Selection of Suitable Precipitation CMIP-5 Sets of GCMs for Iraq Using a Symmetrical Uncertainty Filter

Thair S Khayyun\textsuperscript{1}, Imzahim A Alwan\textsuperscript{2} and Ali M Hayder\textsuperscript{3}

\textsuperscript{1,2} Civil Engineering Dept., University of Technology, Baghdad, Iraq
\textsuperscript{3} Hydraulic Structures and Water Resources Dept., Faculty of Engineering, University of Kufa, Najaf, Iraq and PhD. Student in Civil Engineering Dept., University of Technology, Baghdad, Iraq, E-mail: alim.albayati@uokufa.edu.iq

Abstract. Prediction of future water resources in Iraq for the period 2020 to 2100 based on different scenarios of climatic change models by harnessing many calibrated Arc-SWAT models needs to select future precipitation data from suitable sources. Therefore, the selection of the appropriate source of the possible future precipitation time series data sets was studied by examining 20 models of the fifth phase of the Coupled Model Inter-comparison Project (CMIP-5.0) for General Circulation Models (GCMs). The Symmetrical Uncertainty (SU) approach was used to determine the performance of the 20 CMIP-5.0 sets of GCMs against a monthly scale of historical precipitation datasets at each one of 35 rain gauge stations spread throughout Iraq, and an appropriate ensemble of GCMs selected. The examined models were ranked as 1st, 2nd, 3rd, etc. based on the SU values at each station and the final ranking of the models was carried out using a multi-criteria decision-making (MCDM) method. The results showed that the HadGEM2-AO and HadGEM2-ES were the best (1st ranked) models at 31 stations, while the MIROC-5 and CSIRO-Mk.3.6 were the best models at Zakho and Duhok stations, respectively, and the BCC.CSM1.1.m and FIO.ESM were the best models at Samarra and Hill stations, respectively. There was a variation in simulation preference regionally between the two models HadGEM2-AO and HadGEM2-ES in the second ranking except at Tel-Afer and Najaf stations. No predominant models were found in other ranks throughout Iraq. The results of the final ranking of these 20 CMIP-5 sets by the MCDM method thus showed that there are only four suitable GCMs, HadGEM2-AO, HadGEM2-ES, CSIRO.Mk3.6, and MIROC5 for data projections studying scenarios involving future water resources in Iraq.

Keywords: climate change, general circulation models, Iraq, precipitation, symmetrical uncertainty.

1. Introduction

Water resource management and planning are becoming more challenging due to climate change uncertainty affecting the future [1]. Variations in temperature and precipitation are likely to increase or decrease the total and seasonal water resources [2]; thus, climate change models that predict the state of climate have become to an increasing extent dependent on the impact of climatic changes on hydrological processes, which affects the planning of water resource management. Climate change assessment is generally estimated by using simulations known as General Circulation Models (GCMs), which are coupled numerical models that represent the different systems of earth in terms of the land surface, oceans, sea-ice, and atmosphere; these offer significant potential for climate change studies and variability [3]. Nevertheless, a large uncertainty is associated with climate simulation by GCMs due to their model structures, assumptions, calibration processes, parametrization and so on [4]. Subsequently, not all the GCMs can be used directly to project future climate effects for a particular area. Generally, in order to
reduce the uncertainty associated with the GCMs, a certain number of GCMs are chosen for a particular area, eliminating any models that do not offer similarities to the climate of the selected area [5].

Election of a convenient collection of the GCMs is an essential defy for projection of climate influence assessment, but there are many approaches for selecting these GCMs: (1) the historical performance evaluation approach where selection of a GCM depends only on its ability to replicate past climate datasets without taking into consideration the future projection of the climate [6]; (2) the envelope approach, where selection of a GCM depends only on its agreement with future climate without taking into account the GCM’s ability to simulate historical climate datasets [7]; and (3) the hybrid approach, which combines the two previous methods for selection a GCM, with GCMs screened according to their historical performance and a final set of GCMs selected based on an envelope approach [5]. In order to select an ensemble of GCMs from all considered GCMs, the historical performance evaluation approach is more practical, as the whole group of the GCMs covers the entire domain of projections [8]. Many researchers have thus proposed that this approach (past performance-based approach) is the most appropriate due to the fact that the GCMs which are best in replicating the historical climatic conditions are more likely to forecast future climate accurately [9,10].

