Ranking of global climate models for India using multicriterion analysis

Article *in* Climate Research · August 2014

DOI: 10.3354/cr01222

Srinivasa Raju Komaragiri
BITS Pilani, Hyderabad

D Nagesh Kumar
Indian Institute of Science

Vol. 60: 103–117, 2014
doi: 10.3354/cr01222

CLIMATE RESEARCH
Clim Res

Published online June 17

Ranking of global climate models for India using multicriterion analysis

K. Srinivasa Raju[1], D. Nagesh Kumar[2,3,*]

[1]Department of Civil Engineering, Birla Institute of Technology and Science-Pilani, Hyderabad campus, India
[2]Center for Earth Sciences, Indian Institute of Science, 560012 Bangalore, India

[3]*Present address:* Department of Civil Engineering, Indian Institute of Science, 560012 Bangalore, India

ABSTRACT: Eleven GCMs (BCCR-BCCM2.0, INGV-ECHAM4, GFDL2.0, GFDL2.1, GISS, IPSL-CM4, MIROC3, MRI-CGCM2, NCAR-PCMI, UKMO-HADCM3 and UKMO-HADGEM1) were evaluated for India (covering 73 grid points of 2.5° × 2.5°) for the climate variable 'precipitation rate' using 5 performance indicators. Performance indicators used were the correlation coefficient, normalised root mean square error, absolute normalised mean bias error, average absolute relative error and skill score. We used a nested bias correction methodology to remove the systematic biases in GCM simulations. The Entropy method was employed to obtain weights of these 5 indicators. Ranks of the 11 GCMs were obtained through a multicriterion decision-making outranking method, PROMETHEE-2 (Preference Ranking Organisation Method of Enrichment Evaluation). An equal weight scenario (assigning 0.2 weight for each indicator) was also used to rank the GCMs. An effort was also made to rank GCMs for 4 river basins (Godavari, Krishna, Mahanadi and Cauvery) in peninsular India. The upper Malaprabha catchment in Karnataka, India, was chosen to demonstrate the Entropy and PROMETHEE-2 methods. The Spearman rank correlation coefficient was employed to assess the association between the ranking patterns. Our results suggest that the ensemble of GFDL2.0, MIROC3, BCCR-BCCM2.0, UKMO-HADCM3, MPI-ECHAM4 and UKMO-HADGEM1 is suitable for India. The methodology proposed can be extended to rank GCMs for any selected region.

KEY WORDS: Entropy · Performance indicators · PROMETHEE-2 · River basin

1. INTRODUCTION

Changes in the concentration of greenhouse gases and in the radiative balance of the atmosphere cause corresponding changes in temperature and precipitation patterns (Bates et al. 2008). One of the important impacts of future climate changes on society will be changes in regional water availability for various purposes such as drinking water supply and irrigation, and the occurrence of extreme events, including floods and droughts. To address this problem, global climate models (GCMs) have been developed to simulate the present climate and are gaining importance due to their ability to project future climate changes and consequently their impacts, e.g. on hydrologic systems (Smith & Chiew 2010, Pitman et al. 2012).

The climate system is represented in a simplified form in GCMs, with combinations of models for different components of the climate system. While GCMs demonstrate significant skill at continental and hemispheric spatial scales and incorporate a large proportion of the complexity of the global system (Reichler & Kim 2008), they are inherently unable to represent local subgrid-scale features and dynamics. However, in considering impacts of global climate changes, the focus is primarily on societal responses to the local and regional consequences of the projected large-scale changes.

*Corresponding author: nagesh@civil.iisc.ernet.in

Xu (1999) inferred that the accuracy of GCMs (1) decreases with increasingly finer spatial and temporal scales; (2) decreases from free tropospheric variables to surface variables, while the variables at the ground surface have direct use in water balance computations; (3) decreases from climate-related variables — i.e. wind, temperature, humidity and air pressure — to precipitation, evapotranspiration, runoff and soil moisture, while the latter variables are of key importance in hydrologic regimes.

The uncertainties associated with the formulation of GCMs arise due to a number of factors and lead to significant variability across model simulations of future climates. These factors include effects of aerosols that are differently parameterised in different GCMs, initial and boundary conditions for each GCM, parameter and model structure of GCMs, randomness, and future greenhouse gas emissions. These uncertainties accumulate from various levels, from the GCM to the downscaling levels and may propagate down to the local levels, which may, in turn, affect the adaptation studies that would be used as the basis for implementation. In other words, climate model simulations at a local or regional scale can be highly uncertain (Mujumdar & Nagesh Kumar 2012). Moreover, in many regional hydrologic assessments, time and resource constraints limit the number of GCMs that can be used in determining impacts (Wilby & Harris 2006). An additional complexity is the lack of observed data to evaluate the GCM outputs in modelling a future climate. All of these aspects necessitate evaluating the available GCMs with more accuracy. Knutti et al. (2010) suggested that the skill or performance of the models needs to be defined by comparing simulated patterns of present-day climate with the observed data.

