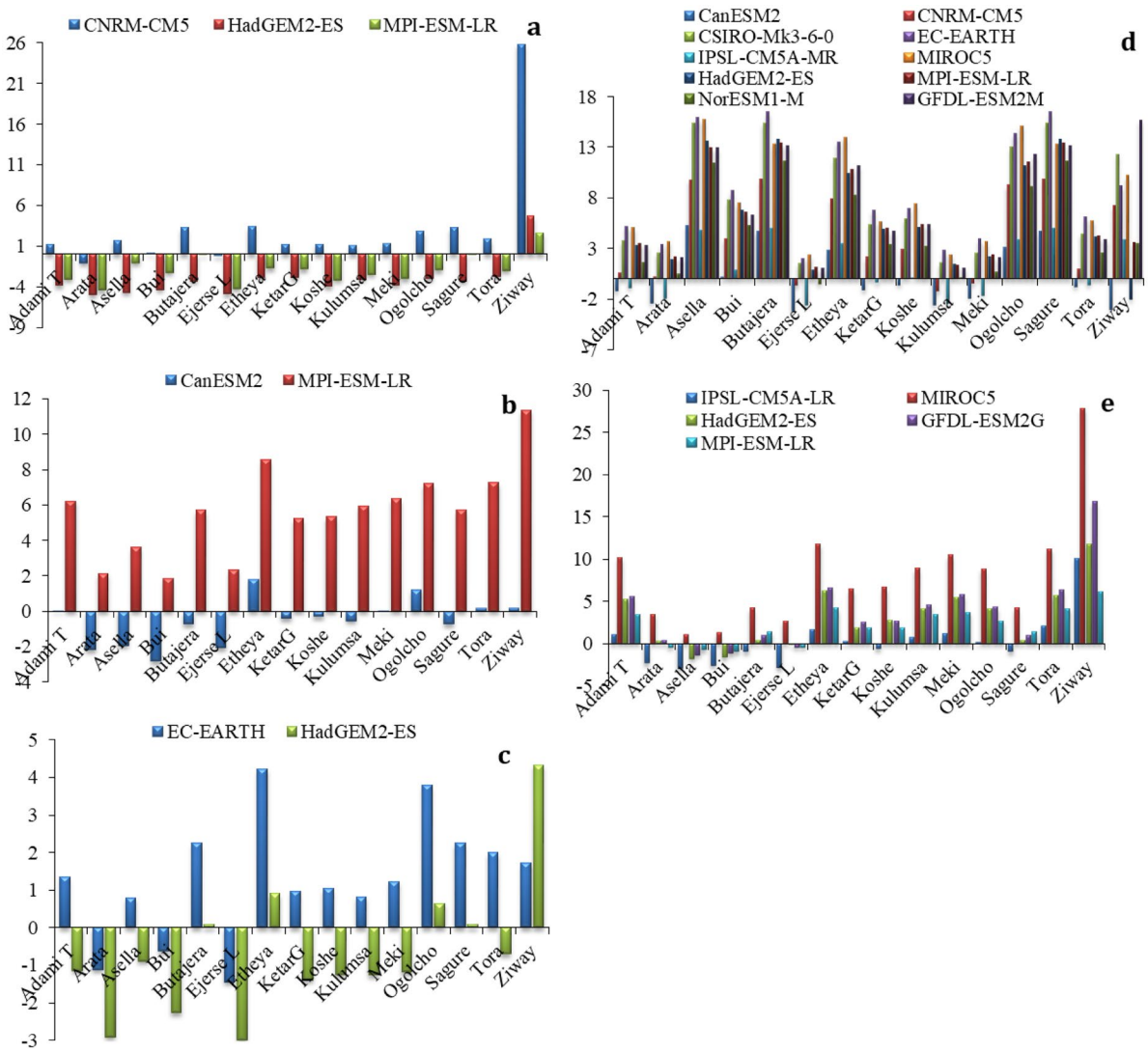

Fig. 13 Annual biases (%) of historical regional climate models (RCMs) (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

The magnitude of the calculated NRMSE of each model is presented in Fig. 11a–e. As Fig. 14a shows, the NRMSE for CCLM4-8–17 driven by CNRM-CM5, HadGEM2-ES, and MPI-ESM-LR climate models varied in the range of 7.09 to 16.18%, 11.38 to 19.36%, and 8.31 to 16.54% respectively. The annual computed NRMSE for the CRCM5 regional climate model with its irrespective GCMs (CanESM2 and MPI-ESM-LR) is spatially distributed over the study region with an error in the range of 3.8 to 10.58% and 8.55 to 31.13% respectively. When comparing these two models, CanESM2 has the lowest error difference over most stations (except for Butajera) than the MPI-ESM-LR climate model (Fig. 14b). Figure 14c shows that the spatial distribution of NRMSE for the RACMO22T regional climate model is complex over stations in the study region. Overall, the RMSE of the EC-EARTH and HadGEM2-ES climate models has varied in the range of 5.14 to 16.94% and 5.14 to 13.21% respectively. As Fig. 14d shows, the spatial distribution of RMSE of RCA4 over the study region is very complex and there is no single model that dominantly performed better over the