The performance of the fifth phase of the Coupled Model Inter-comparison Project (CMIP-5.0) in which datasets of GCMs were evaluated by applying various types of filters such as spectral analysis [11], difference performance indicators [12,13], clustering hierarchy approaches [14,15], Bayesian weighting [16], and weighted skill scores [17,18] is thus relevant. Furthermore, different statistical measures have been conducted in parallel with these filters to allow GCM assessment, ranking and selection [19,20].

At present, symmetrical uncertainty (SU) is one of the most powerful algorithms which is applied for the selection of variables; using this, the ranking of GCMs can be performed without the use of several conventional statistical metrics [8,21]. The SU algorithm is a filter-based tool which selects the variables in an unbiased and reliable manner [22]; this algorithm has ability to track the agreement and disagreement between time series datasets of measured climate variable and historical period results by the GCMs and then select GCMs according to their capability to replicate the measured climate variables. Furthermore, the distributions underlying the data and various conditional dependencies are not required for the SU algorithm, a property which is its major advantage.

The main objectives of this study are thus (1) to assess the ability of 20 CMIP-5.0 GCMs to replicate historical datasets of precipitation from 35 observation rain gauge stations, using the SU algorithm to select the most suitable set of the GCMs for reliable precipitation projection in Iraq, with its varied climate and large seasonal and spatial discrepancies; and (2) to use the selected set of GCMs from this study to predict and evaluate the future hydrology and water resources in Iraq for the period 2020 to 2100 based on four scenarios outlined by representative concentration pathways (RCPs: here, RCP-2.6, RCP-4.5, RCP-6.0, and RCP-8.5) of climatic change models by harnessing multiple calibrated Arc-SWAT models such as in [23].

2. Geography and Climate of Iraq

2.1. Geography
Iraq is located in the southwest of the Asian continent (latitude 29° 03’ 40” N to 37° 22’ 53” N and longitude 38° 47’ 37” E to 48° 37’ 40” E), and covering an area of 437,299 square kilometers. Iraq is bordered by Iran to the east, the Jordan and the Syria to the west, Turkey to the north, and Saudi Arabia and Kuwait to the south. The topography of Iraq is basin-like, being composed of the Great Mesopotamian alluvial plain of the Euphrates and the Tigris rivers. High-altitude areas (mountains and plateaus) in the east and the north are circumfluent to this alluvial plain, while in the south and west, it is surrounded by vast desert areas [24]. Its natural surface levels thus significantly vary, from highs in the north, north-east and west regions to low elevations in the south-east and south-west regions (see Figure 1).
2.2. Climate of Iraq
The climate of Iraq is fundamentally continental, subtropical and semi-arid, with a Mediterranean climate in the mountainous regions at the north and north-eastern areas. Rainfall is very seasonal, being concentrated in the winter from December to February, with the exception of the north and northeast regions of Iraq, where the rainfall extends from November to April. The annual rainfall is between 1,200 mm in the northeast to less than 100 mm over 60 percent of the country in the south, with the overall average estimated at 216 mm. Winters are cool to cold, with daytime temperatures of about 16 °C, dropping at night to 2 °C with the possibility of frost. Summers are dry and hot or extremely hot, with shade temperatures of over 43 °C during July and August, dropping at night to 26 °C [24].

Figure 1. Topographical map of Iraq showing spatial distribution of rain gauge stations.

3. Data and Sources
Precipitation datasets used for analysis in this study came from two sources, observed monthly precipitation and the GCMs based on the CMIP-5.0. These datasets are discussed in detail in the following sections.

3.1. Observed Datasets.
In this study, the observed historical monthly precipitation time series datasets from 35 rain gauge stations spread throughout Iraq were used for a 78-year period between 1929 and 2005. The spatial distribution of these stations is presented in Figure 1.