Simple, effective and meaningful metrics, indicators, measures or criteria are required to rank the GCMs across time and space to evolve a subset of models that can be employed for hydrological modelling applications (Randall et al. 2007, Tebaldi & Knutti 2007). These metrics may strengthen the confidence level of outputs of GCMs, as these are the main inputs to the regional climate models or downscaling methods, and the evolved forecasts of hydrologic variables incorporating the impact of climate changes are used in planning mitigation measures. A brief literature review related to the various performance metrics is presented next.

Giorgi & Mearns (2002) introduced the reliability ensemble averaging (REA) method for calculating average, uncertainty range and a measure of reliability of simulated climate changes at the sub-continental scale from ensembles of 9 different atmosphere–ocean general circulation model simulations with 2 emission scenarios. They considered 2 reliability criteria: (1) performance of the model in reproducing present-day climate and (2) convergence of the simulated changes across models. The REA method was applied to mean seasonal temperature and precipitation changes for the late decades of the 21st century, over 22 land regions of the world. Similar studies were reported by Murphy et al. (2004) and Dessai et al. (2005).

Perkins et al. (2007) evaluated 14 GCMs using a skill score approach which is based on their ability to simulate daily rainfall and daily minimum and maximum temperatures for 12 regions of Australia. The evaluation was based on how well each climate model could capture the observed probability density functions for each variable and each region. They concluded that the skill score methodology is useful in assessing which of the climate models should be used by the impact groups. Similar studies were reported by Maxino et al. (2008), Perkins et al. (2009, 2013) and Anandhi et al. (2011). Suppiah et al. (2007) assessed the performance of 23 models with respect to how well they reproduced patterns of seasonal average temperature, mean sea level pressure and rainfall over the Australian continent. They reduced the size of the sample to 15 by rejecting those models which frequently failed to meet certain root mean square error (RMSE) and spatial correlation thresholds across the 4 seasons. Gleckler et al. (2008) developed the model climate performance index (MCPI) and the model variability index (MVI), which are based on the relative error/RMSE methodology. They suggested developing a broad suite of metrics to characterize the model performance from which it may be possible to identify the optimal subsets for various applications. Tang et al. (2008) assessed the model convergence by exploring ensemble mean square, ensemble spread and the ratio of signal-to-noise, and concluded that predictions from different models vary markedly if the model convergence is poor. Mujumdar & Ghosh (2008) focused on modelling GCM and scenario uncertainty using the possibility theory in projecting stream flow of the Mahanadi River in Hirakud, India. Three GCMs with 2 greenhouse emission scenarios were used. They used the C-coefficient, which is based on a cumulative distribution function of the data as a performance measure, and ranked the GCMs by scenario.

Smith & Chandler (2009) adopted a similar approach to that of Suppiah et al. (2007), but restricted their assessment to rainfall only to assess the performance of 22 GCMs, for the case of Australia. Johnson & Sharma (2009) developed a variable convergence score (VCS) method to rank 8 variables based on the coefficient of variation of an ensemble of 9 GCMs for the case study of Australia. They mentioned that this methodology allows for quantitative assessment between different hydro-climatic variables. They concluded that there is no widely accepted metric for assessing climate models as a whole. Macadam et al. (2010) ranked 17 GCMs for each 20 yr period, based on grid box-wise root mean square differences. They used the 'turnover' principle, which relates the percentage of GCMs ranked in the top 'n' for 1 time period that are not among the top 'n' for the following time period. Reshmi Devi & Nagesh Kumar (2010) presented a state-of-the-art review on intercomparisons of GCMs for hydrologic predictability which can be used as the basis for ranking GCMs. Raje & Mujumdar (2010) formulated an uncertainty modelling methodology whereby, in addition to GCM and scenario uncertainty, the uncertainty in the nature of the downscaling relationship was explored with a generalised uncertainty measure using the Dempster-Shafer evidence theory. They used 3 GCMs for 3 scenarios: A2, A1B, B1. The methodology was tested for projecting the monsoon stream flow of the Mahanadi River at the Hirakud Reservoir in Orissa, India. The results showed an increasing probability of extreme, severe and moderate droughts due to climate change. Fordham et al. (2011) made efforts to select a suitable ensemble of GCMs for Australia based on 20 GCMs for 4 seasons individually. Precipitation was considered as a climate variable. Complementary comparison metrics, namely, bias, correlation, RMSE, Reichler-Kim index and Taylor-index, were considered. Ojha et al. (2013) applied the VCS method for the case of India. They used an ensemble of 17 GCMs and ranked 10 variables. The results indicated higher consistency across GCMs for pressure and temperature, and lower consistency for precipitation and related variables. Regional assessments, while broadly consistent with the overall results, indicated low convergence in atmospheric attributes for the north-eastern part of India.