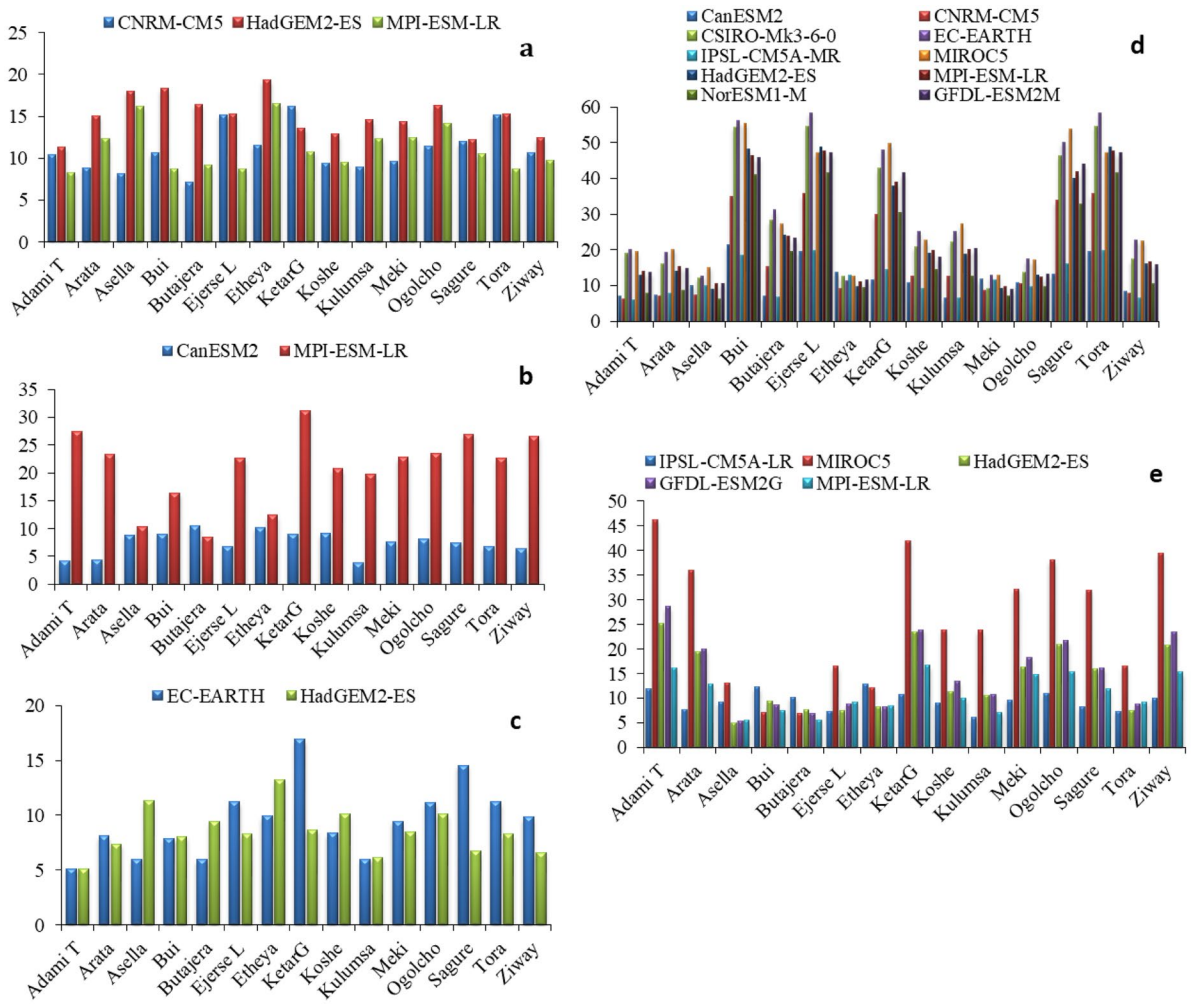

Fig. 14 Annual normalized root means square error (%) of RCMs (**a**) CCLM4-8–17, **b** CRCM5, **c** RACMO22T, **d** RCA4, and (**e**) REMO2009 in the CRV Lakes subbasin

stations. Relatively, CM5A-MR and CanESM2 have been considered suitable models among a group of GCMs RCA4 with an error in the range of 6.11 to 19.87% and 6.43 to 21.26% respectively, whereas for the rest of the climate models, computed values of RMSEs have been presented up to 58.34% and they have poorly simulated annual rainfall of the study region. In addition, the REMO2009 climate model has been characterized by NRMSE in the range of 4.98 to 46.33%. Each GCM of REMO2009 was presented with a different magnitude of NRMSE over the stations (Fig. 14e).

Uncertainty of RCMs

To see the aim of this study in-depth and characterize the difference between the observed and simulated historical monthly rainfall, we compared the bias and root mean square error for the monthly and seasonal (summer, spring, and winter) annual cycles for each RCM that downscaled the same GCM output. An overview of the results of different RCMs but for the same GCM is presented in Figs. 15 and 16 for bias and root mean square errors. The ability of four RCMs (CCLM4-8–17, CRCM5, RCA4,

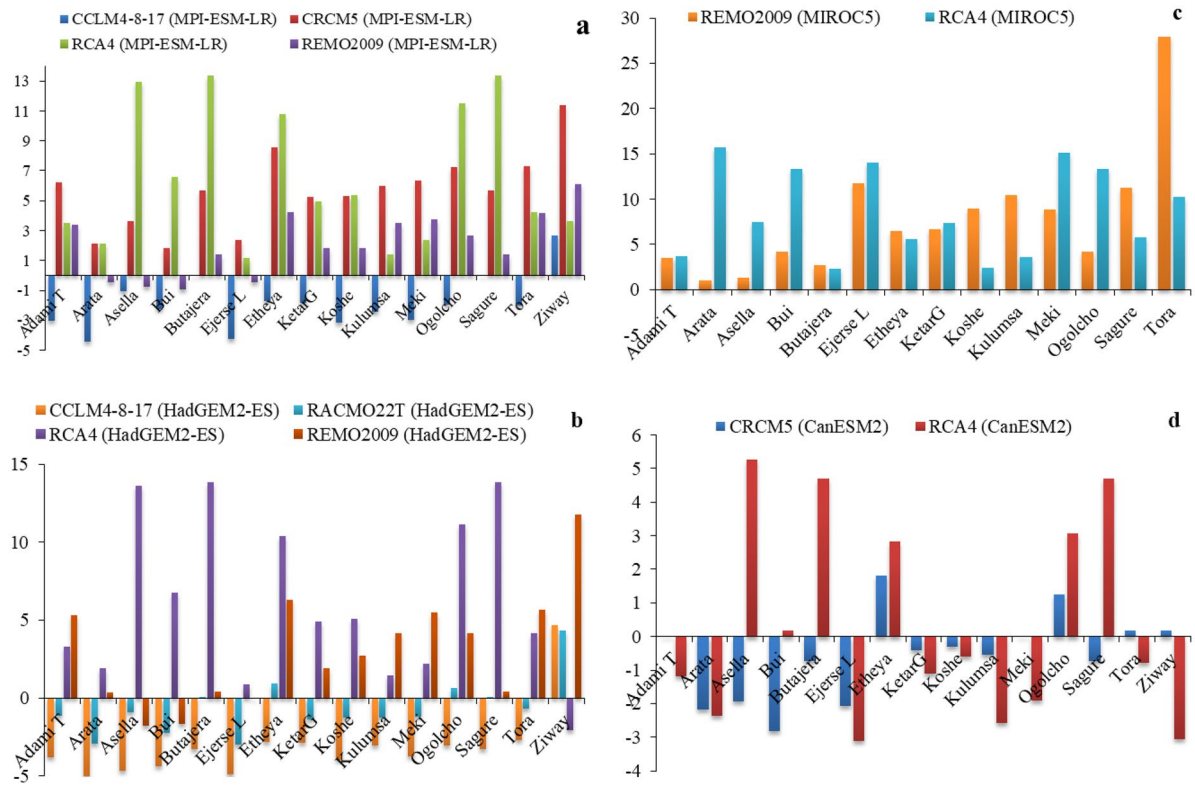

Fig. 15 Uncertainty (bias) of RCMs to downscale the same GCMs over target stations. Words in the brackets are the deriving GCM and words outside of the bracket are RCMs

