Evaluation and selection of CORDEX-SA datasets and bias correction methods for a hydrological impact study in a humid tropical river basin, Kerala

M. S. Saranya* and V. Nair Vinish
Department of Civil Engineering, Rajiv Gandhi Institute of Technology, Velloor P.O., Pampady, Kottayam 686501, Kerala, India
*Corresponding author. E-mail: saranyamscivil@gmail.com
MSS, 0000-0002-2763-1676

ABSTRACT

It is well recognised that the performance of climate model simulations and bias correction methods is region specific, and, therefore, careful validation should always be performed. This study evaluates the performance of five general circulation model–regional climate model (GCM–RCM) combinations selected from CORDEX–SA datasets over a humid tropical river basin in Kerala, India, for climate variables such as precipitation, maximum and minimum temperatures. This involves ranking of the selected climate models based on an EDAS (Evaluation Based on Distance from Average Solution) method and the selection of an appropriate bias correction method for the selected three climate variables. A range of indices are used to evaluate the performance of the bias-corrected climate models to simulate observed climate data. Finally, the hydrological impact of the bias-corrected ranked models is assessed by simulating streamflow over the river basin using individual models and different combinations of models based on rank. According to the findings, hydrological simulation using an average of all GCM–RCM pairs provides the best model output in simulating streamflow, with an NSE value of 0.72. The results confirm the importance of a multimodel ensemble for improving the reliability and minimising the uncertainty of climate predictions for impact studies.

Key words: bias correction, CORDEX–SA datasets, EDAS, rank, SWAT simulation

HIGHLIGHTS

• Multicriteria decision-making approach, EDAS, has been employed to rank CORDEX–SA datasets.
• Bias correction methods evaluated and selected for precipitation and temperature simulation of climate models.
• Hydrological impact of bias-corrected ranked models assessed for the study area using the SWAT model.

INTRODUCTION

The fact that climate change would have a major impact on water resources and hydrology is generally acknowledged. The susceptibility of regional/local hydrology to changing climate conditions results in climate change predictions necessary to evaluate current and future changes in the hydrological cycle (Teutschbein & Seibert 2012). It is a common practice to estimate the present and future impacts of climate change on hydrology by considering climate variables such as temperature and precipitation from climate models. More reliable and accurate predictions of climate variables are becoming increasingly important as the study of current climate–hydrology interactions serves as the basis for future climate change simulations and their impact on water resource management (Smitha et al. 2018).

General circulation models (GCMs) are mathematical models that reflect different earth systems, including the atmosphere, oceans, land surface and sea ice, and provide significant potential for climate change research and variability. They remain relatively coarse in resolution (generally exceed 100 km), however, and are unable to resolve essential features at a finer scale. In terms of spatial and temporal scales, there is a strong mismatch between climate and hydrological modelling, and between GCM accuracy and the hydrological relevance of the variables (Fowler et al. 2007). GCM outputs should, therefore, not be used explicitly to apply hydrological responses to climate change. A downscaling technique is needed to resolve the scale incompatibility issue between the coarse-resolution GCM outputs and the resolution required for regional scale impact assessment (Mendez et al. 2020). Regional climate models (RCMs) are well recognised for providing a better understanding of...
relevant regional/local climate phenomena, their variability and changes. They are widely used to downscale the coarse resolution GCMs and are expected to deliver more precise and reliable climate change forecasts on finer spatial scales for hydrological impact studies (Fowler et al. 2007; Mendez et al. 2020; Pastén-Zapata et al. 2020). Future projections of precipitation from RCMs have been widely used to evaluate the effect of climate change on different systems of water management and hydrological modelling (Mendez et al. 2020).

The RCM outputs possess a certain amount of systematic bias, which is not recommended for direct use in hydrological impact studies. Usually, the climate variable outputs provided by RCMs do not match the statistical properties of the observed time series due to systematic bias. As a result, bias correction (BC) is a necessary step for RCM-simulated meteorological variables before being used to drive a hydrological model. To correct the RCM simulations, many BC techniques have been developed with varying degrees of complexity. These include simple correction approaches to sophisticated distribution mapping (Chen et al. 2013). The BC methods are usually intended for meteorological variables such as precipitation and temperature.

The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (AR5, IPCC) is the latest report from the IPCC, widely referred to for climate change studies. The World Climate Research Program (WCRP) has developed a Coupled Model Intercomparison Project (CMIP) framework for enhancing climate change awareness, which has now been addressed in its fifth phase in the IPCC Fifth Assessment Report (AR5) to communicate the major effect of climate change on natural and human systems across all continents and oceans (Yang et al. 2020). CMIP5 consists of a more varied range of GCMs than CMIP3. As a result, model evaluation from the CMIP5 archives has become more inspiring and insightful all over the world (Raju et al. 2017).

The outcomes of both GCMs and RCMs suffer from the presence of significant uncertainties from various sources, for example, anthropogenic emissions and large-scale and local/regional scale changes that are vulnerable to the model parameterisations and internal dynamics (Diallo et al. 2012). It is impossible to eliminate the uncertainties inherent in climate models; therefore, it is vital to understand and minimise these uncertainties for making a reliable impact assessment of climate change. Multimodel ensemble is widely recommended from the numerous methods that have been generated to reduce uncertainties associated with climate models (Diallo et al. 2012; Yang et al. 2020). According to the study conducted by Diallo et al. (2012), the use of multimodel RCM ensembles for climate change prediction will reduce systematic errors associated with nested RCMs and driving GCMs. Yang et al. (2020) presented a comprehensive index (CI) to develop an optimal multimodel ensemble that can reasonably represent the spatiotemporal variability of both temperature and precipitation and may also be used in hydrological models for climate change assessment.

Impact modellers face a challenging task in selecting appropriate models in terms of reproducibility of observed weather. Several statistical performance measures have been developed by different researchers all over the world to assess the performance of GCMs/RCMs. However, there is no global metric that explicitly identifies a ‘best’ model. Several studies have shown that climate model rankings are responsive to the parameter being considered (Perkins et al. 2007; Gleckler et al. 2008; Raju & Kumar 2014, 2015). Ruan et al. (2018) used an improved score-based approach with multiple parameters to evaluate the performance of CMIP5 GCMs in reproducing observed precipitation at a basin scale. The study conducted by Panjwani et al. (2019) used the Fuzzy Analytic Hierarchy Process and Reliability index to rank global climate models across the Indian region by comparing their temperature and rainfall simulation performance.

The Evaluation based on Distance from Average Solution (EDAS) is a comparatively new Multi-Criteria Decision Making (MCDM) approach, and it uses an average solution to compare alternatives. To the best of our knowledge, no systematic research has been recorded where EDAS has been used for climate model ranking and combined with hydrological impact assessment. Similarly, there is no evidence of a comprehensive study to determine the suitable climate models and BC methods for the study area. As a result, the aim of this research is to rank chosen GCM–RCM pairs downloaded from the CORDEX–SA initiative for climate variables such as precipitation (pr), maximum temperature (Tmax) and minimum temperature (Tmin) independently and evaluate their hydrological impact on the Meenachil river basin, a humid tropical river basin in Kerala, India. Under equal weight (EW) and varying weight (VW) situations, the EDAS method is applied where objective weights are allocated using the CRiteria Importance Through Intercriteria Correlation (CRITIC) method. The hydrological impact of the ranked models of different combinations is evaluated using a semi-distributed hydrological model Soil and Water Assessment Tool (SWAT, Arnold et al. 1998), as several researchers have used it to determine the effect of climate change on hydrological processes.
MATERIALS AND METHODS

Description of the study area

The study area, the Meenachil river basin, geographically lies between 9°25′–9°55′N latitude and 76°30′–77°E longitude (Figure 1). Meenachil river, one of the prominent river basins in Central Kerala, originates at the east of Erattupetta in the district of Kottayam in Kerala at an elevation of 1,097 m above mean sea level (M.S.L.) and flows through Palai, Ettumanoor, Kottayam and Kumarakom before discharging into the Vembanad Lake on the southwest coast of India. It is a natural channel that drains rainwater from the Kottayam district into the sea. The river is 78 km long and has a total drainage area of approximately 1,272 sq. km. The overall elevation of the entire study area ranges from 77 to 1,156 m in the highlands, 8–68 m in the midlands and less than 2 m in the lowlands. The basin receives fairly good rainfall as it falls within the humid tropical domain of climate. The annual catchment precipitation ranges from a minimum of 2,420 mm to a maximum of 4,686 mm and the mean monthly temperature varies from 26.2 to 29.4 °C. The basin experiences both south-west (S-W) monsoon, which sets during June and continues up to the end of August, and north-east (N-E) monsoon, which onsets in October and lasts till November.