It is evident from the above literature review that in most of the studies, either the error-based approaches (Suppiah et al. 2007, Gleckler et al. 2008, Tang et al. 2008, Smith & Chandler 2009, Macadam et al. 2010, Fordham et al. 2011) or the skill score-based approaches (Perkins et al. 2007, 2009, 2013, Maxino et al. 2008, Anandhi et al. 2011) were considered independently while ranking GCMs. No comprehensive study has been reported wherein error-based and skill score approaches were jointly applied for ranking GCMs. No decision-making approach has been used to date for ranking. In addition, a number of performance metrics, used by various researchers elsewhere, have not been used for Indian climate conditions for assessing their suitability and applicability.

Bearing the above knowledge gap in mind, the objectives of the present study are formulated as follows. (1) Identify performance indicators: 5 performance indicators were chosen, namely, the correlation coefficient (R), normalised RMSE (NRMSE), absolute normalised mean bias error (ANMBE), average absolute relative error (AARE) and skill score (SS), for the climate variable 'precipitation rate'. (2) Develop a suitable methodology for determining the weights of these 5 indicators. The Entropy method was employed to determine the weights of the 5 indicators. (3) Develop a methodology for ranking of GCMs: exploring the applicability of a multicriterion decision making outranking method, viz. PROMETHEE-2 (Preference Ranking Organisation Method of Enrichment Evaluation). (4) Demonstrate the applicability of the Spearman rank correlation (r_S) to evaluate the association between the different ranking patterns. (5) Evaluate the suitability of GCMs for India and for 4 river basins in India (Godavari, Krishna, Mahanadi and Cauvery). (6) Develop a simple but effective methodology for forming GCM ensembles.

We selected a total of 11 GCMs to be ranked which is a subset of the 23 available GCMs used as part of the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (Meehl et al. 2007). The abbreviations of the selected 11 GCMs and the organisations involved are presented in Table 1. These 11 GCMs were evaluated for India (covering 73 grids of 2.5° × 2.5°) using the 5 performance indicators described above. An effort was also made to rank the GCMs for 4 river basins (Godavari, Krishna, Mahanadi and Cauvery) in peninsular India. The Upper Malaprabha catchment in Karnataka, India, was chosen to demonstrate the Entropy and PROMETHEE-2 methods. Fig. 1 presents a flow chart of the present methodology, and Fig. 2 shows the location of the Upper Malaprabha catchment and the 4 river basins.

2. PERFORMANCE INDICATORS

A performance indicator or metric is a quantifiable measure for any GCM to determine how well it simulates the observed data. As mentioned in the lit-

Table 1. Details of GCMs considered and their acronyms

S.No	GCM	Organisation	Acronym
1	BCCR-BCCM 2.0	Bjerknes Centre for Climate Research, Norway	BCCR
2	INGV-ECHAM 4	Istituto Nazionale Di Geofisica E Vulcanologia, Italy	ECHAM
3	GFDL2.0	Geophysical Fluid Dynamic Laboratory, USA	GFDL2.0
4	GFDL2.1	Geophysical Fluid Dynamic Laboratory, USA	GFDL2.1
5	GISS	Goddard Institute for Space Studies, USA	GISS
6	IPSL-CM 4	Institut Pierre Simon Laplace, France	IPSL
7	MIROC3	Centre for Climate Research, Japan	MIROC3
8	MRI-CGCM2	Meteorological Research Institute, Japan	CGCM2
9	NCAR-PCMI	Parallel Climate Models, NCAR, USA	PCMI
10	UKMO-HADCM3	UK Met Office, UK	HADCM3
11	UKMO-HAD GEM1	UK Met Office, UK	HADGEM1