and REMO22) to downscale the historical long-term (1983–2005) annual rainfall of the same GCM (M-MPI-ESM-LR) at local meteorological stations has performed differently (Fig. 15a). For example, CCLM4-8–17 underestimated annual rainfall over most stations with a bias in the range of 0.03 to 4.4% and overestimated annual rainfall at Ziway stations with a bias of 2.65%. CRCM5 and RCA overestimated the annual rainfall over all stations with a bias in the range of 1.18 to 11.37% and 1.15 to 12.96%, respectively. Similarly, REMO22 underestimated the annual rainfall over most stations with biases in the range of 0.45 to 0.96% and overestimated over most stations with biases in the range of 1.39 to 6.11%. Compared to the output of these climate models, REMO22 and CCLM4-8–17 are better performed at downscaling the output of the global climate model MPI-ESM-LR. This is also true for other RCMs with their respective GCMs. CCLM4-8–17, RACMO22T, RCA4, and REMO2009 were driven by the same HasGEM2-ES and differently simulated the annual total rainfall over all stations (Fig. 15b). REMO22T and RCA4 are driven by GCM (MIROC5) and their level of performance for simulated annual total rainfall is different over all stations (Fig. 15c). Similarly, canESM2 is simulated by CRCM5 and RCA4 (Fig. 15d), EC-EARTH is simulated by RCA4 and REMO2009, IPSL-CM5A-MR is simulated by GFDL-ESM2M, and CNRM-CM5 is simulated by CCLM4-8–17 and RCA4, and their performances are different from each other.

Similarly to the bias, the NRMSE results showed that each RCM performed differently to downscale the same GCM over the same spatial location (station) (Fig. 16a–d). Figure 16a presents the performance of each GCM to downscale MPI-ESM-LR spatially at the same location (stations). The NRMSE varies in the range of 8.31 to 16.54% for CCLM4-8–17, 8.55 to 31.13% for CRCM5, 9.64 to 47.83% for RCA4, and 5.53 to 16.70% for REMO2009 (Fig. 16a). Similarly, HadGEM2-ES has been downscaled by four RCMs (CCLM4-8–17, RACMO22T, RCA4,

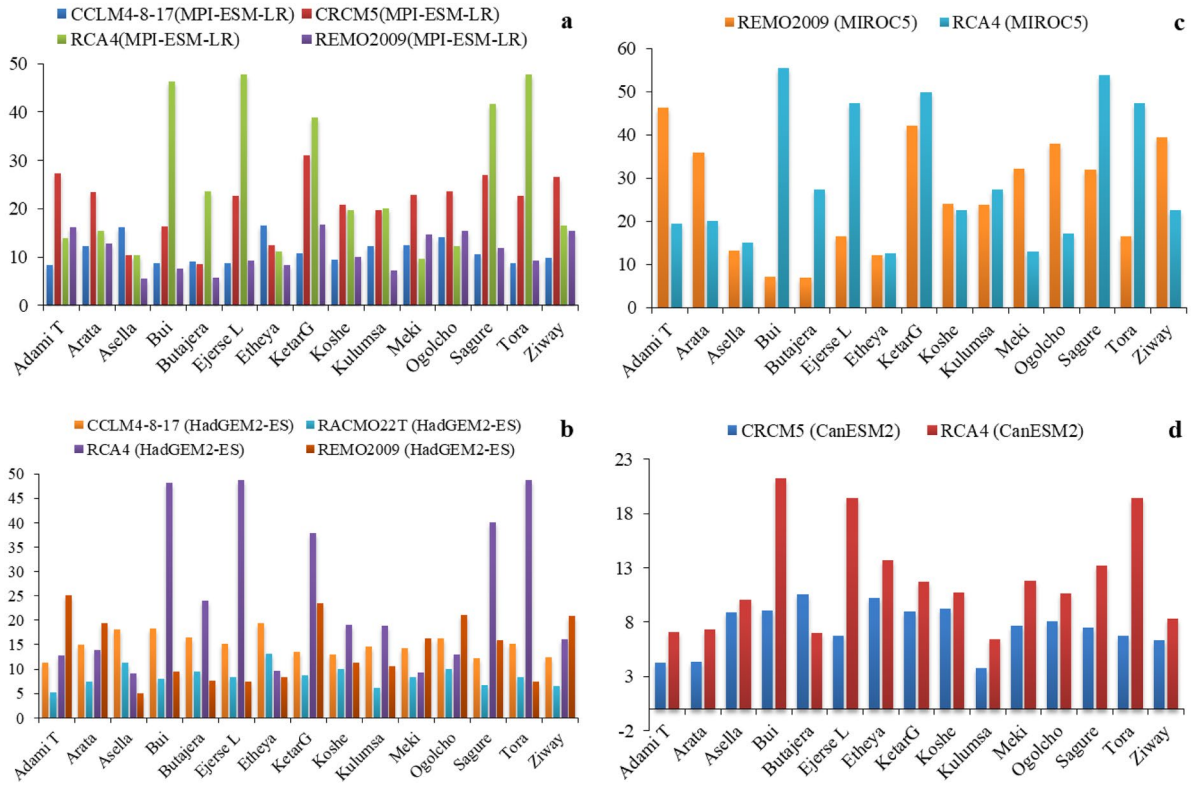

Fig. 16 Uncertainty (NRMSE) of RCMs to downscale the same GCMs over target stations. Words in the brackets are the deriving GCM and words outside of the bracket are RCMs

and REMO2009). The NRMSE varies in the range of 11.38 to 1936% for CCLM4-8–17, 5.14 to 13.21% for RACMO22T, 9.03 to 48.81% for RCA4, and 4.98 to 25.16% for REMO2009 (Fig. 16b). Figure 16c shows that MIROC5 was downscaled by REMO2009 and RCA4 and their performances in terms of NRMSE varied in the range of 6.94 to 46.33% and 12.62 to 55.50%, respectively (Fig. 16c). CRCM5 and RCA4 had downscaled the output of CanESM2, and the RMSE values varied in the range of 3.80 to 10.58% and 6.43 to 21.26%, respectively (Fig. 16d). Among these RCMs, RACMO22T downscales GCM (HadGEM2-ES) very well, with NRMSE values ranging from 5.14 to 13.21 over the stations studied.

Selection of best RCMs

Among the three performance indicators, bias has a significant effect on the ranking of RCMs/GCMs followed by NRMSE and CC over all stations considered for this study. For example, the Arata station was presented for demonstration and the weight of each indicator was given as bias (0.7894), followed by NRMSE (0.1722) and CC (0.0383), and the combined effect of NRMSE and CC is less than 25%. Equations 6–10 were coded in Excel 2016, and the entropy, degree of diversification, normalized weight, payoff matrix (P_{ij}), and L_p metrics were calculated, and subsequently, the RCMs/GCMs for each station were ranked. The entropy method assigns different weights to each indicator and helps to differentiate the degree of their importance. For each RCM/GCM, the values of L_p metrics were calculated using a compromise programming method, and the minimum value of the L_p metric was considered suitable, and the ranking configuration was gained accordingly for all RCMs/GCMs in Excel 2016.