3.2. Model Datasets.
The historical monthly CMIP-5.0 precipitation datasets as simulated by the GCMs were downloaded for the various time periods from the centre for environmental data analysis (CEDA) database as network common data form (NetCDF) files. The future accumulative monthly precipitation projections for Iraq as a whole and for the four possible scenarios of RCPs (RCP-2.6, RCP-4.5, RCP-6.0, and RCP-8.5) were
available only for 20 CMIP-5.0 GCMs; thus, these 20 GCMs were selected for this study, as these four RCP scenarios provide complete range of possible impacts of climate change in the future. The CMIP-5.0s of the GCMs’ names and detailed information regarding modeling institute and resolution are given in Table 1. In order to make comparisons in a similar manner, the selected CMIP-5.0 models for precipitation simulation were interpolated using an inverse distance weighted method to find the precipitation at the same positions for each of 35 rain gauge stations spread throughout Iraq (see Figure 1). Several R-packages (sp, rgdal, raster, and ncdf4) were used for extraction of numeric climate data values from the NetCDF files at each grid point of CMIP-5.0 model.

Table 1. The CMIP-5.0s of GCMs used in the study.

| No | Model Developer/Institution | Name of CMIP-5.0 | Spatial Resolution (Log × Lat) |
|----|-----------------------------|------------------|--------------------------------|
| 1  | Beijing Climate Center, China | BCC-CSM1.1m | 1.96° × 1.13° |
| 2  | Beijing Climate Center, China | BCC-CSM1-1 | 2.79° × 2.81° |
| 3  | Bjerknes Centre for Climate Research, Norwegian Meteorological Institute, Norway | NorESM1-M | 1.90° × 2.50° |
| 4  | Bjerknes Centre for Climate Research, Norwegian Meteorological Institute, Norway | NorESM1-ME | 1.90° × 2.50° |
| 5  | Commonwealth Scientific and Industrial Research Organization, Australia | CSIRO-Mk3.6.0 | 1.86° × 1.88° |
| 6  | Geophysical Fluid Dynamics Laboratory, USA | GFDL-CM3 | 2.00° × 2.50° |
| 7  | Geophysical Fluid Dynamics Laboratory, USA | GFDL-ESM2M | 2.00° × 2.50° |
| 8  | Institute Pierre Simon Laplace, France | IPSL-CM5A-LR | 1.89° × 3.75° |
| 9  | Institute Pierre Simon Laplace, France | IPSL-CM5A-MR | 1.27° × 2.50° |
| 10 | Met Office Hadley Centre, UK | HadGEM2-AO | 1.25° × 1.88° |
| 11 | Met Office Hadley Centre, UK | HadGEM2-ES | 1.25° × 1.88° |
| 12 | Meteorological Research Institute, Japan | MRI-CGCM3 | 1.12° × 1.13° |
| 13 | National Aeronautics and Space Administration / Goddard Institute for Space Studies, USA | GISS-E2-H | 2.50° × 2.50° |
| 14 | National Aeronautics and Space Administration / Goddard Institute for Space Studies, USA | GISS-E2-R | 2.50° × 2.50° |
| 15 | National Center for Atmospheric Research USA | CESM1-CAM5 | 0.94° × 1.25° |
| 16 | National Center for Atmospheric Research, USA | CCM4 | 0.94° × 1.25° |
| 17 | National Institute for Environmental Studies, Japan Agency for Marine-Earth Science and Technology, and University of Tokyo, Japan | MIROC5 | 1.40° × 1.41° |
| 18 | National Institute for Environmental Studies, Japan Agency for Marine-Earth Science and Technology, and University of Tokyo, Japan | MIROC-ESM | 2.79° × 2.81° |
| 19 | National Institute for Environmental Studies, Japan Agency for Marine-Earth Science and Technology, and University of Tokyo, Japan | MIROC-ESM-CHEM | 2.80° × 2.81° |
| 20 | The First Institute of Oceanography, China | FIO-ESM | 2.80° × 2.81° |

4. Methodology

4.1. Procedure
In this study, the CMIP-5.0s of GCMs are predicted to be capable of simulating the precipitation in Iraq reasonably well. The observed precipitation at 35 station (see Figure 1) and the association of CMIP-5.0 datasets with station data were thus compared at a monthly scale to assess the performance of the 20 GCMs. The procedure followed in this study can be summarised in the following steps:
1- Historical available datasets of precipitation at each of the 35 station were prepared and missing data within the time series interpolated using the inverse distance weighting interpolation method.
2- The same historical time series dataset range of precipitation at observation stations were generated for each of the 20 GCMs.
3- The SU algorithm was applied for each set of rain gauge station data versus CMIP-5.0 data for the 20 GCMs.
4- Ranking of GCMs were obtained by multi-criteria decision-making (see section 4.3).