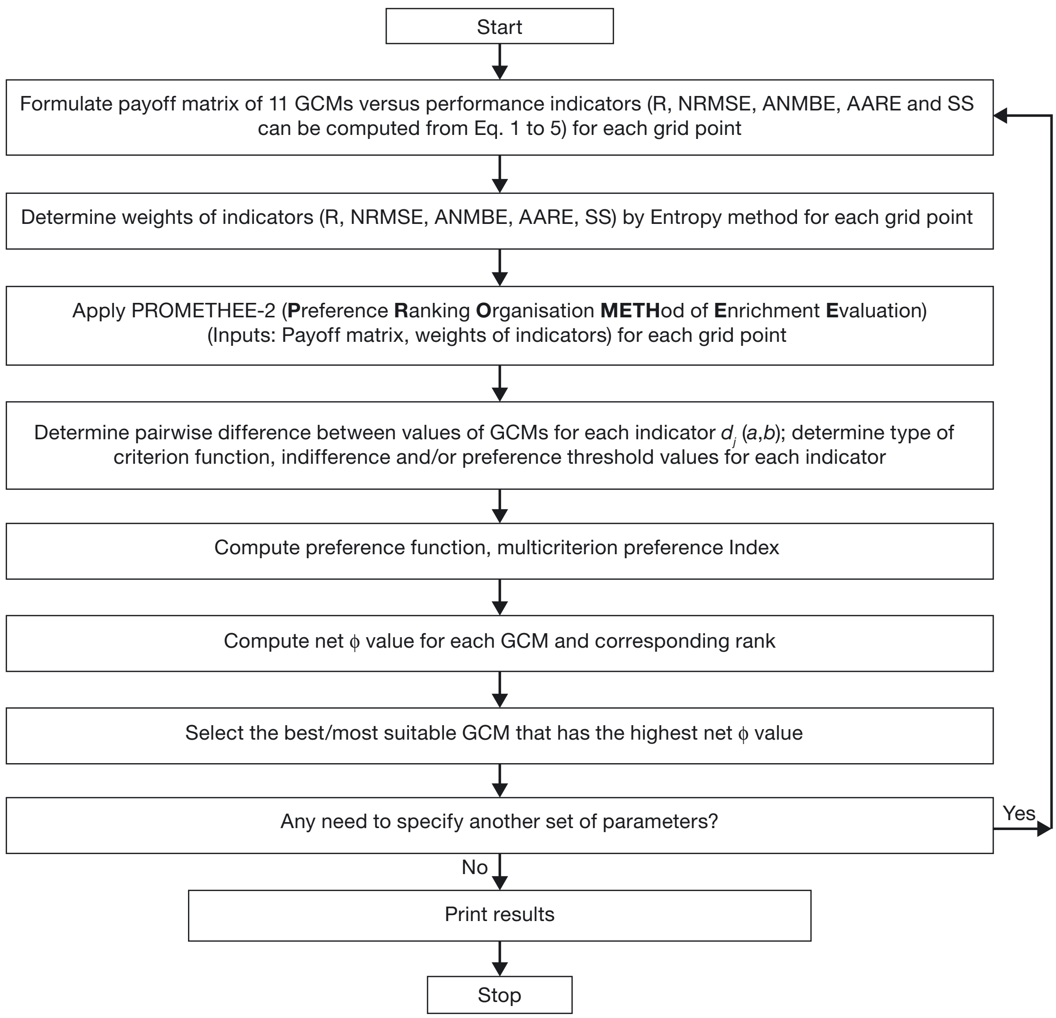

Fig. 1. Methodology for the selection of suitable global climate models (GCMs). Performance indicators are R: correlation coefficient, NRMSE: normalised root mean square error, ANMBE: absolute normalised mean bias error, AARE: average absolute relative error, SS: skill score

erature review above, different researchers have chosen different performance indicators. In the present study, we considered 5 indicators, namely, R, NRMSE, ANMBE, AARE, and SS, among the numerous performance indicators that are currently available (Milton & Arnold 2007, Wilks 2011). Of these 5,

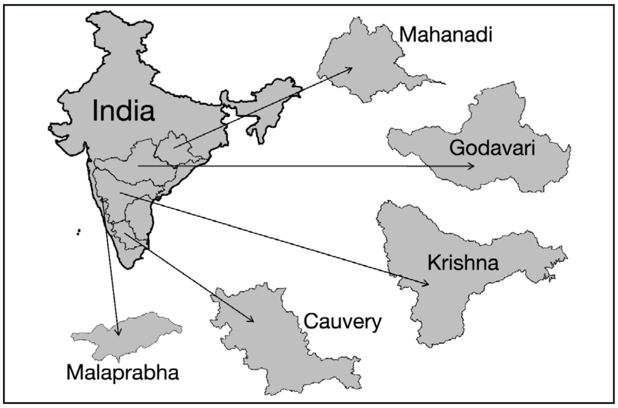

Fig. 2. Location of the Upper Malaprabha catchment and the Godavari, Krishna, Mahanadi and Cauvery River basins in India

NRMSE, ANMBE and AARE are in the error-based category. As there are neither agreed measures of performance nor superiority of any one indicator that can be used to rank models, an attempt was made by employing the 5 performance indicators simultaneously to evaluate the GCMs.

(1) The R provides information on the strength of the linear relationship between the observed and the computed values. R values vary between −1 and +1, where a value close to 1.0 indicates good model performance.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - x_{av})(y_i - y_{av})}{(n-1)s_x s_y} \tag{1}$$

where r_{xy} is the R between observed and projected values x and y; x_i is the observed value; x_{av} is the mean of observed values; y_i is the projected value; y_{av} is the mean of projected values; s_x and s_y are the standard deviations of x and y, respectively; and n is the number of observations.

(2) NRMSE is a measure of the difference between the observed values and the model projected values.

$$\text{NRMSE} = \frac{\left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}(x_i - y_i)^2\right]^{1/2}}{\left(\frac{1}{n}\right)\sum_{i=1}^{n}x_i} \tag{2}$$

Smaller values of NRMSE indicate better performance of the model. Ideally, a value of 0 is preferred.