As the results are shown in Table 2, the value of the L_p metric is wide-ranging from 0.000 (first rank) to 1.049 (least rank) over 22 GCMs. The regional climate model of REMO2009 driven by MIROC5 occupied the first rank, followed by CRCM5 driven

Table 2 Normalized payoff matrix (P_{ij}) of NRMSE, bias, CC, L_p metric value, and ranking pattern for Arata Meteorological Station for precipitation under different weight scenarios

Normalized payoff matrix						
RCMs	GCMs	NRMSE	bias	CC	L_p	Rank
CCLM4-8–17	CNRM-CM5	0.051	0.018	0.020	0.813	15
	HadGEM2-ES	0.036	0.054	0.037	0.504	10
	MPI-ESM-LR	0.028	0.043	0.047	0.626	13
CRCM5	CanESM2	0.044	0.000	0.042	1.049	22
	MPI-ESM-LR	0.070	0.088	0.047	0.171	2
RACMO22T	EC-EARTH	0.027	0.019	0.050	0.872	16
	HadGEM2-ES	0.025	0.016	0.045	0.914	19
RCA4	CanESM2	0.029	0.017	0.047	0.891	17
	CNRM-CM5	0.026	0.009	0.046	0.993	21
	CSIRO-Mk3-6–0	0.047	0.053	0.048	0.475	8
	EC-EARTH	0.056	0.073	0.050	0.297	5
	CM5A-MR	0.038	0.013	0.039	0.913	18
	MIROC5	0.042	0.071	0.052	0.355	6
	HadGEM2-ES	0.067	0.046	0.047	0.470	7
	MPI-ESM-LR	0.052	0.050	0.049	0.489	9
	NorESM1-M	0.042	0.023	0.045	0.776	14
	GFDL-ESM2M	0.050	0.047	0.042	0.522	11
REMO2009	CM5A-LR	0.027	0.015	0.046	0.917	20
	MIROC5	0.079	0.144	0.052	0.000	1
	HadGEM2-ES	0.064	0.075	0.048	0.266	4
	GFDL-ESM2G	0.062	0.079	0.048	0.246	3
	MPI-ESM-LR	0.037	0.048	0.052	0.553	12

by MPI-ESM-LR and REMO2009 driven by GFDL-ESM2G occupied the second and third ranks with L_p metric values of 0.000, 0.171, and 0.246 respectively. The 4th and 5th ranks were occupied by REMO2009 driven by HadGEM2-ES and RCA4 driven by EC-EARTH with L_p metric values of 0.266 and 0.297 respectively, whereas RCA4 driven by CNRM-CM5 and CRCM5 driven by CanESM2 occupied the 21st and 22nd ranks with L_p metric values of 0.993 and 1.049 respectively. The ranking of RCMs/GCMs can be used to select the best suitable RCMs/GCMs from a wide range of alternatives to CORDEX output and further processes such as bias correction and climate change impact study using hydrological models.

HadGEM2-ES was downscaled by REMO2009, RCA4, CCLM4-8–17, and RACMO22T modeling institutions. However, their ranking positions are 4th, 7th, 10th, and 19th respectively. Similarly, the MPI-ESM-LR global climate model results have been downscaled by CRCM5, RCA4, CCLM4-8–17, and REMO2009 institutions and occupied the ranking positions of 2nd, 9th, 12th, and 13th respectively. MIROC5 was downscaled by two different RCM (REMO2009 and RCA4) institutes and occupied the ranking positions of 1st and 6th. This is also true for other similar GCMs but downscaled by different modeling institutions. This may be because of the difference between the downscaling institutes and their parametrization of the physical condition of the climate systems. A similar process has been repeatedly applied to the rest of the meteorological stations and their results are presented in Tables 3 and 4.

Discussion

One of the most significant meteorological characteristics with notable regional and temporal variation is precipitation. For sustainable catchment management and water infrastructure projects, it is urgently necessary to have accurate future climate projections. RCMs are the most advanced tool for simulating future climate conditions. However, it is necessary to assess its efficiency in simulating spatial and temporal

Table 3 Normalized payoff matrix (Pij) of NRMSE, bias, and CC, L_p metric value, and ranking pattern for each RCM and their driving GCMs for stations in the Katar watershed

RCMs	GCMs	Arata		Asella		Etheya		Katar G.		Kulumsa		Ogolcho		Sagure	
		L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank
CCLM4-8–17	CNRM-CM5	0.813	15	0.181	16	0.188	14	1.049	15	0.446	15	0.825	15	1.311	15
	HadGEM2-ES	0.504	10	0.035	1	0.017	1	1.286	18	0.258	10	0.520	6	1.400	17
	MPI-ESM-LR	0.626	13	0.055	2	0.036	2	1.497	21	0.340	14	0.639	9	1.625	19
CRCM5	CanESM2	1.049	22	0.133	12	0.143	9	1.331	19	0.581	20	1.088	22	1.662	20
	MPI-ESM-LR	0.171	2	0.118	11	0.108	5	0.428	9	0.102	7	0.179	2	0.714	10
RACMO22T	EC-EARTH	0.872	16	0.232	17	0.197	16	1.041	14	0.532	18	0.909	17	1.255	14
	HadGEM2-ES	0.914	19	0.111	8	0.095	3	1.600	22	0.516	17	0.918	19	1.837	21
RCA4	CanESM2	0.891	17	0.138	13	0.108	6	1.204	17	0.568	19	0.809	14	1.345	16
	CNRM-CM5	0.993	21	0.295	21	0.251	18	0.528	10	0.312	12	1.053	21	0.461	8
	CSIRO-Mk3-6–0	0.475	8	0.101	7	0.184	13	0.171	5	0.081	4	0.689	10	0.152	5
	EC-EARTH	0.297	5	0.061	3	0.152	10	0.045	1	0.047	1	0.476	5	0.045	1
	CM5A-MR	0.913	18	0.142	14	0.111	7	1.025	13	0.604	22	0.828	16	1.126	11
	MIROC5	0.355	6	0.079	4	0.156	11	0.126	3	0.080	3	0.551	8	0.115	3
	HadGEM2-ES	0.470	7	0.090	6	0.184	12	0.077	2	0.048	2	0.690	11	0.098	2
	MPI-ESM-LR	0.489	9	0.115	10	0.207	17	0.178	6	0.088	5	0.699	12	0.191	6
	NorESM1-M	0.776	14	0.249	18	0.267	19	0.422	8	0.255	9	0.991	20	0.427	7
	GFDL-ESM2M	0.522	11	0.113	9	0.192	15	0.157	4	0.094	6	0.729	13	0.129	4
REMO2009	CM5A-LR	0.917	20	0.153	15	0.107	4	1.490	20	0.599	21	0.912	18	1.973	22
	MIROC5	0.000	1	0.089	5	0.137	8	0.255	7	0.103	8	0.000	1	0.595	9
	HadGEM2-ES	0.266	4	0.268	19	0.298	21	0.691	12	0.306	11	0.267	4	1.166	12
	GFDL-ESM2G	0.246	3	0.269	20	0.276	20	0.680	11	0.318	13	0.247	3	1.171	13
	MPI-ESM-LR	0.553	12	0.298	22	0.303	22	1.084	16	0.450	16	0.526	7	1.518	18