4.2. Symmetric Uncertainty Filter Approach
It was expected that the periodic differences in the seasonal and annual historical gauged climate data sets (rainfall and temperature) could be replicated by using the capabilities of the GCMs [8]. In the Symmetric Uncertainty SU algorithm, the degree of similarity of the GCM data set with the observed climate data was used for ranking the GCMs throughout the entire sequence of time series data. Where the SU could track the information included in the different GCMs, the most relevant GCM could be selected for a given variable. Moreover, the SU provides a generic measure that associates dependent and independent variables without relying on knowledge of the type of the underlying distributions and conditional dependencies; this feature is one of the significant advantages of the SU [25]. SU is also effective for feature selection in large datasets [26]. Moreover, the SU filter is a useful method for ranking GCMs without the use of the conventional statistical metrics such as coefficient of determination and normalised root mean square error [8].

The SU is a filtering approach that works on the basis of the concept of entropy and the Mutual Information (MI) technique or Information Gain (IG). Entropy can be measured as the uncertainty for a random variable [26,27]. Let $O_r$ be the observed monthly rainfall and $P_r$ the simulated gridded historical rainfall from the GCMs at a grid point. The entropy of $O_r$ is thus defined as

$$H(O_r) = -\sum_{i=1}^{n} p(O_{r_i}) \left(p(O_{r_i})\right)$$

and the entropy of $O_r$ after observing values of $P_r$ is defined as

$$H(P_r) = -\sum_{i=1}^{n} p(P_{r_i}) \sum_{i=1}^{n} p(P_{r_i}) \left(p(P_{r_i})\right)$$

where, $p(O_r)$ is the prior probability density functions of $O_r$, $p(O_r, P_r)$ is the posterior joint probability density function of $P_r$, and $n$ is the total number of paired observed and simulated rainfall values.

The MI (IG) is an essential technique to express how much information is correlated between two attributes, as shown in Figure 2; this can be expressed as the difference between the sum of the marginal entropies and their joint entropy. Thus, the MI (IG) between the $O_r$ and the $P_r$ simulated by GCMs can be expressed as

$$IG(P_r) = H(O_r) - H(P_r)$$

The IG is symmetrical for two random variables. However, IG is biased toward those features that have higher values. This can be overcome using an SU algorithm in which the estimated IG is divided by the sum of entropies of observed monthly rainfall ($O_r$) data and the simulated gridded historical rainfall ($P_r$) from the GCMs data [26,27], given as follows:

$$SU(P_r) = 2 \frac{IG(P_r)}{H(O_r) + H(P_r)}$$
If the value of SU \( (O_r, P_r) \) is equal to 1, then a strong agreement between the \( O_r \) data and the \( P_r \) of GCMs data exists; if the value of SU \( (O_r, P_r) \) is equal to zero, it indicates absolute disagreement between them [28].

The SU value has two main functions: (1) it can remove features with SUs lower than threshold and (2) every feature’s weight is used to guide the initialisation of the population for genetic algorithms in the memetic framework. Those featured which have larger SU valued will get higher weights, while those features with less SU value are removed.

\[
H(O_r) - IG(O_r | P_r) - H(P_r)
\]

**Figure 2.** The concept of mutual information (Information Gain), After [8].

### 4.3. Ranking GCMs Using the Weighting Method

A number of researchers in recent years have applied multi-criteria decision-making methods (MCDM) for ranking and selection of GCMs based on their performance at each grid point [21,29]. Ranking of GCMs at a single station point is simple and can be easily assessed; in contrast, however, the assessment and selection of GCMs becomes difficult when all stations at various stations points are evaluated, as many GCMs give different results at different station points. To simplify this complexity, the MCDM method was used for ranking the GCMs for the whole of Iraq by applying the following steps:

1. The GCMs were ranked \( (1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 5^{th}, \ldots) \) at each station based on the values of the SU.
2. A specified weight \( (w_i) \) which assigned as the inverse of the rank given to each GCM at the different station points.
3. The frequency of each rank \( (FR_i) \) for each GCM was enumerated.
4. Total ranking weight \( (TRW) \) was based on the MCDM for each GCM and calculated as

\[
TRW = \sum_{i}^{5} FR_i \times w_i
\]

5. The final ranking of GCMs was obtained by sorting the total ranking weights in descending order.

In this study, grades of GCMs up to the 5\(^{th}\) rank at each station point were taken into consideration; other ranks were ignored because it was assumed that they did not have the ability to simulate precipitation at those stations sufficiently well.