(3) ANMBE is computed as

$$\text{ANMBE} = \left|\frac{\frac{1}{n}\left(\sum_{i=1}^{n}(y_i - x_i)\right)}{\frac{1}{n}\sum_{i=1}^{n}x_i}\right| \tag{3}$$

Here again, smaller values indicate better performance of the model, and ideally, a value of 0 is preferred.

(4) AARE is the average of the absolute values of relative errors and is expressed as

$$\text{AARE} = \frac{1}{n}\sum_{i=1}^{n}|RE_i| \tag{4}$$

where $RE_i = \frac{(y_i - x_i)}{x_i}$.

Smaller values again indicate better performance of the model, and a value of 0 is preferred.

(5) SS (Perkins et al. 2007) provides a measure of similarity between 2 probability density functions (PDFs), which allows comparison across the entire PDF, and is expressed as

$$\text{SS} = \sum_{i=1}^{nb}\min(f_m, f_o) \tag{5}$$

where nb is the number of bins used to calculate the PDF for a given region, f_m is the frequency of values in the given bin from the chosen GCM, and f_o is the frequency of values in the given bin from the observed data. If a model simulates the observed conditions perfectly, the SS is 1, which is the total sum of the binned values in a given PDF. If a model simulates the observed PDF poorly, the SS will be close to 0.

3. METHODS

(1) The Entropy method was employed to determine the weights of indicators (Pomerol & Romero 2000, Raju & Nagesh Kumar 2010). The weights of indicators for each grid depend on the formulated gridwise payoff matrix, i.e. GCMs versus performance indicator array (see Table 1 for the Upper Malaprabha catchment). The main advantage of the Entropy method is that the weights are determined for each indicator without the intervention of a decision maker, which is expected to reduce the undue bias towards any indicator. An added advantage of this method is that the variation of weights of indicators across the various grid points provides an opportunity for the water resources planner to understand their importance to the outcome. The methodology is based on the amount of information available (measured by its entropy value) and its relationship with the importance of the criterion. Pomerol & Romero (2000) explained the Entropy method in the following steps:

For the given normalised payoff matrix, p_{ij}, (where i is the index for GCMs; j is the index for indicators), entropy E_j for indicator j for the set of GCMs is computed as

$$E_j = -\frac{1}{\ln(N)}\sum_{i=1}^{N} p_{ij}\ln(p_{ij}) \text{ for } j = 1, \ldots, J \tag{6}$$

where $i = 1, \ldots, N$ is the number of GCMs (11 shown in Table 1) and j is the number of indicators (5 shown in Table 1).

Degree of diversification, D_j, of the information provided by the outcomes of indicator j is

$$D_j = 1 - E_j \text{ for } j = 1, \ldots, J \tag{7}$$

Normalised weights of indicators are computed as

$$w_j = \frac{D_j}{\sum_{j=1}^{J} D_j} \tag{8}$$

If the entropy value is high, the uncertainty contained in the criterion vector is high (step 1; Eq. 6), diversification of the information is low (step 2; Eq. 7), and correspondingly the criterion is less important (step 3; Eq. 8).

(2) PROMETHEE-2, a multicriterion decision-making (MCDM) method of outranking nature, was employed to rank the GCMs (Brans et al. 1986, Pomerol & Romero 2000, Raju & Nagesh Kumar 2010). This method uses the preference function. Mathematically, preference function $P_j(a,b)$ is based on the pairwise difference $d_j(a,b)$ between the evaluations $f_j(a)$ and $f_j(b)$ of GCMs a and b for indicator j, the chosen criterion function, the indifference q_j and the preference thresholds p_j. Six types of criterion functions are available: (i) the usual criterion, (ii) a quasi criterion, (iii) a criterion with linear preference and no indifference area, (iv) a level criterion, (v) a criterion with linear preference and indifference area, and (vi) a Gaussian criterion. In this study, the usual criterion function was employed in which even a slight positive difference $d_j(a,b)$ counts while comparing 2 GCMs i.e.

$$P_j(a,b) = \begin{bmatrix} 0 & \text{if} & d_j(a,b) \le 0 \\ 1 & \text{if} & d_j(a,b) > 0 \end{bmatrix} \tag{9}$$

In the other criterion functions, either the indifference parameter q_j (representing the largest difference that is considered negligible in comparing 2 GCMs using that indicator) or the parameter p_j (representing the smallest difference that justifies a strict preference for one of the two GCMs using that indicator) or both are required. It may be difficult to precisely identify these parameters, which may create more uncertainty (Brans et al. 1986). The multicriterion preference index, $\pi(a,b)$, a weighted average of the preference functions $P_j(a,b)$ for all the indicators, is defined as:

$$\pi(a,b) = \frac{\sum_{j=1}^{J} w_j P_j(a,b)}{\sum_{j=1}^{J} w_j} \tag{10}$$

$$\phi^+(a) = \frac{\sum_{N}\pi(a,b)}{(N-1)} \tag{11}$$

$$\phi^-(a) = \frac{\sum_{N}\pi(b,a)}{(N-1)} \tag{12}$$

$$\phi(a) = \phi^+(a) - \phi^-(a) \tag{13}$$

where w_j is the weight assigned to the indicator j; $\phi^+(a)$ is the outranking index of a in the GCM set N; $\phi^-(a)$ is the outranked index of a in the GCM set N; $\phi(a)$ is the net ranking of a in the GCM set N, and J is the number of indicators. The GCM having the highest $\phi(a)$ value is considered to be the most suitable GCM.