Table 4 Normalized payoff matrix (P_{ij}) of NRMSE, bias, and CC, L_p metric value, and ranking pattern for Meki watershed stations for each RCM and their driving GCMs

RCMs	GCMs	Adami-T		Bui		Butajera		Ejerse-L		Koshe		Meki		Tora	
		L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank	L_p	Rank
CCLM4-8–17	CNRM-CM5	1.46	13	1.80	13	0.78	20	1.63	12	0.34	13	0.60	14	0.48	19
	HadGEM2-ES	1.88	16	1.51	11	0.47	9	1.78	14	0.29	10	0.41	7	0.31	11
	MPI-ESM-LR	2.12	19	2.12	19	0.67	15	2.44	22	0.40	16	0.50	12	0.37	15
CRCM5	CanESM2	1.92	17	1.93	15	0.54	10	2.26	17	0.48	21	0.71	20	0.44	18
	MPI-ESM-LR	0.43	3	1.60	12	0.64	14	1.32	10	0.10	5	0.08	2	0.04	2
RACMO22T	EC-EARTH	1.46	14	2.13	20	0.79	21	1.97	15	0.45	19	0.70	19	0.51	20
	HadGEM2-ES	2.32	21	2.10	18	0.62	13	2.39	21	0.41	17	0.63	16	0.40	16
RCA4	CanESM2	2.29	20	1.09	9	0.68	16	1.33	11	0.43	18	0.47	10	0.32	12
	CNRM-CM5	1.95	18	0.64	8	0.39	8	0.67	8	0.33	12	0.66	18	0.59	22
	CSIRO-Mk3-6–0	0.88	6	0.07	2	0.07	3	0.07	3	0.07	3	0.41	8	0.27	8
	EC-EARTH	0.78	5	0.08	3	0.05	2	0.05	2	0.03	1	0.47	9	0.13	3
	CM5A-MR	2.33	22	1.14	10	0.60	11	1.23	9	0.45	20	0.49	11	0.35	13
	MIROC5	0.92	7	0.12	4	0.15	5	0.33	7	0.11	8	0.55	13	0.17	4
	HadGEM2-ES	1.16	9	0.02	1	0.02	1	0.04	1	0.03	2	0.60	15	0.22	6
	MPI-ESM-LR	1.17	10	0.22	5	0.12	4	0.19	4	0.08	4	0.79	22	0.24	7
	NorESM1-M	1.57	15	0.30	7	0.18	7	0.32	6	0.17	9	0.64	17	0.42	17
	GFDL-ESM2M	1.20	11	0.27	6	0.16	6	0.25	5	0.10	6	0.72	21	0.28	10
REMO2009	CM5A-LR	1.37	12	1.83	14	0.62	12	2.30	19	0.56	22	0.01	1	0.53	21
	MIROC5	0.00	1	2.18	21	0.83	22	1.68	13	0.11	7	0.25	4	0.02	1
	HadGEM2-ES	0.48	4	2.03	16	0.68	17	2.38	20	0.36	14	0.21	3	0.28	9
	GFDL-ESM2G	0.37	2	2.09	17	0.72	18	2.27	18	0.30	11	0.35	5	0.21	5
	MPI-ESM-LR	1.02	8	2.22	22	0.77	19	2.20	16	0.39	15	0.35	5	0.35	14

scales. Therefore, the reliability and accuracy of five RCMs have been evaluated to simulate monthly, seasonal, and annual cycles of rainfall using the statistical metrics of bias, NRMSE, and CC for the Katar and Mekis subbasins of the CRV. Each RCM and its driving GCM simulated rainfall temporal variability with a complex spatial configuration (under- and oversimulation). There is no single RCM or GCM output that has better performance in simulating the rainfall cycle of the study regions. The results agree with similar studies in different regions (Luhunga et al., 2016; Sarr et al., 2015; Stefanidis et al., 2020; Worku et al., 2018). Intensive studies have been conducted to evaluate the skill of global climate models (GCMs) to simulate climate variables, mainly rainfall and temperature, in different parts of the globe (Shi et al., 2018; Su et al., 2013), and there is a large difference between intermodel performance and simulated climate variables of the same region or station. Similar studies were conducted in Upper Blue Nile Basin of Jemma subbasin by Worku et al. (2018), and they recommend that the ensemble mean of the models was superior in capturing the annual and seasonal pattern of the rainfall and had better correlation with observed than any single model (RCM).

Computed values of L_p can be present in Tables 3 and 4 for each RCM driven by a similar GCM over the stations. Based on the results, there is no single RCM that performs better consistently over the study region. The ability of each RCM to simulate similar driving GCM is compared for the whole study area and the level of their performance is not the same. Based on these three-performance evaluation criteria, each RCM cannot persuasively simulate the output of driving GCM to regional conditions and is unable to represent the spatiotemporal variability of rainfall. A similar uncertainty analysis of RCMs has been conducted by Lahunga et al. (2016) in Tanzania on rainfall and temperature. The results show that there is no similarity

Table 5 Top five GCMs/RCMs identified for the Katar subbasin

RCMs	GCMs	Arata	Asella	Iteya	Katar	Kulumsa	Ogolcho	Sagure	Weighted rank	
		Individual rank at each station								
CRCM5	MPI-ESM-LR	2	11	5	9	7	2	10	46	5
RCA4	EC-EARTH	5	3	10	1	1	5	1	26	1
	MIROC5	6	4	11	3	3	8	3	38	2
	HadGEM2-ES	7	6	12	2	2	11	2	42	4
REMO2009	MIROC5	1	5	8	7	8	1	9	39	3

between RCMs in simulating the respective driving GCM outputs. Christensen and Christensen (2007) and Foley (2010) have made intermodal comparisons for EU PRUDENCE (Prediction of regional scenarios and uncertainty for Defining EuropeaN Climate Changes, risks, and effects) using the same driving GCM, and the difference in RCM design can result in quite different outcomes. Generally, there are uncertainties arising from natural climatic fluctuations and future atmospheric emissions, in addition to those derived from model construction (Christensen & Christensen, 2007). And also, there is no single RCM or driving GCM that is dominantly occupied by the first three ranks in the study regions (Tables 3 and 4). This has been confirmed by a similar study conducted by Su et al. (2013) and Raju et al. (2016). Stefanidis et al.(2020) have evaluated the performance of RCMs to simulate seasonal precipitation over the mountainous region of Central Pindus (Greece) and they found that, except REMO, most RCMs fail to represent the temporal variability of precipitation time series in their study region.