### 5. Results

#### 5.1. Initial Ranking of CMIP-5.0s of the GCMs at Stations

The SU values were estimated for the 20 GCMs based on CMIP-5.0 historical datasets for each of the 35-station points. Subsequently, the rankings of the CMIP-5.0s of the GCMs were provided with regard to their ability to replicate rainfall at each station (step 1 in section 4.3): a heatmap plot was constructed to
explain the spatial distribution of results, shown in Figure 3. In general, there are several predominant GCMs in most spatial distributions and at various stations points (Figure 3). In contrast, the intermediate zone of the plot has different ranks of models in spatial distribution. Nevertheless, the results indicate that there is no unique predominant first ranked CMIP-5.0 model for all regions of Iraq.

Figure 3. GCMs Ranking at each observation station in Iraq according to performance in replicating precipitation.

5.2. Spatial Distribution of Top Ranked GCMs

Although multiple models of datasets were top ranked at various stations, as shown in the previous section, an estimate of the initial spatial distribution of the ranking effects on the final ranking is required. To determine the rank frequency of each GCM at the different stations, a second heatmap plot was thus prepared, as shown in Figure 4. This plot indicates that the models HadGEM2-AO, HadGEM2-ES, CSIRO-Mk-3.6, and GISS-E2-H have the highest frequencies in the first five ranks of spatial distribution for CMIP-5.0 models in Iraq.

Moreover, Figure 5 was prepared to show the detailed spatial distributions of the CMIP-5.0 models ranked 1st, 2nd, 3rd, 4th, and 5th, based on their SU values. In this figure, various colours were used in order to represent different CMIP-5.0 models. The figure illustrates that the HadGEM2-AO and HadGEM2-ES are the best (1st ranked) among the CMIP-5.0 models for the largest number of rain gauge stations in Iraq (31 stations), while MIROC-5 and CSIRO-Mk-3.6 were the best models at the Zakho and Duhok stations in the northern region, and bcc.csm1.1.m and FIO.ESM were best at Samarra and Hilla stations, respectively. Furthermore, at the second-rank level, there is an exchange in the simulation preference regionally between the two models (HadGEM2-AO and HadGEM2-ES) with the exception of Tel-Afer and Najaf stations. These two models were thus 2nd rank for 31 of the 35 stations. In contrast, no predominant models were found at the third, fourth, and fifth ranks for Iraq as a whole; the CSIRO-Mk-3.6 model had the highest frequency in the 3rd and 4th ranking (13 and 11 times), while the GISS.E2.H model featured at 8 stations in the 5th ranking.
Figure 4. Frequency of each GCM rank at the 35 observation precipitation stations in Iraq.

Figure 5. The CMIP-5.0 ranked 1st, 2nd, 3rd, 4th, and 5th spatial distribution positions at different rain gauge stations for simulating precipitation in Iraq.

5.3. Ranking and selection of CMIP-5.0s of the GCMs for Iraq

The MCDM was applied to each of the CMIP-5.0 models at different stations up to the 5th rank, using equation (5) to get the final total ranking weight (TRW) of each analysed model. Subsequently, the final ranking of each model was generated relative to the whole of Iraq. Table 2 and Figure 6 display the total ranking weight values, the final rankings of the examined models, and the bar plot for this. There were significant differences between the TRWs of models. The HadGEM2-AO model was found to have the highest TRW with values of 24.9, being ranked as the 1st model while HadGEM2-ES had the second highest
TRW values of 22.7, being ranked as the 2nd. Thus, HadGEM2-AO and HadGEM2-ES are the two best CMIP-5.0 models for precipitation for Iraq as a whole.

In modelling hydrological processes by using hydrological models such as the SWAT model, precipitation time series datasets are required to study the climate change impact on future hydrological processes (runoff) for the Tigris and Euphrates rivers catchment areas inside Iraqi territory. Therefore, CMIP-5.0 datasets of GCMs which can simulate precipitation properly are desirable for water resource analysis in Iraq. In this study, the selection of final suitable set of GCMs was based on their TRW, and it was assumed that GCM is qualified for simulation of precipitation when the GCM achieved a TRW value greater than the average value for all GCMs (3.97). As shown in Table 2, only four GCMs achieved this: HadGEM2-AO, HadGEM2-ES, CSIRO.Mk3.6, and MIROC5. Tus, only datasets of precipitation from these four models are suitable for studying scenarios of future water resource management in Iraq.