Data used

Rather than testing a random set of CORDEX–SA datasets that are produced and distributed under CMIP Phase 5 (CMIP5), the CMIP5 GCMs that perform well across the Indian region for climate variables such as precipitation and temperature were identified from several studies (Sperber et al. 2013; Choudhary et al. 2018; Panjwani et al. 2019, 2020; Prajapat et al. 2020). Sperber et al. (2013) performed a comparative study between CMIP5 and CMIP3 GCM simulations and found that CMIP5 performed better than CMIP3 models in all diagnostics. The CNRM_CM5 and NorESM1_M models are top five performers in representing the June-September climatology over the Asian summer monsoon. Choudhary et al. (2018) evaluated the performance of CORDEX–SA regional climate simulation in representing the seasonal mean summer monsoon precipitation over India. According to their study, seasonal average precipitation is better captured by the GFDL_ESM2M-RegCM4 GCM–RCM combination. The performance evaluation of climate models in simulating average and extreme temperature events is found in Prajapat et al. (2020) and Panjwani et al. (2019, 2020), respectively. Using statistical metrics, Prajapat et al. (2020) found that in the simulation of near-surface mean temperature over India, CanESM and NorESM experiments ranked in the top five performers. According to Panjwani et al. (2019), NorESM1 has performed better in simulating maximum temperature over the Indian region. The model GFDL-ESM2M captures the interannual variability of extreme temperature events better than other models, according to Panjwani et al. (2020).

Therefore, the current study selected five GCM–RCM pairs for precipitation, namely, CMIP5 GCMs such as CNRM-CM5, GFDL-ESM2M and NorESM1-M and RCMs such as RegCM4 and RCA4. A further range of five GCM–RCM pairs containing CMIP5 GCMs such as CanESM2, GFDL-ESM2M, MIROC5 and NorESM1 M and RCMs such as RegCM4 and RCA4 for maximum and minimum temperatures were selected. A brief description of the RCMs and GCMs used in this study is provided in Tables 1 and 2, respectively. The daily precipitation, maximum temperature and minimum temperature data from the selected CMIP5 GCMs, dynamically downscaled to 0.5°/C2 resolution using RegCM4 and RCA4 RCMs available in the framework of CORDEX over South Asia, were received from the Centre for Climate Change Research (CCCR), Indian Institute of Tropical Meteorology (IITM), Pune (http://cccr.tropmet.res.in/home/data_cccrdx.jsp) accessed 3 June 2020). The observed daily precipitation data corresponding to meteorological stations such as Kottayam, Kozha and Kumarakom and daily maximum and minimum temperatures corresponding to the Kottayam station were obtained from the India Meteorological Department (IMD) for the historical/baseline period 1980–2010. Precipitation data corresponding to the Erattupetta station were collected from the Water Resources Department, Government of Kerala, for the same period. Kottayam is the only station for which continuous data for 51 years are available for all three observed climate variables (pr, Tmax and Tmin). It is, therefore, the reference station considered in the present study to rank and correct GCM–RCM outputs. Since the selected climate models have historical data only up to 2005, the remaining data for the 2006–2010 period were chosen from the climate projection based on Representative Concentration Pathway 8.5 (RCP 8.5). It is a ‘baseline’ scenario that does not have any particular climate reduction goal and is consistent with the observed pattern of annual CO₂ emission growth rates from 2005 to 2012 (Khan & Koch 2018).
In order to carry out the hydrological modelling, watershed was delineated using 30 m resolution Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM) V003, available at https://search.earthdata.nasa.gov/search (accessed 18 July 2020). The land-use land-cover map was prepared using 30 m resolution USGS LANDSAT 5 satellite imagery for the year 1992, available at https://earthexplorer.usgs.gov/ (accessed 19 July 2020). The soil map corresponding to the study area was extracted from the digital soil map of the world (version 3.6) available at 1:5,000,000 scale from the Land and Water Development Division, Food and Agricultural Organisation (FAO), Rome.

Figure 1 | Study area map of the Meenachil river basin showing the location of meteorological stations and stream network.
EDAS is a comparatively new multicriteria decision-making method developed by Ghorabaee et al. (2015), where alternatives are compared using average solution (AV). Positive distance from average (PDA) and negative distance from average (NDA) are the two significant measures used in this approach, depending on whether the criteria considered are beneficial or non-beneficial.

Unlike other MCDM methods, the best alternative is dependent on the distance from average solution. The difference between each alternative solution and average solution can be represented using both PDA and NDA. It is desirable to have a higher PDA value and a lower NDA value for the selection of an alternative. In other words, an alternative solution is better than the average solution when it has higher PDA values and/or lower NDA values. The steps involved in the EDAS approach according to Ghorabaee et al. (2015) are as follows:

1. Selection of relevant criteria to illustrate the performance of chosen climate models and the creation of a decision-making matrix (X)

\[
X = [X_{ij}]_{n \times m} = \begin{bmatrix}
X_{11} & X_{12} & \ldots & X_{1m} \\
X_{21} & X_{22} & \ldots & X_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
X_{n1} & X_{n2} & \ldots & X_{nm}
\end{bmatrix}
\]  

(1)

where \(X_{ij}\) corresponds to the performance value of the \(i\)th climate model on the \(j\)th criterion

2. Evaluation of average solution (AV) based on all criteria

\[
AV = [AV_j]_{1 \times m}
\]

(2)

where

\[
AV_j = \frac{\sum_{i=1}^{n} X_{ij}}{n}
\]

(3)
3. Formulation of the PDA and NDA matrices based on whether the criteria chosen are beneficial or non-beneficial.

\[
PDA = \begin{bmatrix} PDA_{ij} \end{bmatrix}_{1 \times m} \tag{4}
\]

\[
NDA = \begin{bmatrix} NDA_{ij} \end{bmatrix}_{1 \times m} \tag{5}
\]

When the \( j \)th criterion is beneficial,

\[
PDA_{ij} = \max(0, \frac{(X_{ij} - AV_j)}{AV_j}) \tag{6}
\]

\[
NDA_{ij} = \max(0, \frac{(AV_j - X_{ij})}{AV_j}) \tag{7}
\]

When the \( j \)th criterion is non-beneficial,

\[
PDA_{ij} = \max(0, \frac{(AV_j - X_{ij})}{AV_j}) \tag{8}
\]

\[
NDA_{ij} = \max(0, \frac{(X_{ij} - AV_j)}{AV_j}) \tag{9}
\]

4. Calculation of the weighted sum of PDA and NDA for all the GCM–RCM combinations

\[
SP_i = \sum_{j=1}^{m} w_j PDA_{ij} \tag{10}
\]

\[
SN_i = \sum_{j=1}^{m} w_j NDA_{ij} \tag{11}
\]

where \( w_j \) is the weight of the \( j \)th criterion, \( SP_i \) is the weighted sum of PDA for the \( i \)th climate model and \( SN_i \) is the weighted sum of NDA for the \( i \)th climate model.

5. Normalisation of \( SP \) and \( SN \) and calculation of appraisal score (AS) for all GCM–RCM combinations.

\[
NSP_i = \frac{SP_i}{\max(SP)} \tag{12}
\]

\[
NSN_i = 1 - \frac{SN_i}{\max(SN)} \tag{13}
\]

\[
AS_i = \frac{1}{2} (NSP_i + NSN_i) \tag{14}
\]

where AS varies from 0 to 1.

Among the alternatives considered, the GCM–RCM combination with the highest AS value is considered to be the most suitable climate model.

### Selection of Performance Criteria

Performance criteria/indicators are needed in order to rank climate models based on how well they can simulate observed data. A single performance indicator that would be uniformly considered the most suitable cannot be specified. Similarly, there are no approved performance indicators that can be used to rank models. Hence, many researchers have suggested various skill scores (SSs) and many have used a variety of performance metrics to measure model simulation skills (Taylor 2001; Ruan et al. 2018; Panjwani et al. 2019). Thus, based on literature review, five simple yet significant statistical performance
measures such as correlation coefficient \((R)\), SS (Taylor 2001), root-mean-square error (RMSE), mean absolute error (MAE) and percentage bias (PBIAS) are chosen for model performance evaluation. The mathematical expression for these indices is given below. The terms \(S_i\) and \(O_i\) are the \(i\)th simulated and observed values, respectively, \(\bar{S}\) and \(\bar{O}\) are the mean of simulated and observed values and \(n\) is the number of observations.

\[
R = \frac{\sum_{i=1}^{n} (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{n} (S_i - \bar{S})^2} - \sum_{i=1}^{n} (O_i - \bar{O})^2}
\]

(15)

\[
PBIAS = \frac{\sum_{i=1}^{n} (O_i - S_i)}{\sum_{i=1}^{n} O_i} \times 100
\]

(16)

\[
MAE = \frac{\sum_{i=1}^{n} |O_i - S_i|}{n}
\]

(17)

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (O_i - S_i)^2}{n}}
\]

(18)

\[
SS = \frac{4(1 + R)^4}{\left(\sigma_n + \frac{1}{\sigma_n}\right)^2 (1 + R_0)^4}
\]

(19)

Based on Equation (19), SS depends on correlation coefficient \((R)\), normalised standard deviation of simulation \(\sigma_n\) and maximum correlation coefficient \(R_0\). Normalised standard deviation is obtained by dividing the standard deviation of the simulated value with the standard deviation of the observed value.