The relevance of the compromise programming of the multicriteria decision-making approach was applied to identify the best climate model for the study regions. Following the procedures, the top five (which has a lower weight rank) were selected for both the Katar and Meki study regions, and the results are summarized in Tables 5 and 6. Similar studies were conducted by Raju et al. (2016) to select the best climate model for India.

Conclusion

In this study, we applied a multicriteria score-based methodology to evaluate the outputs of various regional climate models throughout the period of 1983–2005 in Ethiopia's CRV Lakes subbasin. This study evaluated the performance of 22 GCMs' (5 RCMs downscaled) precipitation outputs using a variety of evaluation criteria. To choose the right GCMs for the study region, a multi-criteria decision method was employed that could provide an independent decision based on the criteria pertinent to their particular application and research objectives. Based on this, the following conclusions were drawn out by the study.

1. Each RCM cannot reproduce the monthly, seasonal, and annual cycle of rainfall over most stations, and

Table 6 Top Five GCMs/RCMs identified for the Meki subbasin

RCMs	GCMs	Adami-T	Bui	Butajera	Ejerse-L	Koshe	Meki	Ziway	Weighted rank	
		Individual rank at each station								
CRCM5	MPI-ESM-LR	3	12	14	10	5	2	2	50	4
RCA4	CSIRO-Mk3-6–0	6	2	3	3	3	8	8	41	2
	EC-EARTH	5	3	2	2	1	9	4	29	1
	MIROC5	7	4	5	7	8	13	6	54	5
	HadGEM2-ES	9	1	1	1	2	15	7	42	3

there is intermodal uncertainty in simulating the same GCM. This difference might have resulted from the assumption and formulation of RCMs that represent the physical conditions or the climate conditions of the study regions. For example, similar driving GCM but downscaled by different institutions simulated the climate condition of the same station differently for bias and NRMSE.

2. The rank obtained for each station over the CRV subbasin varies and there are no RCMs/GCMs occupying the first three ranks for each station. This demonstrates the importance of evaluating the RCMs before applying them for impact or other related studies and choosing the best climate model among available alternatives (Tables 3 and 4).
3. Instead of using a single climate model to characterize the climate conditions over the study region, it is recommended to use ensembles of the best climate models or the best of each climate model.
4. The approach might be used in many places and serve as a guide for choosing the best GCMs for use in studies on the impacts of regional climate change.

Acknowledgements The authors acknowledge the National Meteorological Agency of Ethiopia for the provision of climate data. The climate data used for this study were sourced from the CORDEX-Africa database (https://esgf-node.llnl.gov/search/esgf-llnl/). The authors would like to thank Mr. Jacob Agyekum from Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana, for providing Zoom training on Climate Data Operator (CDO).

Author contribution All authors significantly contributed to the development of this manuscript. Sisay Kebede Balcha oversaw the conceptualization, data collection, software, data analysis, investigation, and preparation of the original draft. The manuscript was reviewed, edited, and improved by Taye Alemayehu Hulluka and Gebiaw T. Ayele. The overall research work for this study was overseen by Adane Abebe Awass and Amare Bantider. All authors have read and agreed to the published version of the manuscript.

Funding This work was supported by the Water Security and Sustainable Development Hub funded by the UK Research and Innovation's Global Challenges Research Fund (GCRF) (grant number: ES/S008179/1).

Data availability All datasets, raw or preprocessed, are available upon request from the corresponding author. However, permission is required for observed data collected from the National Meteorological Agency of Ethiopia.

Declarations

Competing interests The authors declare no competing interests.

References

Ademe, F., Kibret, K., Beyene, S., Getinet, M., & Mitike, G. (2020). Rainfall analysis for rain-fed farming in the Great Rift Valley Basins of Ethiopia. *Journal of Water and Climate Change, 11*, 812–828. https://doi.org/10.2166/wcc.2019.242

Alemseged, T. H., & Tom, R. (2015). Evaluation of regional climate model simulations of rainfall over the Upper Blue Nile basin. *Atmospheric Research, 161–162*, 57–64. https://doi.org/10.1016/j.atmosres.2015.03.013

Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology, 6*, 661–675. https://doi.org/10.1002/joc.3370060607

Barrios, A., Trincado, G., & Garreaud, R. (2018). Alternative approaches for estimating missing climate data: Application to monthly precipitation records in south-central Chile. *Forest Ecosystems, 5*, 10. https://doi.org/10.1186/s40663-018-0147-x

Beula, T. M. N., & Prasad, G. E. (2012). Multiple criteria decision making with compromise programming. *International Journal of Engineering, Science and Technology, 4*, 4083–4086.

Bhattacharya, T., Khare, D., & Arora, M. (2020). Evaluation of reanalysis and global meteorological products in Beas river basin of North - Western Himalaya. *Environmental Systems Research*, 1–29. https://doi.org/10.1186/s40068-020-00186-1

Bosilovich, M. G., Kennedy, J., Dee, D., Allan, R., & Neill, A. O. (2013). On the reprocessing and reanalysis of observations for climate, in: Hurrell, A.G. and J.W. (Ed.), Climate science for serving society: Research, modeling and prediction priorities. Springer, pp. 51–71. https://doi.org/10.1007/978-94-007-6692-1

Breach, P. A., Simonovic, S. P., & Yang, Z. (2016). Global climate model selection for analysis of uncertainty in climate change impact assessments of hydro-climatic extremes. *American Journal of Climate Change, 5*, 502–525. https://doi.org/10.4236/ajcc.2016.54036

Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology, 58*, 11–27.

Christensen, J. H., & Christensen, O. B. (2007). A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climate Change, 81*, 7–30. https://doi.org/10.1007/s10584-006-9210-7

Di Luca, A., de Elía, R., & Laprise, R. (2015). Challenges in the quest for added value of regional climate dynamical downscaling. *Current Climate Change Reports, 1*, 10–21. https://doi.org/10.1007/s40641-015-0003-9

Di Virgilio, G., Evans, J. P., Di Luca, A., Olson, R., Argüeso, D., Kala, J., Andrys, J., Hoffmann, P., Katzfey, J. J., & Rockel, B. (2019). Evaluating reanalysis-driven CORDEX regional climate models over Australia: Model performance and errors. *Climate Dynamics, 53*, 2985–3005. https://doi.org/10.1007/s00382-019-04672-w

Dibaba, W. T., Miegel, K., & Demissie, T. A. (2019). Evaluation of the CORDEX regional climate models performance in simulating climate conditions of two catchments in Upper Blue Nile Basin. *Dynamics of Atmospheres and Oceans, 87*, 101104. https://doi.org/10.1016/j.dynatmoce.2019.101104

Endris, H. S., Omondi, P., Jain, S., Lennard, C., Hewitson, B., Chang'a, L., Awange, J. L., Dosio, A., Ketiem, P., Nikulin, G., Panitz, H.-J. R., Chner, M. B., Stordal, F., & Tazalika, L. (2013). Assessment of the performance of CORDEX regional climate models in simulating East African rainfall. *Journal of Climate*, 26. https://doi.org/10.1175/JCLI-D-12-00708.1.