(3) The Spearman rank correlation, r_S, is useful to determine the measure of association between the ranks achieved in 2 different scenarios. If U_a and V_a denote the ranks achieved by the above situation(s) for the same GCM a, then R is defined as (Gibbons 1971):

$$R = 1 - \frac{6\sum_{a=1}^{N} {D_a}^2}{N(N^2-1)} \tag{14}$$

where D_a is the difference between ranks U_a and V_a achieved by the same GCM a, and N is the number of GCMs. R values vary between −1 and +1.

(4) Nested bias correction (NBC; Johnson & Sharma 2011, 2012) is an approach that compensates for some of the shortcomings of GCM-predicted rainfall values, corrects for systematic biases of GCM outputs (e.g. mean, standard deviation, lag-one correlation) at multiple timescales, and allows the use of GCM outputs directly in hydrologic studies. When combined with spatial disaggregation, bias correction techniques can provide model inputs at a range of scales suitable for hydrologic studies (Hashino et al. 2007, Johnson & Sharma 2009, 2011, 2012, Mehrotra & Sharma 2010, Mehrotra et al. 2013). In the present study, we followed the approach suggested by Johnson & Sharma (2011, 2012), with the exception that bias in the lag-one autocorrelation statistic at the various time scales was neglected.

The NBC approach represents a nested procedure which addresses bias across pre-specified multiple timescales which are as follows: Denoting a variable for month i in year k as $y_{i,k}$:

(i) Standardisation to create $y'_{i,k}$ by subtracting the model monthly mean ($\mu_{\text{mod},i}$) and dividing by the standard deviation ($\sigma_{\text{mod},i}$) for that month as shown in Eq. (15):

$$Y'_{i,k} = \frac{y_{i,k} - \mu_{\text{mod},i}}{\sigma_{\text{mod},i}} \tag{15}$$

(ii) Interposition of the mean ($\mu_{obs,i}$) and standard deviation ($\sigma_{obs,i}$) of reanalysis data to create a transformed time series $y^*_{i,k}$ at the monthly level:

$$y^*_{i,k} = y'_{i,k}\sigma_{obs,i} + \mu_{obs,i} \quad (16)$$

(iii) Aggregation of transformed monthly series ($y^*_{i,k}$) into the annual scale z_k. The standardisation and transformation steps are repeated at the annual time step.

(iv) Transformation of the annual time series to z*k which exhibits the mean and standard deviation in the recorded annual data.

Subsequent to the above steps, the raw GCM simulation at the monthly time step is transformed by the NBC to:

$$Y_{i,k} = y_{i,k}\left[\frac{y^*_{i,k}}{y_{i,k}}\right]\left[\frac{z^*_k}{z_k}\right] \quad (17)$$

where $Y_{i,k}$ represents the NBC-transformed variable. Using the transformation of Eq. (17), the corrections at monthly and annual scales can be applied to the monthly time series at the same time to create a 1-step correction (Srikanthan 2009). In Eq. (17), $\left[\frac{y^*_{i,k}}{y_{i,k}}\right]\left[\frac{z^*_k}{z_k}\right]$ is a weighing factor, i.e. the ratio of the monthly corrected value to the raw GCM value for month i and year k, multiplied by the ratio of the yearly corrected value to the aggregated GCM rainfall for year k. The above equations were used to transform the GCM simulations for the current climate.

4. DATA AND RESULTS

4.1. Data

The National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP/NCAR) reanalysis precipitation rate data were used as a proxy to the observed data (Kalnay et al. 1996, Kistler et al. 2001). The NCAP/NCAR monthly data at a grid resolution of 2.5° × 2.5° for the period 1950–1999 (50 yr and 12 mo = 600 monthly data) were used for the analysis. While there are known uncertainties in this dataset due to changes in the observing systems, the modelling deficiencies and human-induced errors in assimilation, the uncertainty in NCEP/NCAR datasets varies across variables (Kistler et al. 2001, Ojha et al. 2013). The chosen GCMs use grid cells of different sizes with non-uniform placement. The outputs of all the GCMs were interpolated to a common grid (2.5° × 2.5°), since these were to be compared against the reanalysis data. These outputs were interpolated using the weighted mean of the 4 nearest values, with weights being assigned based on the inverse square distance relationship (Johnson & Sharma 2009). To smooth the changes in the variable projections, a spatial filtering process was adopted wherein each grid cell value was replaced by the average of the 9 neighbouring grid cells. Bias correction as discussed in Section 3 recognizes that the GCM outputs for the current climate may have biases when compared to the observed data. The correction process adjusts the GCM outputs to match the mean and variance of the recorded historical data. In this case, grid cell averaged data were corrected for bias. Bias correction was carried out for each pixel on a monthly basis.