As there is no temporal alignment between the daily climate model simulation and the observed data, the observed daily datasets and climate model simulations of precipitation, maximum temperature and minimum temperature corresponding to the study area are converted to a monthly scale for the period 1980–2010 (372 monthly data comprising of 31 years). These monthly data have been used to evaluate the performance of GCM–RCM combinations based on the performance criteria.

**CRITERIA IMPORTANCE THROUGH INTERCITERA CORRELATION**

The CRITIC method is proposed by Diakoulaki et al. (1995) for determining the objective weights of relative importance for criteria considered under MCDM problems. The method developed is based on the evaluation matrix’s analytical examination to collect all information found in the performance indicators/criteria. The derived weights provide both contrast strength and conflict that are found in the decision problem framework. The standard deviation of each criterion, as well as its correlation with other criteria, is required for the CRITIC method to determine objective weights. Here, we have five GCM–RCM combinations (alternatives) and five evaluation criteria for each climate variable (precipitation, maximum temperature and minimum temperature). The CRITIC method can be explained in the following steps (Diakoulaki et al. 1995):

1. Formulation of the decision matrix \(X\) with \(n\) alternatives (GCM–RCM combinations) and \(m\) evaluation criteria.

\[
X = [X_{ij}]_{n \times m} = \begin{bmatrix}
X_{11} & X_{12} & \cdots & X_{1m} \\
X_{21} & X_{22} & \cdots & X_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
X_{n1} & X_{n2} & \cdots & X_{nm}
\end{bmatrix}
\]

(20)

where \(X_{ij}\) corresponds to the performance value of the \(i\)th GCM–RCM combination on the \(j\)th criterion.
2. Normalisation of the decision matrix based on the equation:

\[ X'_{ij} = \frac{X_{ij} - X_{j}^{\text{worst}}}{X_{j}^{\text{best}} - X_{j}^{\text{worst}}} \]  

(21)

where \( X'_{ij} \) is the normalised performance value of the \( i \)th GCM–RCM combination on the \( j \)th criterion; \( X_{j}^{\text{best}} \) is the best performance value under the \( j \)th criterion and \( X_{j}^{\text{worst}} \) is the worst performance value under the \( j \)th criterion, derived directly from the existing alternatives. In this manner, the initial decision matrix is converted into a matrix of relative score varying from 0 to 1.

3. Estimation of the standard deviation (\( \sigma_j \)) of the performance values under each criterion and the correlation between each criterion for the process of evaluation of criteria weights. Objective weights are obtained by normalising the amount of information delivered by each criterion to unity.

\[ C_j = \sigma_j \sum_{k=1}^{m} (1 - r_{jk}) \]  

(22)

\[ w_j = \frac{C_j}{\sum_{k=1}^{m} C_k} \]  

(23)

where \( C_j \) is the amount of information contained in the \( j \)th criterion, \( w_j \) is the objective weight of the \( j \)th criterion and \( r_{jk} \) is the correlation coefficient between the \( j \)th and the \( k \)th criteria. The higher \( C_j \) value means that the corresponding criterion transmits a greater amount of information and its relative significance is higher for the decision-making process.

**BIAS CORRECTION**

As the direct use of RCM output is not recommended due to systematic bias, it is important to correct the biases in the precipitation and temperature predictions from the chosen GCM–RCM combinations. Out of the numerous BC methods available to remove systematic bias, the present study evaluates the performance of four BC methods for precipitation and three BC methods for temperature. The selected BC methods are simple to complex in nature and include a significant type of existing BC methods. A brief description of the chosen BC methods is given in Table 3. The BC methods were executed using the CMhyd tool (Rathjens et al. 2016). In order to evaluate the ability of the bias-corrected GCM–RCM pairs to simulate observed precipitation and temperature, a number of time-series-based and distribution-based indices were chosen.

The BC of the climate model-simulated time-series data for the baseline period (1980–2010) was performed using the observed data of the same duration.

| BC method | Correcting variable | Description |
|-----------|---------------------|-------------|
| Linear scaling LS | Precipitation and temperature | It applies monthly correction values to exactly match the monthly average of corrected RCM simulations (Lenderink et al. 2007). |
| LOCI | Precipitation | Using monthly correction, it corrects the mean, wet-day frequencies and wet-day intensities of precipitation (Schmidli et al. 2006). |
| MPT | Precipitation | A hybrid method combining both the LOCI and the power transformation (PT) method. It is a non-linear correction that adjusts both mean and variance of RCM simulation using an exponential form (Leander & Buishand 2007; Smitha et al. 2018). |
| VS | Temperature | It adjusts both the mean and variance of temperature similar to the method of power transformation (Chen et al. 2011). |
| DM | Precipitation and temperature | It corrects the mean, standard deviation and quantiles of raw RCM data by matching the distribution function with the data observed. Gamma distribution and normal distribution are best suited for precipitation and temperature, respectively (Teutschbein & Seibert 2012). |
The reliability of the BC methods to represent observed climate variables was evaluated using the indices mentioned in Table 4 (Sillmann et al. 2013; Pastén-Zapata et al. 2020) for the baseline period. Based on their ability to simulate all indices relative to other BC methods and uncorrected GCM–RCM data, the BC methods and raw GCM–RCM were then ranked for each GCM–RCM pair. The evaluation of the BC methods was carried out separately for precipitation, maximum temperature and minimum temperature. For precipitation with four BC methods, ranks were assigned in such a way that a value of 1 for the best simulation and a value of 5 for the worst simulation correspond to each performance indicator. Similarly, a rank of 1 for the best simulation and a rank of 4 for the worst simulation were given for maximum/minimum temperature with three BC methods. Based on the rank assigned for each predictor, the average score corresponding to each BC method and the raw GCM–RCM data were calculated. The BC method with the lowest average score was chosen for the GCM–RCM pair under consideration.

HYDROLOGICAL SIMULATION

The hydrological model SWAT was used in this study to evaluate the impact of ranked and bias-corrected GCM–RCM combinations in hydrological processes. The amount of streamflow that the SWAT model simulates depends on the accuracy of the input climate variables, such as precipitation and temperature. It is, therefore, necessary to carefully select climate models for hydrological simulation.

In order to understand the effect of bias-corrected ranked models on the hydrological response of the river basin, streamflow was simulated using different combinations of the ranked models. At first, streamflow was simulated using the observed precipitation, maximum temperature and minimum temperature. Observed climate data were then replaced by corresponding simulations of the GCM–RCM pairs ranked 1 and then the streamflow was simulated by keeping all other parameters the same between simulations. The simulation was then repeated with the average simulation of the top 2 models, top 3 models, top 4 models and finally with the average of all the five GCM–RCM pairs. To understand the hydrological effect of uncorrected GCM–RCM pairs, simulation was performed with the ensemble average of uncorrected GCM–RCM pairs. Also, each bias-corrected model was independently used for streamflow simulation. As the aim of the current study was to understand the hydrological effect of the selected climate models, the parameters of the SWAT hydrological model were not calibrated. All streamflow simulations using climate model data were compared with streamflow simulations using the observed data, evaluating whether the climate model predictions were able to simulate streamflows comparable to those simulated using the observed data.

RESULTS AND DISCUSSION

The selection of appropriate models in terms of reproducibility of observed weather is an important task in climate change impact studies. As a result, it is vital to assess the relative performance of climate models using a variety of criteria that are appropriate for impact studies.

Evaluation of the GCM–RCM simulation skill

The daily data collected from the IMD corresponding to the Kottayam station were converted to a monthly scale (31 × 12 = 372 monthly data) for variables such as precipitation (pr), maximum temperature (Tmax) and minimum temperature (Tmin). The outputs of the selected GCM–RCM pairs on a monthly scale for the three variables were then compared with the corresponding data observed for the baseline period 1980–2010. GCM–RCM pairs comprising a set of five models for precipitation (see Table 1) and another set of five models common for maximum and minimum temperatures (see Table 2) were chosen. The performance of the selected GCM–RCM pairs over the Meenachil river basin was assessed using performance indicators R, SS, RMSE, PBIAS and MAE.