Foley, A. M. (2010). Uncertainty in regional climate modelling : A review. *Progress in Physical Geography, 34*, 647–670. https://doi.org/10.1177/0309133310375654

Fordham, D. A., Wigley, T. M. L., & Brook, B. W. (2011). Multi-model climate projections for biodiversity risk assessments. *Ecological Applications, 21*, 3317–3331.

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. *Journal of Geophysical Research, 118*, 4154–4167. https://doi.org/10.1002/jgrd.50269

Garibay, V. M., Gitau, M. W., Kiggundu, N., & Moriasi, D. (2021). Evaluation of reanalysis precipitation data and potential bias correction methods for use in data - scarce areas. *Water Resources Management, 35*, 1587–1602. https://doi.org/10.1007/s11269-021-02804-8

Giorgi, F. (2019). Thirty years of regional climate modeling : Where are we and where are we going next ? Journal of Geophysical Research : Atmospheres. *Journal of Geophysical Research: Atmospheres, 124*, 5696–5723. https://doi.org/10.1029/2018JD030094

Giorgi, F., & Gutowski, W. J. (2016). Coordinated experiments for projections of regional climate change. *Current Climate Change Reports, 2*, 202–210. https://doi.org/10.1007/s40641-016-0046-6

Giorgi, F., Jones, C., & Asrar, G. (2009). Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bulletin, 53*, 175–183.

Goodarzi, M. R., Pooladi, R., & Niazkar, M. (2022). Evaluation of satellite-based and reanalysis precipitation datasets with gauge-observed data over Haraz-Gharehsoo. *Sustainability, 14*, 13051. https://doi.org/10.3390/su142013051

Gordon, N., & Shaykewich, J. (2000). Guidelines on performance assessment of public weather services. WMO/TD No.1023 (p. 67). Geneva, Switzerland.

Gyamfi, C., Tindan, J. Z. O., & Kifanyi, G. E. (2021). Evaluation of CORDEX Africa multi-model precipitation simulations over the Pra River Basin. *Ghana. Journal of Hydrology: Regional Studies, 35*, 100815. https://doi.org/10.1016/j.ejrh.2021.100815

Hartkamp, A. D., Beurs, K. D., Stein, A., & White, J. W. (1999). Interpolation techniques for for climate variables. *NRG-GIS Ser., 99–01*, 1–35.

Hassler, B., & Lauer, A. (2021). Comparison of reanalysis and observational precipitation datasets including EAR5 and WFDE5. *Atmosphere (basel)., 12*, 1462. https://doi.org/10.3390/atmos12111462

Hidalgo, H. G., & Alfaro, E. J. (2015). Skill of CMIP5 climate models in reproducing 20th century. *International Journal of Climatology, 35*, 3397–3421. https://doi.org/10.1002/joc.4216

Hoffman, R. N., Privé, N., & Bourassa, M. (2017). Comments on "Reanalyses and observations: What's the difference?" *Bulletin of the American Meteorological Society, 98*, 2455–2460. https://doi.org/10.1175/BAMS-D-17-0008.1

Ilori, O. W., & Balogun, I. A. (2021). Evaluating the performance of new CORDEX-Africa regional climate models in simulating West African rainfall. *Modeling Earth Systems and Environment*, 1–25. https://doi.org/10.1007/s40808-021-01084-w

Ingole, P. V., & Nichat, M. K. (2015). Landmark based shortest path detection by using Dijkestra Algorithm and Landmark based shortest path detection by using Dijkestra Algorithm and Haversine Formula. *International Journal of Engineering Research and Applications, 3*, 162–165.

Jury, M. R. (2015). Statistical evaluation of CMIP5 climate change model simulations for the Ethiopian highlands. *International Journal of Climatology, 35*, 37–44. https://doi.org/10.1002/joc.3960

Kalognomou, E.-A., Lennard, C., Shongwe, M., Pinto, I., Favre, A., Kent, M., Hewitson, B., Dosio, A., Nikulin, G., Panitz, H.-J.R., & Chner, M. B. (2013). A diagnostic evaluation of precipitation in CORDEX models over Southern Africa. *Journal of Climate, 26*, 9477–9506. https://doi.org/10.1175/JCLI-D-12-00703.1

Kamworapan, S., & Surussavadee, C. (2019). Evaluation of CMIP5 global climate models for simulating climatological temperature and precipitation for southeast Asia. *Advances in Meteorology, 2019*, 18. https://doi.org/10.1155/2019/1067365

Kim, J., Waliser, D. E., Mattmann, C. A., Goodale, C. E., Hart, A. F., Zimdars, P. A., Crichton, D. J., Jones, C., Nikulin, G., Hewitson, B., Jack, C., Lennard, C., & Favre, A. (2014). Evaluation of the CORDEX-Africa multi-RCM hindcast: Systematic model errors. *Climate Dynamics, 42*, 1189–1202. https://doi.org/10.1007/s00382-013-1751-7

Luhunga, P., Botai, J., & Kahimba, F. (2016). Evaluation of the performance of CORDEX regional climate models in simulating present climate conditions of Tanzania. *Journal of Southern Hemisphere Earth Systems Science, 66*(1), 32–54. https://doi.org/10.22499/3.6601.005

Lin, C.-Y., & Tung, C.-P. (2017). Procedure for selecting GCM datasets for climate risk assessment. *Terrestrial,*

Atmospheric and Oceanic Sciences, 28, 43–55. https://doi.org/10.3319/TAO.2016.06.14.01(CCA)1

Luhunga, P., Botai, J., & Kahimba, F. (2016). Evaluation of the performance of CORDEX regional climate models in simulating present climate conditions of Tanzania. *Journal of Southern Hemisphere Earth Systems Science, 66*, 32–54.

Ly, S., River, M., Vientiane, C., Charles, C., & Degre, A. (2011). Geostatistical interpolation of daily rainfall at catchment scale : The use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences., 15*, 2259–2274. https://doi.org/10.5194/hess-15-2259-2011

Majmder, M. (2015). Multi criteria decision making, in: Impact of urbanization on water shortage in face of climatic aberrations,. pp. 35–48. https://doi.org/10.1007/978-981-4560-73-3.