Table 2. The total ranking weight values and final ranking of the CMIP-5.0 datasets of GCMs that examined in this study.

| Model Name            | Total Ranking Weight | Final Rank |
|-----------------------|----------------------|------------|
| bcc.csm1.1            | 0.6                  | 11         |
| bcc.csm1.1.m          | 3                    | 7          |
| CCSM4                 | 0                    | 15         |
| CESM1.CAM5            | 0                    | 15         |
| CSIRO.Mk3.6           | 10                   | 3          |
| FIO.ESM               | 1.2                  | 9          |
| GFDL.CM3              | 1                    | 10         |
| GFDL.ESM2M            | 0                    | 15         |
| GISS.E2.H             | 3.3                  | 6          |
| GISS.E2.R             | 0.2                  | 14         |
| HadGEM2-AO            | 24.9                 | 1          |
| HadGEM2-ES            | 22.7                 | 2          |
| IPSL.CM5A.LR          | 1.9                  | 8          |
| IPSL.CM5A.MR          | 3.7                  | 5          |
| MIROC5                | 5.9                  | 4          |
| MIROC.ESM             | 0                    | 15         |
| MIROC.ESM.CHEM        | 0                    | 15         |
| MRLC.GCM3             | 0                    | 15         |
| NorESM1.M             | 0.4                  | 13         |
| NorESM1.ME            | 0.5                  | 12         |
Figure 6. Bar plot of the total ranking weight values and final ranking of the CMIP-5.0 datasets of GCMs that examined in this study.

6. Conclusions

The ability of CMIP-5.0 historical datasets from 20 GCMs to simulate precipitation was assessed by using historical datasets from 35 rain gauge stations in Iraq. The main conclusions can be summarised as

(1) the results show that there are variations in the performance of GCMs in generating precipitation at the studied areas using the SU filter;

(2) HadGEM2-AO, HadGEM2-ES, CSIRO-Mk-3.6, and GISS-E2-H models were the most frequent in the first five ranks of spatial distribution for the CMIP-5.0s of GCMs models in Iraq as a whole;

(3) the 1st ranked models in spatial distribution for the 35 stations were HadGEM2-AO for 17 stations, HadGEM2-ES for 14 stations, MIROC-5 for 1 station, CSIRO-Mk-3.6 for 1 station, bcc.csm1.1.m for 1 station, and FIO.ESM for 1 station;

(4) HadGEM2-AO and HadGEM2-ES models showed spatial switching in their performance at most stations; and

(5) the final ranking of GCMs when MCDM was used, showed that HadGEM2-AO, HadGEM2-ES, CSIRO.Mk3.6, and MIROC5 were suitable models for the projection of precipitation in Iraq.

1. References

[1] Ficklin D L, Stewart I T and Maurer E P 2013 Climate change impacts on streamflow and subbasin-scale hydrology in the Upper Colorado River Basin PLoS One 8 e71297

[2] Parry M, Parry M L, Canziani O, Palutikof J, Van der Linden P and Hanson C 2007 Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC vol 4 (Cambridge University Press)

[3] Setegn S G, Rayner D, Melesse A M, Dargahi B, Srinivasan R and Wörman A 2011 Climate change impact on agricultural water resources variability in the Northern Highlands of Ethiopia Nile River Basin (Springer) pp 241–66

[4] Yip S, Ferro C A T, Stephenson D B and Hawkins E 2011 A simple, coherent framework for
partitioning uncertainty in climate predictions J. Clim. 24 4634–43

[5] Lutz A F, ter Maat H W, Biemans H, Shrestha A B, Wester P and Immerzeel W W 2016 Selecting representative climate models for climate change impact studies: an advanced envelope-based selection approach Int. J. Climatol. 36 3988–4005

[6] Srinivasa Raju K and Nagesh Kumar D 2015 Ranking general circulation models for India using TOPSIS J. Water Clim. Chang. 6 288–99

[7] Warszawski L, Frieler K, Huber V, Piontek F, Serdeczny O and Schewe J 2014 The inter-sectoral impact model intercomparison project (ISI–MIP): project framework Proc. Natl. Acad. Sci. 111 3228–32