The correlation coefficient reflects the strength of the relationship between the observed and the simulated climate variable. An R value of 1 is indicative of good model efficiency. SS is a least complex performance measure developed by Taylor (2001) to evaluate model performance. The value of the SS ranges from 0 to 1; the greater the SS value, the better a model’s ability to simulate observed data. The PBIAS, RMSE and MAE are error measures, and, hence, the least value is preferred. Table 5 displays the values of selected performance indicators PBIAS, R, SS, MAE and RMSE calculated to evaluate the performance of selected GCM–RCM pairs in simulating precipitation (pr), maximum temperature (Tmax) and minimum temperature (Tmin). The table shows that GFDL-RCA4 has the best value for each performance indicator when simulating precipitation. Similarly, GFDL-RegCM4 produces better results when simulating Tmax, and MIROC5-RCA4 produces better results when
simulating Tmin. The aim of the present study was to look into all five performance indicators at the same time in order to assess their applicability while ranking the GCM–RCM pairs.

Despite the fact that the same set of models was used for both Tmax and Tmin, the values of performance indicators were not the same for the two variables. The R and SS values revealed that the selected models performed better in simulating Tmax than in simulating Tmin. The MIROC5-RCA4 model predicted higher R values for both Tmax and Tmin. Although the SS value for CanESM2-RegCM4 was relatively good for Tmax, the SS values for MIROC5-RCA4 and NorESM1-RCA4 were better for minimum temperature. The above results indicate that a model that performs well in maximum temperature simulation does not guarantee effective minimum temperature simulation skills, and vice versa. The findings also revealed

Table 4 | Brief description of the indices used for evaluating BC methods for precipitation, maximum temperature and minimum temperature (Sillmann et al. 2013; Pastén-Zapata et al. 2020)

| Index name | Index definition | Performance measure |
|------------|------------------|---------------------|
| Precipitation | | |
| 95th percentile | 95th percentile of daily precipitation | Bias (mm/day) |
| 90th percentile | 90th percentile of daily precipitation | Bias (mm/day) |
| 50th percentile | 50th percentile of daily precipitation | Bias (mm/day) |
| 25th percentile | 25th percentile of daily precipitation | Bias (mm/day) |
| Correlation coefficient (R) | Pearson's correlation coefficient between monthly observed and simulated precipitation | Index |
| AMP | Annual accumulated precipitation over the chosen period | Mean percentage error |
| MMP | Accumulated precipitation for a given month of the year over the chosen period | Mean percentage error |
| Maximum 1-day precipitation (RX1 Day) | Monthly maximum 1-day precipitation for a given month over the chosen period | Mean percentage error |
| Number of heavy precipitation days (R10) | Average number of days with daily precipitation amount of ≥10 mm within a year | Bias (days) |
| Number of very heavy precipitation days (R20) | Average number of days with daily precipitation amount of ≥20 mm within a year | Bias (days) |
| Simple Daily Intensity Index (SDII) | Ratio of sum of daily precipitation amount on wet days (≥1 mm) to the number of wet days over the chosen period | Index |
| Maximum and minimum temperatures | | |
| 99th percentile | 99th percentile of the daily maximum/minimum temperature | Bias (°C/day) |
| 1st percentile | 1st percentile of the daily maximum/minimum temperature | Bias (°C/day) |
| Correlation coefficient (R) | Pearson's correlation coefficient between monthly observed and simulated maximum/minimum temperature | Index |
| Max. TX (TXx) | Maximum value of daily maximum temperature for a given month over the chosen period | Mean percentage error |
| Min. TX (TXn) | Minimum value of daily maximum temperature for a given month over the chosen period | Mean percentage error |
| Max. TN (TNx) | Maximum value of daily minimum temperature for a given month over the chosen period | Mean percentage error |
| Min. TN (TNn) | Minimum value of daily minimum temperature for a given month over the chosen period | Mean percentage error |
| AMT | Annual average daily maximum/minimum temperature over the chosen period | Mean percentage error |
| MMT | Monthly average daily maximum/minimum temperature over the chosen period | Mean percentage error |
that a GCM-RCM pair can perform well for one criterion, while performing poorly for another. As a result, rather than employing a single criterion method, it is essential to consider the GCM-RCMs extensively using a multiple criteria method.

Ranking of RCMs using EDAS under EW and VW scenarios

Based on the performance indicators determined above, RCMs ranked under EW and VW scenarios were chosen independently for precipitation, maximum and minimum temperatures. As the indicators selected were 5 in number, an EW of 0.2 was assigned for each indicator under an EW scenario, whereas objective weights were calculated using the CRITIC method for each criterion to carry out ranking under a VW scenario. A normalised decision matrix consisting of five GCM–RCM combinations and five performance criteria was formulated to achieve objective weights without the involvement of the decision-maker (see Equation (21)). Since the normalisation process involves the best and worst values corresponding to each criterion, the highest value was considered to be the best value for beneficial criteria such as RR and SS, while the lowest value was considered to be the best value for non-beneficial criteria such as PBIAS, MAE and RMSE.

The CRITIC method requires the standard deviation ($\sigma_j$) and the linear correlation ($r_{jk}$) of the normalised performance values to determine the criteria weights, where $\sigma_j$ corresponds to the contrast intensity of the criteria and $r_{jk}$ is a measure of the conflict created by each criterion in relation to the rest of the criteria. A symmetric matrix with a 5×5 dimension (since the number of performance indicators chosen was equal to 5) was created with generic element $r_{jk}$ in such a way that $r_{jk}$ corresponds to the linear correlation between the $j$th and the $k$th criterion. Using the above matrix and the standard deviation of the performance values for each criterion, the amount of information transmitted by each criterion ($C_j$) was then calculated using Equation (22). The $C_j$ values for performance indicators such as PBIAS, R, SS, MAE and RMSE were found to be 0.18, 0.23, 0.30, 0.29 and 0.21 for precipitation, respectively. Likewise, the $C_j$ values calculated were 0.83, 2.90, 1.08, 0.84 and 0.87 for maximum temperature and 0.65, 0.80, 0.99, 1.54 and 0.63 for minimum temperature. After normalising the $C_j$ values, the objective weights of PBIAS, R, SS, MAE and RMSE for precipitation were found to be 0.14, 0.19, 0.25, 0.24 and 0.17, respectively. Accordingly, objective weights were found to be 0.13, 0.44, 0.17, 0.15 and 0.13 for the maximum temperature and 0.14, 0.17, 0.22, 0.33 and 0.14 for the minimum temperature. The SS has the highest weight in evaluating

| GCM-RCM combination | PBIAS (%) | $R$ | SS | MAE | RMSE |
|----------------------|-----------|----|----|-----|------|
| **Precipitation**    |           |    |    |     |      |
| CNRM-RCA4            | 22.96     | 0.61 | 0.35 | 4.11 | 6.31 |
| CNRM-RegCM4          | 35.12     | 0.59 | 0.25 | 4.61 | 6.77 |
| GFDL-RCA4            | 1.47      | 0.68 | 0.49 | 3.95 | 5.88 |
| GFDL-RegCM4          | 41.11     | 0.62 | 0.37 | 4.58 | 6.81 |
| NorESM1-RCA4         | 50.33     | 0.59 | 0.21 | 4.58 | 7.40 |
| **Maximum temperature** |         |    |    |     |      |
| CanESM2-RegCM4       | 7.03      | 0.64 | 0.43 | 2.29 | 2.6  |
| GFDL-RCA4            | 12.30     | 0.78 | 0.35 | 3.93 | 4.43 |
| GFDL-RegCM4          | 5.46      | 0.69 | 0.41 | 1.95 | 2.34 |
| MIROC5-RCA4          | 10.75     | 0.82 | 0.37 | 3.46 | 4.00 |
| NorESM1-RCA4         | 7.82      | 0.75 | 0.36 | 2.62 | 3.17 |
| **Minimum temperature** |         |    |    |     |      |
| CanESM2-RegCM4       | 12.42     | 0.44 | 0.23 | 2.93 | 3.22 |
| GFDL-RCA4            | 13.34     | 0.53 | 0.28 | 3.15 | 3.40 |
| GFDL-RegCM4          | 14.25     | 0.79 | 0.22 | 3.35 | 3.58 |
| MIROC5-RCA4          | 12.11     | 0.56 | 0.29 | 2.85 | 3.13 |
| NorESM1-RCA4         | 12.72     | 0.54 | 0.29 | 4.11 | 3.25 |
precipitation. The highest weight was assigned to $R$ in the case of $T_{\text{max}}$, while the highest weight was assigned to $\text{MAE}$ in the case of $T_{\text{min}}$. For all three cases, the CRITIC method assigned the least weight to the performance indicator $\text{PBIAS}$. Furthermore, the least weight was assigned to $\text{MAE}$ and $\text{RMSE}$ for maximum temperature and $\text{RMSE}$ for minimum temperature. The CRITIC method yielded an objective weight that combined the contrast intensity of each criterion as well as the conflict between the criteria.