In the present study, gridwise (various latitude and longitude combinations resulting in 73 grids) studies were explored to assess how the individual GCMs should be prioritised with reference to the 5 chosen indicators using PROMETHEE-2. However, the average ranking approach (Bui 1987) was employed to aggregate the perspective for all of India (for all 73 grid points), as well as for the 4 chosen river basins (Godavari, 22 grids; Krishna, 17; Mahanadi, 15; Cauvery, 8).

4.2. Case study of upper Malaprabha catchment

4.2.1. Analysis of indicators

For demonstration of the Entropy and PROMETHEE-2 methods, we chose the Upper Malaprabha catchment, Karnataka State, India, located between 15° 00′ and 16° 12′ N and between 74° 14′ and 76° 05′ E. The catchment area of the river up to the dam site (2564 km^2) was considered in this study.

Table 1 presents the values obtained for the 5 indicators (R, NRMSE, ANMBE, AARE and SS, Eqs. 1–5) for the 11 chosen GCMs. Minimum or 0 error is desirable in the case of NRMSE, ANMBE and AARE, whereas an ideal value of 1 is desirable for R and SS. A negative sign is placed before the values related to NRMSE, ANMBE and AARE to make them maximisation type and to be compatible with R and SS in Table 1. Table 1 shows that in the case of R (Eq. 1), MIROC3 was correlated well with the observed data with a value of 0.8416, whereas a minimum R was observed for PCMI with a value of 0.3553. R was computed from a temporal perspective (600 data sets) for each grid point. This aspect is extended spatially for all 73 grid points covering India.

Table 2. Values of 5 performance indicators obtained for the 11 chosen global climate models (GCMs) for the Upper Malaprabha catchment, India. GCMs are described in the 'Introduction'. Performance indicators are R: correlation coefficient, NRMSE: normalised root mean square error, ANMBE: absolute normalised mean bias error, AARE: average absolute relative error, SS: skill score

GCM	R	NRMSE	ANMBE	AARE	SS
BCCR	0.7751	−0.7960	−0.2744	−1.7127	0.7717
ECHAM	0.7866	−0.7573	−0.1619	−1.8639	0.6833
GFDL2.0	0.7868	−0.8286	−0.4157	−0.8080	0.8150
GFDL2.1	0.7395	−0.7871	−0.1551	−1.2731	0.8350
GISS	0.8275	−0.8221	−0.4786	−0.7539	0.7783
IPSL	0.4740	−1.2539	−0.7082	−1.0124	0.6583
MIROC3	0.8416	−0.6224	−0.0613	−1.3811	0.8567
CGCM2	0.7708	−0.9386	−0.4985	−0.6556	0.7550
PCMI	0.3553	−1.1779	−0.4899	−1.6149	0.6283
HADCM3	0.8018	−0.8793	−0.5092	−0.8002	0.8100
HADGEM1	0.8064	−0.9422	−0.5686	−0.7010	0.7883
Total entropy (Eq. 6)	0.9896	0.9922	0.9416	0.9719	0.9982
Degree of diversification (Eq. 7)	0.0104	0.0078	0.0584	0.0281	0.0018
Normalised weights (Eq. 8)	0.0976	0.0729	0.5481	0.2640	0.0174

Similar trends were observed for SS (Eq. 5). In this case, MIROC3 shows 85.67 % similarity with the observed PDFs, whereas PCMI shows only 62.83 % similarity. In the case of the NRMSE indicator (Eq. 2), MIROC3 is the preferred GCM, with an NRMSE value of −0.6224, whereas IPSL is the least preferred (NRMSE = −1.2539). For ANMBE (Eq. 3) as well, MIROC3 is the preferred GCM (−0.0613) whereas IPSL is the least preferred (−0.7082). For AARE (Eq. 4), CGCM2 and ECHAM are the most and least preferred GCMs, respectively. The above analysis indicates that each indicator responds differently for various GCMs. In the present study, an effort was made to explore all 5 indicators simultaneously to assess their applicability (and their relative contribution) while ranking the GCMs.