Mardani, A., Jusoh, A., MD Nor, K., Khalifah, Z., Zakwan, N., & Valipour, A. (2015). Multiple criteria decision-making techniques and their applications – A review of the literature from 2000 to 2014. *Economic Research*, *28*, 516–571.https://doi.org/10.1080/1331677X.2015.1075139

Min, E., Hazeleger, W., Oldenborgh, G. J. van, & Sterl, A. (2013). Evaluation of trends in high temperature extremes in north-western Europe in regional climate models. *Environmental Research Letters*, 8. https://doi.org/10.1088/1748-9326/8/1/014011.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidline for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers., 50*, 885–900.

Nkiaka, E., Nawaz, N. R., & Lovett, J. C. (2017). Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region : Case study of the Logone catchment. *Lake Chad Basin, 18*, 9–18. https://doi.org/10.1002/met.1600

Okafor, G., Annor, T., Odai, S., & Agyekum, J. (2019). Volta basin precipitation and temperature climatology: Evaluation of CORDEX-Africa regional climate model simulations. *Theoretical and Applied Climatology, 137*, 2803–2827. https://doi.org/10.1007/s00704-018-2746-4

Pettitt, A. N. (1979). A non-parametric to the approach problem. *Applied Statistics, 28*, 126–135.

Pierce, D. W., Barnett, T. P., Santer, B. D., & Gleckler, P. J. (2009). Selecting global climate models for regional climate change studies. *Environmental Sciences, 106*, 8441–8448.

Pomerol, J.-C., & Barba-Romero, S. (2000). *Multicriterion decision in management principles and practice*. Kluwer Academic.

Raju, K. S., & Kumar, D. N. (2014). Ranking of global climate models for India using multicriterion analysis. *Climate Research, 60*(2), 103–117. https://doi.org/10.3354/cr01222

Raju, K. S., & Kumar, D. N. (2015). Ranking general circulation models for India using TOPSIS. *Journal of Water and Climate Change*, *6,* 288–299. https://doi.org/10.2166/wcc.2014.074

Raju, K. S., & Kumar, D. N. (2017). Impact of climate change on water resources with modelling Techniques and Case Studies. *Springer Nature Singapore*. https://doi.org/10.1007/978-981-10-6110-3

Raju, K. S., & Kumar, D. N. (2018). Impact of climate change on water resources in India. *Journal of Environmental Engineering (united States)*. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001394

Raju, K. S., Sonali, P., & Kumar, D. N. (2016). Ranking of CMIP5-based global climate models for India using compromise programming. *Theoretical and Applied Climatology, 128*, 563–574. https://doi.org/10.1007/s00704-015-1721-6

Rummukainen, M., Rockel, B., Bärring, L., Christensen, J. H., & Reckermann, M. (2015). Twenty-first-century challenges in regional climate modeling. *Bulletin of the American Meteorological Society, 96*, 135–138. https://doi.org/10.1175/BAMS-D-14-00214.1

Sabaei, D., Erkoyuncu, J., & Roy, R. (2015). A review of multi-criteria decision making methods for enhanced maintenance delivery. *Procedia CIRP, 37*, 30–35. https://doi.org/10.1016/j.procir.2015.08.086

Sarr, A. B., Camara, M., & Diba, I. (2015). Spatial distribution of CORDEX regional climate models biases over West Africa. *International Journal of Geosciences, 6*, 1018–1031. https://doi.org/10.4236/ijg.2015.69081

Seleshi, Y., & Zanke, U. (2004). Recent changes in rainfall and rainy days in Ethiopia. *International Journal of Climatology, 24*, 973–983. https://doi.org/10.1002/joc.1052

Shi, F., Wang, Z., Qi, L., & Chen, R. (2018). An assessment of GCM performance at a regional scale using a score-based method. *Advances in Meteorology, 2018*, 2018. https://doi.org/10.1155/2018/7641019

Stefanidis, S., Dafis, S., & Stathis, D. (2020). Evaluation of regional climate models (RCMs) performance in simulating seasonal precipitation over mountainous Central Pindus (Greece). *Water, 12*, 1–10. https://doi.org/10.3390/w12102750

Su, F., Duan, X., Chen, D., Hao, Z., & Cuo, L. (2013). Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau. *American Meteorological Society, 26*, 3187–3208. https://doi.org/10.1175/JCLI-D-12-00321.1

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics, 56*, 79–107. https://doi.org/10.1002/2017RG000574

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experment design. *Bulletin of the American Meteorological Society, 93*, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Teegavarapu, R. S. V. (2009). Estimation of missing precipitation records integrating surface interpolation techniques and spatio-temporal association rules. *Journal of Hydroinformatics, 11*, 133–146. https://doi.org/10.2166/hydro.2009.009

Teegavarapu, R. S. V., & Chandramouli, V. (2005). Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *Journal of Hydrology, 312*, 191–206. https://doi.org/10.1016/j.jhydrol.2005.02.015

Tesfaye, S., Taye, G., Birhane, E., & van der Zee, S. E. (2019). Observed and model simulated twenty-first century hydro-climatic change of Northern Ethiopia. *Journal of Hydrology: Regional Studies, 22*, 100595. https://doi.org/10.1016/j.ejrh.2019.100595

Vieux, B. E. (2004). *Distributed hydrologic modelling using GIS, second* (edi). Kluwer Academic Publishers.

von Neumann, J. (1941). Distribution of the ratio of the mean square successive difference to the variance. *The Annals of Mathematical Statistics, 12*, 367–395. https://doi.org/10.1214/aoms/1177731677

Warnatzsch, E. A., & Reay, D. S. (2019). Temperature and precipitation change in Malawi: Evaluation of CORDEX-Africa climate simulations for climate change impact assessments and adaptation planning. *Science of the Total Environment, 654*, 378–392. https://doi.org/10.1016/j.scitotenv.2018.11.098

Wijngaard, J. B., Klein Tank, A. M. G., & Können, G. P. (2003). Himogeneity of the 20th century European daily temprature and precipitation series. *International Journal of Climatology, 23*, 679–692. https://doi.org/10.1002/joc.906

Worku, G., Teferi, E., Bantider, A., Dile, Y. T., & Taye, M. T. (2018). Evaluation of regional climate models performance in simulating rainfall climatology of Jemma sub-basin, Upper Blue Nile Basin, Ethiopia. *Dynamics of Atmospheres and Oceans, 83*, 53–63. https://doi.org/10.1016/j.dynatmoce.2018.06.002

Zhu, Y., Tian, D., & Yan, F. (2020). Effectiveness of entropy weight method in decision-making. *Mathematical Problems in Engineering, 20*, 1–5. https://doi.org/10.1155/2020/3564835Research