Based on the performance indicators calculated, the average solution corresponding to each criterion was identified. Moreover, for the three climate variables, the weighted sum of both positive and negative distances from the average solution ($\text{PDA}$ and $\text{NDA}$) was determined using an EW of 0.2 and the respective unequal weights found by the CRITIC method. Table 6 displays the normalised weighted sum of $\text{PDA}$ and $\text{NDA}$ ($\text{NSP}_i$ and $\text{NSN}_i$) and the average score ($\text{AS}_i$) used to rank each GCM–RCM pair for precipitation, maximum and minimum temperatures.

Models such as GFDL-RCA4, GFDL-RegCM4 and MIROC5-RCA4 ranked top in simulating precipitation, maximum temperature and minimum temperature, respectively. The selection was made based on the highest average score, which was suitable for selecting climate models using the EDAS method. The same ranking pattern for all three climate variables was observed in the EW and VW scenarios. Although the GCM–RCM pair GFDL-RCA4 was the best model for precipitation simulation, it was the least preferred model based on the maximum temperature simulation. However, it fell under the top three models for minimum temperature simulation. Similarly, the best model (GFDL-RegCM4) for maximum temperature simulation was ranked as the last model for minimum temperature simulation. Nevertheless, the GCM–RCM pair NorESM1-RCA4 had good simulation skills for both maximum and minimum temperatures, as it was one of the top three models. In general, a model with better simulation skills for a chosen climate variable does not guarantee similar simulation skills for other variables. Therefore, in climate change impact studies, a careful selection of climate models is vital. With the addition of different GCM–RCM pairs, performance criteria, climate variables and changes in criteria weights, the resulting ranking patterns may change (Raju & Kumar 2014).

Table 6 | Calculated values of $\text{NSP}_i$, $\text{NSN}_i$, $\text{AS}_i$ and rank assigned for each GCM–RCM pair under EW and VW scenarios for precipitation, maximum temperature and minimum temperature

| GCM–RCM combination | EW scenario | VW scenario |
|----------------------|-------------|-------------|
|                      | NSP$_i$, NSN$_i$, AS$_i$, Rank | NSP$_i$, NSN$_i$, AS$_i$, Rank |
| Precipitation | | |
| CNRM-RCA4 | 0.227, 0.990, 0.608, 2 | 0.217, 0.989, 0.603, 2 |
| CNRM-RegCM4 | 0.000, 0.569, 0.284, 4 | 0.000, 0.510, 0.255, 4 |
| GFDL-RCA4 | 1.000, 1.000, 1.000, 1 | 1.000, 1.000, 1.000, 1 |
| GFDL-RegCM4 | 0.067, 0.649, 0.358, 3 | 0.091, 0.700, 0.395, 3 |
| NorESM1-RCA4 | 0.000, 0.000, 0.000, 5 | 0.000, 0.000, 0.000, 5 |
| Maximum temperature | | |
| CanESM2–RegCM4 | 0.680, 0.894, 0.787, 2 | 0.699, 0.643, 0.671, 2 |
| GFDL–RCA4 | 0.057, 0.000, 0.028, 5 | 0.191, 0.000, 0.095, 5 |
| GFDL–RegCM4 | 1.000, 0.952, 0.976, 1 | 1.000, 0.837, 0.918, 1 |
| MIROC5–RCA4 | 0.108, 0.425, 0.266, 4 | 0.363, 0.427, 0.395, 4 |
| NorESM1–RCA4 | 0.229, 0.951, 0.590, 3 | 0.261, 0.939, 0.600, 3 |
| Minimum temperature | | |
| CanESM2–RegCM4 | 0.344, 0.577, 0.461, 4 | 0.381, 0.174, 0.277, 4 |
| GFDL–RCA4 | 0.371, 0.904, 0.637, 2 | 0.384, 0.926, 0.655, 2 |
| GFDL–RegCM4 | 0.000, 0.000, 0.000, 5 | 0.000, 0.000, 0.000, 5 |
| MIROC5–RCA4 | 1.000, 1.000, 1.000, 1 | 1.000, 1.000, 1.000, 1 |
| NorESM1–RCA4 | 0.443, 0.553, 0.498, 3 | 0.406, 0.542, 0.474, 3 |
Uncorrected GCM–RCM pair simulations

The ability to simulate the precipitation, maximum temperature and minimum temperature of uncorrected GCM–RCM pairs was evaluated. The monthly average precipitation over the Meenachil river basin (Figure 2) reveals that significant bias exists between the uncorrected GCM–RCM output and the observed precipitation. Among the five uncorrected GCM–RCM pairs, GFDL-RCA4 has the least bias with respect to the observed precipitation. All GCM–RCM pairs except GFDL-RCA4 underestimate the S-W monsoon and the N-E monsoon; GFDL-RCA4 overestimates the S-W monsoon and underestimates the N-E monsoon. The NorESM1-RCA4 and GFDL-RegCM4 models underestimate precipitation during all seasons. Since the peak of the S-W monsoon for observed data is in June, the peak simulated by GFDL-RCA4, GFDL-RegCM4 and NorESM1-RCA4 is in July. For CNRM-RCA4 and CNRM-RegCM4, the peak of the S-W monsoon is in August. Furthermore, an analysis of the average maximum and average minimum temperature trends over the Meenachil river basin (Figure 3) shows that direct application of GCM–RCM outputs in local climate change studies is not recommended.

Although the average maximum temperature during all seasons is underestimated by all the uncorrected GCM–RCM pairs, there is a major bias during the period from June to August. However, the underestimation of the average minimum temperature during the period of June–August is minimal. The average maximum temperature during the November–February period is well simulated by GFDL-RegCM4 with a bias of less than 2%. In addition, the peak of the maximum temperature during the month of March is well simulated by NorESM1-RCA4 and MIROC5-RCA4. Although some of the GCM–RCM pairs are better at simulating maximum temperature, their minimum temperature simulation is poor.

Furthermore, the raw GCM–RCM outputs were compared with the measured climate data in terms of probability of exceedance. Figure 4 depicts the probability of exceedance for the chosen climate models when simulating (a) precipitation, (b) maximum temperature and (c) minimum temperature. When comparing the exceedance probability of the GCM–RCM pairs with the observed precipitation, it was found that GFDL-RCA4 adequately represents the observed precipitation, whereas NorESM1-RCA4 exhibits a significant mismatch. This means that the GFDL-RCA4 precipitation simulation is more reliable. For precipitation with a low probability of exceedance, such as an extremely high precipitation, the differences between observations and raw GCM–RCM pair precipitation outputs are larger. Furthermore, for precipitation values of low magnitude, relatively small variations are observed. When climate models were evaluated for maximum temperature simulation, the upper extreme of maximum temperature was satisfactorily simulated, while the lower extreme was greatly underestimated by all GCM–RCM pairs. Climate models, on the other hand, underestimate both the lower and the upper extremes of minimum temperature. The models’ mismatch in representing maximum temperature distributions increases in the direction of high probability of exceedance (i.e. towards the lower extreme of maximum temperature), with GFDL-RegCM4 exhibiting the least mismatch and GFDL-RCA4 exhibiting the most mismatch. When compared with the other
models, the CanESM2-RegCM4 model has a higher bias in simulating maximum temperature at a low probability of exceedance and a lower bias at a high probability of exceedance. The mismatch between the observed climate data and the raw GCM-RCM pair output is generally proportional to the bias in climate model output. The distribution of precipitation, maximum temperature and minimum temperature is better represented by GFDL-RCA4, GFDL-RegCM4 and MIROC5-RCA4, respectively, among the selected GCM–RCM pairs. These findings also support the EDAS method–based ranking of the models.