[8] Khan N, Shahid S, Ahmed K, Ismail T, Nawaz N and Son M 2018 Performance assessment of general circulation model in simulating daily precipitation and temperature using multiple gridded datasets Water 10 1793

[9] Hidalgo H G, Das T, Dettinger M D, Cayan D R, Pierce D W, Barnett T P, Bala G, Mirin A, Wood A W and Bonfils C 2009 Detection and attribution of streamflow timing changes to climate change in the western United States J. Clim. 22 3838–55

[10] Raju K S, Sonali P and Kumar D N 2017 Ranking of CMIP5-based global climate models for India using compromise programming Theor. Appl. Climatol. 128 563–74

[11] Jiang X, Waliser D E, Xavier P K, Petch J, Klingaman N P, Woolnough S J, Guan B, Bellon G, Crueger T and DeMott C 2015 Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations J. Geophys. Res. Atmos. 120 4718–48

[12] Afshar A A, Hasanzadeh Y, Besalatpour A A and Pourreza-Bilondi M 2017 Climate change forecasting in a mountainous data scarce watershed using CMIP5 models under representative concentration pathways Theor. Appl. Climatol. 129 683–99

[13] Xuan W, Ma C, Kang L, Gu H, Pan S and Xu Y-P 2017 Evaluating historical simulations of CMIP5 GCMs for key climatic variables in Zhejiang Province, China Theor. Appl. Climatol. 128 207–22

[14] Yokoi S, Takayabu Y N, Nishii K, Nakamura H, Endo H, Ichikawa H, Inoue T, Kimoto M, Kosaka Y and Miyasaka T 2011 Application of cluster analysis to climate model performance metrics J. Appl. Meteorol. Climatol. 50 1666–75

[15] Knutti R, Masson D and Gettelman A 2013 Climate model genealogy: Generation CMIP5 and how we got there Geophys. Res. Lett. 40 1194–9

[16] Min S and Hense A 2006 A Bayesian approach to climate model evaluation and multi-model averaging with an application to global mean surface temperatures from IPCC AR4 coupled climate models Geophys. Res. Lett. 33

[17] Maxino C C, McAvaney B J, Pitman A J and Perkins S E 2008 Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation Int. J. Climatol. A J. R. Meteorol. Soc. 28 1097–112

[18] Perkins S E, Pitman A J, Holbrook N J and McAneney J 2007 Evaluation of the AR4 climate models’ simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions J. Clim. 20 4356–76
[19] Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram J. Geophys. Res. Atmos. 106 7183–92

[20] Gu H, Yu Z, Wang J, Wang G, Yang T, Ju Q, Yang C, Xu F and Fan C 2015 Assessing CMIP5 general circulation model simulations of precipitation and temperature over China Int. J. Climatol. 35 2431–40

[21] Salman S A, Shahid S, Ismail T, Ahmed K and Wang X-J 2018 Selection of climate models for projection of spatiotemporal changes in temperature of Iraq with uncertainties Atmos. Res. 213 509–22

[22] Witten I H, Paynter G W, Frank E, Gutwin C and Nevill-Manning C G 2005 Kea: Practical automated keyphrase extraction Design and Usability of Digital Libraries: Case Studies in the Asia Pacific (IGI Global) pp 129–52

[23] Khayyun T S, Alwan I A and Hayder A M 2019 Hydrological model for Hemren dam reservoir catchment area at the middle River Diyala reach in Iraq using ArcSWAT model Appl. Water Sci. 9 133

[24] Frenken K 2009 Irrigation in the Middle East region in figures AQUASTAT Survey-2008. Water Reports

[25] Wu Y and Zhang A 2004 Feature selection for classifying high-dimensional numerical data Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. vol 2 (IEEE) pp II–II

[26] Kannan S S and Ramaraj N 2010 A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm Knowledge-Based Syst. 23 580–5

[27] Singh B, Kushwaha N and Vyas O P 2014 A feature subset selection technique for high dimensional data using symmetric uncertainty J. Data Anal. Inf. Process. 2 95

[28] Shreem S S, Abdullah S and Nazri M Z A 2016 Hybrid feature selection algorithm using symmetrical uncertainty and a harmony search algorithm Int. J. Syst. Sci. 47 1312–29

[29] Raju K S and Kumar D N 2014 Ranking of global climate models for India using multicriterion analysis Clim. Res. 60 103–17