**Evaluation of BC methods**

As the realistic estimates of climate variables are essential for hydrological impact study, the BC of climate model simulations was executed for the baseline period 1980–2010 at a daily timescale. The BC of the GCM–RCM pairs of precipitation using linear scaling (LS), local intensity scaling (LOCI), modified power transformation (MPT) and distribution mapping (DM) (see Table 3) considerably improved their simulation skill. Based on 11 performance indicators mentioned in Table 4, BC methods and uncorrected model were ranked for each GCM–RCM pair (Table 7). The BC method with the lowest average score for each GCM–RCM pair was chosen for the respective model. In other words, for hydrological simulation in SWAT, the output of each climate model bias corrected by the respective top-ranked BC method was considered. Thus, simulations of the CNRM-RCA4 model bias corrected with the MPT method and those of all other models done by the LOCI method were chosen for the hydrological impact study. A similar approach was adopted for the GCM–RCM pairs of maximum and minimum temperatures with the BC methods, LS, VS and DM. Tables 8 and 9, respectively, display the ranking of the BC methods based on seven performance metrics (see Table 4) for maximum and minimum temperatures. For further impact analysis, maximum temperature outputs of CanESM2-RegCM4, GFDL-RegCM4 and MIROC5-RCA4 bias corrected by the VS method, NorESM1-RCA4 bias corrected by VS or DM and GFDL-RCA4 bias corrected by the DM method were selected (see Table 8). Similarly, the minimum temperature simulations of CanESM2-RegCM4 and GFDL-RegCM4 bias corrected by the DM method, GFDL-RCA4 and NorESM1-RCA4 bias corrected by the DM or VS method and MIROC5-RCA4 bias corrected by the LS or VS method were considered (see Table 8).

According to the results, all of the BC methods were able to improve the GCM–RCM outputs to some extent. Their corrected statistics, on the other hand, are not the same and depend on the correction method used. When compared with the precipitation BC methods, the variability in the performance of the BC methods is much lower for temperature.

**Performance of precipitation BC methods**

Figure 5 shows the amount of bias associated with upper and lower precipitation extremes for all models regardless of whether it is positive or negative bias. We analysed the biases in the upper and lower precipitation extremes using percentiles and found that all of the BC methods performed flawlessly (zero bias) in simulating the 25th percentile. While the LS method slightly overestimated the 50th percentile (<5 mm/day) for all models, all other methods showed no bias in the 50th percentile.
All uncorrected GCM–RCM pairs underestimated the 95th and 90th percentiles, whereas all BC methods reduced the amount of biases, with the DM method significantly reducing them. R increased for all models while using the LS and LOCI methods (Figure 6(c)). The SDII ratio underestimated by the uncorrected GCM–RCM pairs exhibited satisfactory results for the same when bias corrected by all the methods except the LS method. The SDII ratio for the observed data was 21.33. Figure 6(d) shows the SDII ratio of the uncorrected and bias-corrected GCM–RCM pairs. Most BC methods underestimate maximum 1-day precipitation (RX1 day) for all models (Figure 7(a)), while the LS method for CNRM-RCA4 and the LOCI method for GFDL-RegCM4 simulate it with the least PBIAS.

The performance of the MPT and DM methods while simulating R10 was found to be better (a bias of 3–12 days) than that of the LS and LOCI methods, except for the GFDL-RegCM4 model, where all BC methods performed equally well (Figure 7(b)). However, the LOCI, MPT and DM methods showed satisfactory results for the simulation of R20 for the GCM–RCM pairs with bias between 0 and 9 days (Figure 7(c)). For the GCM–RCM pairs, the LS, LOCI and MPT methods
| GCM–RCMs and BC methods | 95th percentile | 90th percentile | 50th percentile | 25th percentile | R | AMP | MMP | RX1 | R10 | R20 | SDII | Average score | Rank |
|--------------------------|----------------|----------------|----------------|----------------|---|-----|-----|-----|-----|-----|-----|------|----------------|------|
| CNRM–RCA4 Uncorrected    | 5              | 5              | 4              | 1              | 5 | 5   | 4   | 2   | 1   | 5   | 5   | 3.82 | 5               |      |
| LS                       | 4              | 4              | 5              | 1              | 1 | 3   | 1   | 1   | 4   | 4   | 4   | 2.91 | 4               |      |
| LOCI                     | 5              | 3              | 1              | 1              | 2 | 2   | 1   | 3   | 5   | 1   | 3   | 2.27 | 2               |      |
| MPT                      | 2              | 2              | 1              | 1              | 3 | 1   | 1   | 5   | 3   | 2   | 2   | 2.09 | 1               |      |
| DM                       | 1              | 1              | 1              | 1              | 4 | 4   | 5   | 4   | 2   | 3   | 1   | 2.45 | 3               |      |
| CNRM–RegCM4 Uncorrected  | 5              | 5              | 5              | 1              | 5 | 5   | 4   | 2   | 5   | 5   | 5   | 4.27 | 5               |      |
| LS                       | 4              | 4              | 1              | 1              | 1 | 3   | 3   | 1   | 3   | 3   | 4   | 2.55 | 3               |      |
| LOCI                     | 3              | 1              | 1              | 1              | 2 | 2   | 1   | 3   | 4   | 1   | 2   | 1.91 | 1               |      |
| MPT                      | 2              | 3              | 1              | 1              | 3 | 1   | 1   | 5   | 2   | 2   | 3   | 2.18 | 2               |      |
| DM                       | 1              | 2              | 1              | 1              | 4 | 4   | 5   | 4   | 1   | 4   | 1   | 2.55 | 3               |      |
| GFDL–RCA4 Uncorrected    | 5              | 5              | 5              | 1              | 2 | 1   | 4   | 3   | 4   | 5   | 5   | 3.64 | 5               |      |
| LS                       | 4              | 4              | 4              | 1              | 1 | 4   | 1   | 4   | 3   | 4   | 4   | 3.09 | 4               |      |
| LOCI                     | 3              | 2              | 1              | 1              | 3 | 2   | 1   | 1   | 5   | 1   | 2   | 2.00 | 1               |      |
| MPT                      | 2              | 3              | 1              | 1              | 4 | 3   | 1   | 5   | 1   | 3   | 3   | 2.45 | 3               |      |
| DM                       | 1              | 1              | 1              | 1              | 5 | 5   | 5   | 2   | 2   | 1   | 2.36 | 2               |      |
| GFDL–RegCM4 Uncorrected  | 5              | 5              | 4              | 1              | 5 | 5   | 5   | 1   | 5   | 5   | 5   | 4.18 | 5               |      |
| LS                       | 4              | 4              | 5              | 1              | 2 | 2   | 1   | 5   | 2   | 4   | 4   | 3.09 | 4               |      |
| LOCI                     | 1              | 2              | 1              | 1              | 3 | 1   | 1   | 3   | 1   | 2   | 1   | 1.55 | 1               |      |
| MPT                      | 3              | 1              | 1              | 1              | 1 | 4   | 1   | 4   | 3   | 1   | 3   | 2.09 | 2               |      |
| DM                       | 2              | 3              | 1              | 1              | 3 | 3   | 4   | 2   | 4   | 3   | 2   | 2.55 | 3               |      |
| NorESM1–RCA4 Uncorrected | 5              | 5              | 4              | 1              | 4 | 5   | 5   | 4   | 5   | 5   | 5   | 4.36 | 5               |      |
| LS                       | 4              | 4              | 5              | 1              | 1 | 3   | 1   | 2   | 3   | 4   | 4   | 2.91 | 4               |      |
| LOCI                     | 3              | 2              | 1              | 1              | 2 | 2   | 1   | 1   | 1   | 1   | 1   | 1.45 | 1               |      |
| MPT                      | 2              | 3              | 1              | 1              | 3 | 1   | 1   | 5   | 2   | 3   | 3   | 2.27 | 2               |      |
| DM                       | 1              | 1              | 1              | 1              | 5 | 4   | 4   | 3   | 1   | 2   | 2   | 2.27 | 2               |      |
simulated monthly mean precipitation (MMP) with zero bias, while the DM method simulated MMP with a 31–36 PBIAS. Similarly, all BC methods provided acceptable results for annual mean precipitation (AMP), with the PBIAS varying from 1 to 4%.

Table 8 | Rank assigned for BC methods and uncorrected GCM–RCM pairs of maximum temperature based on performance indices

| GCM–RCMs and BC methods | 99th percentile | 1st percentile | R | TXx | TXn | AMT | MMT | Average score | Rank |
|-------------------------|----------------|----------------|---|-----|-----|-----|-----|---------------|------|
| CanESM2–RegCM4          | Uncorrected    | 2              | 4 | 4   | 3   | 3   | 4   | 4             | 3.43 | 4    |
|                         | LS             | 4              | 4 | 1   | 3   | 4   | 1   | 1             | 2.14 | 2    |
|                         | VS             | 1              | 2 | 1   | 1   | 2   | 1   | 1             | 1.29 | 1    |
|                         | DM             | 3              | 3 | 1   | 2   | 4   | 1   | 1             | 2.14 | 2    |
| GFDL–RCA4               | Uncorrected    | 3              | 4 | 1   | 4   | 4   | 4   | 4             | 3.43 | 4    |
|                         | LS             | 4              | 1 | 4   | 3   | 1   | 1   | 1             | 2.14 | 3    |
|                         | VS             | 2              | 2 | 3   | 2   | 2   | 1   | 1             | 1.86 | 2    |
|                         | DM             | 1              | 3 | 2   | 1   | 3   | 1   | 1             | 1.71 | 1    |
| GFDL–RegCM4             | Uncorrected    | 3              | 4 | 4   | 4   | 4   | 4   | 4             | 3.43 | 4    |
|                         | LS             | 4              | 1 | 3   | 3   | 1   | 1   | 1             | 2.00 | 3    |
|                         | VS             | 2              | 2 | 1   | 1   | 2   | 1   | 1             | 1.43 | 1    |
|                         | DM             | 1              | 3 | 2   | 2   | 3   | 1   | 1             | 1.86 | 2    |
| MIROC5–RCA4             | Uncorrected    | 2              | 4 | 1   | 4   | 4   | 4   | 4             | 3.29 | 4    |
|                         | LS             | 4              | 1 | 4   | 3   | 1   | 1   | 1             | 2.14 | 3    |
|                         | VS             | 3              | 2 | 2   | 2   | 2   | 1   | 1             | 1.86 | 1    |
|                         | DM             | 1              | 3 | 3   | 1   | 3   | 3   | 1             | 2.14 | 2    |
| NorESM1–RCA4            | Uncorrected    | 3              | 4 | 1   | 4   | 3   | 4   | 4             | 3.29 | 4    |
|                         | LS             | 4              | 1 | 4   | 2   | 1   | 1   | 1             | 2.00 | 3    |
|                         | VS             | 2              | 2 | 3   | 3   | 2   | 1   | 1             | 1.86 | 1    |
|                         | DM             | 1              | 3 | 1   | 1   | 4   | 1   | 1             | 1.57 | 3    |

Table 9 | Rank assigned for BC methods and uncorrected GCM–RCM pairs of minimum temperature based on performance indices

| GCM–RCMs and BC methods | 99th percentile | 1st percentile | R | TNx | TNn | AMT | MMT | Average score | Rank |
|-------------------------|----------------|----------------|---|-----|-----|-----|-----|---------------|------|
| CanESM2–RegCM4          | Uncorrected    | 4              | 4 | 4   | 4   | 4   | 4   | 4             | 4.00 | 4    |
|                         | LS             | 3              | 3 | 3   | 3   | 3   | 1   | 3             | 2.71 | 3    |
|                         | VS             | 2              | 2 | 1   | 2   | 2   | 2   | 1             | 1.71 | 2    |
|                         | DM             | 1              | 1 | 1   | 1   | 2   | 1   | 1             | 1.14 | 1    |
| GFDL–RCA4               | Uncorrected    | 4              | 4 | 4   | 4   | 4   | 4   | 4             | 4.00 | 4    |
|                         | LS             | 2              | 3 | 3   | 3   | 1   | 1   | 1             | 2.00 | 3    |
|                         | VS             | 1              | 2 | 1   | 3   | 2   | 1   | 1             | 1.57 | 1    |
|                         | DM             | 3              | 1 | 2   | 2   | 1   | 1   | 1             | 1.57 | 1    |
| GFDL–RegCM4             | Uncorrected    | 4              | 4 | 4   | 4   | 4   | 4   | 4             | 4.00 | 4    |
|                         | LS             | 3              | 3 | 3   | 2   | 3   | 1   | 1             | 2.29 | 3    |
|                         | VS             | 1              | 2 | 1   | 3   | 2   | 1   | 1             | 1.57 | 2    |
|                         | DM             | 1              | 1 | 1   | 1   | 1   | 1   | 1             | 1.00 | 1    |
| MIROC5–RCA4             | Uncorrected    | 4              | 4 | 4   | 4   | 4   | 4   | 4             | 4.00 | 4    |
|                         | LS             | 1              | 1 | 3   | 3   | 1   | 1   | 1             | 1.57 | 1    |
|                         | VS             | 2              | 2 | 1   | 2   | 2   | 1   | 1             | 1.57 | 1    |
|                         | DM             | 3              | 3 | 2   | 3   | 1   | 1   | 1             | 2.00 | 3    |
| NorESM1–RCA4            | Uncorrected    | 4              | 4 | 4   | 4   | 4   | 4   | 4             | 4.00 | 4    |
|                         | LS             | 1              | 3 | 3   | 1   | 3   | 1   | 1             | 1.86 | 3    |
|                         | VS             | 2              | 2 | 1   | 2   | 2   | 1   | 1             | 1.57 | 1    |
|                         | DM             | 3              | 1 | 1   | 3   | 1   | 1   | 1             | 1.57 | 1    |
Figure 5 | Biases in simulating the 95th, 90th, 50th and 25th percentiles of precipitation by uncorrected and bias-corrected model (a) CNRM-RCA4, (b) CNRM-RegCM4, (c) GFDL-RCA4, (d) GFDL-RegCM4 and (e) NorESM1-RCA4 [LS- linear scaling; LOCI- local intensity scaling; MPT-modified power transformation; DM- distribution mapping].
Performance of temperature BC methods

All BC methods increased the $R$ value of the uncorrected GCM–RCM pairs for the minimum temperature (see Figure 6(b)), while the LS method decreased the $R$ value of the GFDL-RCA4, MIROC5-RCA4 and NorESM1-RCA4 models; the VS method decreased the $R$ value of the NorESM1-RCA4 model while simulating the maximum temperature (see Figure 6(a)).

The biases in simulating the lower and upper extremes of temperature by various BC methods are shown in Figure 8.

The LS method of BC for maximum temperature had the least bias when simulating the first percentile, but had a higher bias when simulating the 99th percentile. All of the BC methods, on the other hand, significantly reduced biases when simulating both the 1st and the 99th percentiles for minimum temperature. Both the DM and VS methods performed equally well with the least bias for Tmin. Despite the fact that the same set of GCM–RCM combinations was used for both maximum and minimum temperatures, generally, the performance of the BC methods varies depending on the climate variables being studied.
The raw GCM–RCM pairs overestimate the TXx and TXn of maximum temperature by 2–13% and TNx and TNn of minimum temperature by 7–27% (Figure 9). All BC techniques, however, have a reduced PBIAS, as seen in Figure 9. In addition, all BCs resulted in a zero PBIAS in simulating annual and monthly mean temperatures for both maximum and minimum temperatures.

Overall, BC methods perform better for temperature than for precipitation when it comes to bringing raw GCM–RCM simulations closer to observations.

Streamflow simulation
The hydrological impact of the bias-corrected ranked models was evaluated by simulating monthly streamflow in the semi-distributed hydrological model SWAT. Streamflow simulations in the hydrological model were carried out over the period 1980–2010, with a warm-up period of 5 years, in order to reduce the impact of the initial conditions on the performance of the model. Streamflow simulations from 1985 to 2010 were, therefore, considered for further analysis. The model performance was evaluated in terms of Nash–Sutcliffe efficiency (NSE), RMSE to the standard deviation of observed data (RSR) and
Figure 8 | Biases in simulating the 1st and 99th percentiles of maximum and minimum temperatures (Tmax and Tmin) by uncorrected and bias-corrected GCM-RCM pairs (a) CanESM2-RegCM4 (b) GFDL-RCA4 (c) GFDL-RegCM4 (d) MIROC5-RCA4 and (e) NorESM1-RCA4 [LS- linear scaling; VS- variance scaling; DM- distribution mapping].
Moriasi et al. (2007) recommended PBIAS as general performance ratings for a monthly time step. Monthly hydrographs and flow duration curves are shown in Figure 10(a) for observed data and uncorrected GCM–RCM pairs. Model 1, representing streamflow simulated using rank 1 precipitation, maximum and minimum temperature models, is depicted. Model 2, model 3, model 4, and model 5 are also shown in the figure. Figure 10(b) illustrates the monthly average hydrograph corresponding to these models after bias correction using the best-performing BC method. The uncorrected GCM–RCM pairs simulate biased streamflow, with the highest deviation shown by streamflow simulated using the model ranked 5 for pr, Tmax, and Tmin (Figure 10(a)).

However, after BC, PBIAS has decreased substantially for all GCM–RCM pairs, ranging from $-1.66$ to $+0.77$. For model 1, model 2, model 3, model 4, and model 5, the RSR values were 0.69, 0.73, 0.72, 0.72, 0.75 and the NSE values were 0.52, 0.46, ...
0.49, 0.48 and 0.43, respectively. According to Moriasi et al. (2007), model simulation can be considered to be satisfactory if NSE $> 0.50$, RSR $< 0.70$ and PBIAS $< \pm 25\%$ for streamflow. As a result, only model 1 achieved satisfactory results in this case. The results presented above confirm the applicability of the EDAS method and the performance criteria used to rank climate models since the amount of bias in ranked models while simulating streamflow was directly proportional to the rank assigned to them by the EDAS approach. The impact of bias-corrected RCM simulated precipitation and temperature on streamflow simulation precludes direct application of raw climate model output to hydrological models.

Figure 10 represents the monthly average hydrograph of the simulations with observed climate variables (pr, Tmax and Tmin) and different combinations of GCM–RCM pairs. Here, rank 1 represents the streamflow simulated using model
ranked 1 for pr, Tmax and Tmin. Similarly, rank 1–2 corresponds to streamflow simulated using the average simulation of GCM–RCM pairs ranked 1 and 2 for each climate variable. The remaining combinations considered for streamflow simulation are rank 1–3 (average simulation of the top 3 models), rank 1–4 (average simulation of the top 4 models), rank 1–5 (average simulation of all the models) and uncorrected 1–5 (average simulation of all the uncorrected models). It can be observed that all model combinations except the average of uncorrected models closely follow the pattern of the simulated streamflow using observed climate variables.

The hydrological impact of bias-corrected ranked models was further evaluated using monthly flow duration curves as shown in Figure 12. The streamflow simulated using rank 1, rank 1–2, rank 1–3, rank 1–4 and rank 1–5 models closely follows the flow distribution simulated using observed climate variables. The deviation of flow distribution for the uncorrected 1–5 models from streamflow simulated using observed climate variables can also be seen in Figure 12. The PBIAS values for rank 1, rank 1–2, rank 1–3, rank 1–4 and rank 1–5 were 0.21, −5.27, −0.74, −0.81 and −1.04, respectively; the RSR values were 0.69, 0.62, 0.57, 0.55 and 0.52, respectively, and the NSE values were 0.52, 0.61, 0.67, 0.7 and 0.72, respectively.

Thus, the streamflow simulation of all model combinations can be considered to be satisfactory according to Moriasi et al. (2007). The performance of the model combination rank 1–3, rank 1–4 and rank 1–5 also falls under the performance rating ‘good’ as the value of the evaluation criteria PBIAS< ±15, RSR<0.6 and NSE>0.65 corresponds to the performance rating ‘good’. The above results also show that best performance was achieved (NSE=0.72) when streamflow was simulated using rank 1–5 (average of all models) models. When we considered streamflow simulation using different combinations of GCM–RCM pairs, the performance rating values improved. The uncertainty resulting from each individual model is reduced by the multimodel ensemble average (Yang et al. 2020). Since the performance of each individual model simulation has a direct impact on the performance of the multimodel ensemble, selecting the most appropriate model and/or model ensembles based on statistical criteria to evaluate various impacts is critical.

CONCLUSIONS

This study evaluated the hydrological impact of selected GCM–RCM pairs on the Meenachil river basin, Kerala. A total of five GCM–RCM pairs was considered for precipitation and a further set of five GCM–RCM pairs for both maximum and minimum temperatures. Daily simulations of GCM–RCM pairs chosen for precipitation, maximum and minimum temperatures based on literature study were ranked using the EDAS method based on the PBIAS, R, SS, MAE and RMSE performance

![Figure 12](https://example.com/figure12.png)

**Figure 12** | Flow duration curve of observed climate data and uncorrected and bias-corrected ranked GCM–RCM pairs of precipitation (pr), maximum temperature (Tmax) and minimum temperature (Tmin) [Rank 1 – rank 1 model of pr, Tmax and Tmin; Rank 1–2 – average of top 2 models; Rank 1–3 – average of top 3 models; Rank 1–4 – average of top 4 models; Rank 1–5 – average of all models; Uncorrected 1–5 – average of all uncorrected models].
The ranking was performed under EW and VW scenarios and the CRITIC method was employed to determine the objective weights of performance indicators. Then, the hydrological impact of different combinations of bias-corrected ranked models was assessed by simulating monthly streamflow in the hydrological model SWAT. This is the first attempt to evaluate CORDEX datasets using the EDAS method for ranking and the CRITIC method for measuring objective weights. The results of this study show that the EDAS method has the ability to rank models based on performance. Since an evaluation of climate models and BC methods for a new region is vital, the results of this study can be considered for further studies on climate change impacts and future projections of hydrological processes and climate change in the study area under consideration.

The major findings of this study are as follows:

1. The top-performing GCM–RCM pairs identified for precipitation (pr), maximum temperature (Tmax) and minimum temperature (Tmin) were GFDL-RCA4, GFDL-RegCM4 and MIROC5-RCA4, respectively.

2. In assessing the objective weights assigned to the performance indicators by the CRITIC method, the highest weight was assigned to SS with a weight of 0.25 for pr, R with a weight of 0.44 for Tmax and MAE with a weight of 0.35 for Tmin. PBIAS was the common lowest weighting performance indicator in all three cases. Furthermore, the lowest weight was assigned to RMSE and MAE, with a weight of 0.13 for Tmax and RMSE with a weight of 0.14 for Tmin.

3. Despite the fact that performance metrics were given unequal weights in the VW scenario, the ranks allocated to the GCM–RCM pairs were the same as those in the EW scenario.

4. The best BC method identified for the GCM–RCM pairs of precipitation based on 11 performance indicators was the MPT method for CNRM-RCA4 and the LOCI method for the rest of the models.

5. Unlike precipitation, BC methods for Tmax and Tmin were evaluated using seven performance measures. As a result, the best method of BC chosen for Tmax simulations of CanESM2-RegCM4, GFDL-RegCM4 and MIROC5-RCA4 was the VS method. Both the VS and the DM methods performed equally well in the correction of NorESM1-RCA4 Tmax simulations. Similarly, the DM method was the best BC method for the Tmin simulations of CanESM2-RegCM4 and GFDL-RegCM4, both the DM and VS methods for GFDL-RCA4 and NorESM1-RCA4, and both the LS and VS methods for the MIROC5-RCA4 model.

6. The hydrological impact of ranked models in simulating monthly streamflow was assessed using various combinations of ranked models, and it was discovered that ‘good’ streamflow simulation could be achieved using the average of bias-corrected top three models (rank 1–3), top four models (rank 1–4) and the average of all models (rank 1–5). Even with only the bias-corrected rank 1 model, a satisfactory model output could be achieved. Above all, with NSE=0.72, RSR=0.52 and PBIAS=−1.04, the highest performing model combination was the average of all models.

**DATA AVAILABILITY STATEMENT**

All relevant data are included in the paper or its Supplementary Information.

**REFERENCES**

Arnold, J. G., Srinavasan, R., Mutthia, R. S. & Williams, J. R. 1998 Large-area hydrologic modeling and assessment: Part I. Model development. *J. Am. Water Resour. Assoc.* 34, 73–89.

Chen, J., Brissette, F. P. & Leconte, R. 2011 Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *J. Hydrol.* 401 (3–4), 190–202. http://dx.doi.org/10.1016/j.jhydrol.2011.02.020.

Chen, J., Brissette, F. P., Chaumont, D. & Braun, M. 2015 Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resour. Res.* 49, 4187–4205. doi:10.1002/wrcr.20351.

Choudhary, A., Dimri, A. P. & Maharana, P. 2018 Assessment of CORDEX-SA experiments in representing precipitation climatology of summer monsoon over India. *Theor. Appl. Climatol.* 134, 283–307. https://doi.org/10.1007/s00704-017-2274-7.

Diakoulaki, D., Mavrotas, G. & Papayannakis, L. 1995 Determining objective weights in multiple criteria problems: the critic method. *Comput. Oper. Res.* 22 (7), 763–770. https://doi.org/10.1016/0305-0548(94)00059-H.

Diallo, I., Sylla, M. B., Giorgi, F., Gaye, A. T. & Camara, M. 2012 Multimodel GCM-RCM ensemble-based projections of temperature and precipitation over West Africa for the early 21st century. *Int. J. Geophys.* 19. https://doi.org/10.1155/2012/972896.

Fowler, H. J., Blenkinsop, S. & Tehalf, C. 2007 Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* 27, 1547–1578. https://doi.org/10.1002/joc.1556.

Ghorabaee, M. K., Zavadskas, E. K., Olfat, L. & Turskis, Z. 2015 Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS). *Informatica* 26 (3), 435–451.