4.2.2. Application of the Entropy method

All indicator values in Table 2 were normalised using $\frac{f_j(i)}{\sum_{i=1}^{N} f_j(i)}$ to ensure that the criterion with a larger range would not dominate the criterion with a smaller range, and to make them compatible with the requirements of the Entropy method. Table 2 also presents total entropy E_j (Eq. 6), the degree of diversification D_j (Eq. 7) and the normalised weight w_j of each indicator (Eq. 8). Among the 5 indicators, ANMBE has the highest importance value (54.81 %), which means that its effect on ranking of GCMs is significant. The total contribution of R, NRMSE and SS is less than 20 %, whereas AARE contributes 26.4 %.

4.2.3. Application of PROMETHEE-2

The usual criterion function was employed (Brans et al. 1986). In the case of the usual criterion (Eq. 9), the elements of the preference function matrix are either 0 or 1 (Raju & Nagesh Kumar 2010). For example, the pairwise difference of values in Table 2, between BCCR and GFDL2.0, for R is 0.7751 − 0.7868 = −0.0117 and so the corresponding value of the preference function under the usual criterion (Eq. 9) is 0 (as −0.0117 < 0). Conversely, the pairwise difference between GFDL2.0 and BCCR for R is 0.0117 and the corresponding value of the preference function is 1 (as 0.0117 > 0). In this way, the pairwise preference function values are computed for each indicator. Weights estimated by the Entropy method, hereafter termed the varying weight (VW) scenario, were 0.0976, 0.0729, 0.5481, 0.2640 and 0.0174 (for R, NRMSE, ANMBE, AARE and SS, respectively), and were used to compute the weighted preference function values, i.e. the multicriterion preference index (Eq. 10). These values are presented in Table 3. All diagonal values are 0 because of the comparison of a GCM with itself. Table 4 presents ϕ^+, ϕ^- and ϕ values corresponding to each GCM and the ranking pattern of the GCMs. For example, in the case of MIROC3, the average of the values in the row corresponding to MIROC3 (Eq. 11) in Table 3, i.e. the sum of all elements in that row/(number of elements − 1) = (1 + 1 +

Table 3. Multicriterion preference index values using entropy-based weights (varying weight scenario) for the Upper Malaprabha catchment for 11 global climate models

	BCCR	ECHAM	GFDL2.0	GFDL2.1	GISS	IPSL	MIROC3	CGCM2	PCMI	HADCM3	HADGEM1
BCCR	0.0000	0.2814	0.6210	0.0976	0.6210	0.7360	0.0000	0.7360	0.7360	0.6210	0.6210
ECHAM	0.7186	0.0000	0.6210	0.1705	0.6210	0.7360	0.0000	0.7186	0.7360	0.6210	0.6210
GFDL2.0	0.3790	0.3790	0.0000	0.3616	0.5655	1.0000	0.2640	0.7360	1.0000	0.6384	0.6384
GFDL2.1	0.9024	0.8295	0.6384	0.0000	0.6384	0.7360	0.2640	0.6384	1.0000	0.6384	0.6384
GISS	0.3790	0.3790	0.4345	0.3616	0.0000	1.0000	0.2640	0.7360	1.0000	0.9826	0.7186
IPSL	0.2640	0.2640	0.0000	0.2640	0.0000	0.0000	0.2640	0.0000	0.3790	0.0000	0.0000
MIROC3	1.0000	1.0000	0.7360	0.7360	0.7360	0.7360	0.0000	0.7360	1.0000	0.7360	0.7360
CGCM2	0.2640	0.2814	0.2640	0.3616	0.2640	1.0000	0.2640	0.0000	0.4519	0.8121	0.8850
PCMI	0.2640	0.2640	0.0000	0.0000	0.0000	0.6210	0.0000	0.5481	0.0000	0.5481	0.5481
HADCM3	0.3790	0.3790	0.3616	0.3616	0.0174	1.0000	0.2640	0.1879	0.4519	0.0000	0.6384
HADGEM1	0.3790	0.3790	0.3616	0.3616	0.2814	1.0000	0.2640	0.1150	0.4519	0.3616	0.0000

Table 4. Values of ϕ^+, ϕ^-, ϕ (defined in Section 3) and ranks of global climate models (GCMs) for the Upper Malaprabha catchment

GCM	Varying weight scenario				Equal weight scenario			
	ϕ^+	ϕ^-	ϕ	Rank	ϕ^+	ϕ^-	ϕ	Rank
BCCR	0.5071	0.4929	0.0142	6	0.4600	0.5400	−0.0800	8
ECHAM	0.5564	0.4436	0.1127	5	0.4800	0.5200	−0.0400	7
GFDL2.0	0.5962	0.4038	0.1924	4	0.6200	0.3800	0.2400	4
GFDL2.1	0.6924	0.3076	0.3848	2	0.6400	0.3600	0.2800	3
GISS	0.6255	0.3745	0.2511	3	0.6600	0.3400	0.3200	2
IPSL	0.1435	0.8565	−0.7130	11	0.1400	0.8600	−0.7200	11
MIROC3	0.8152	0.1848	0.6304	1	0.8600	0.1400	0.7200	1
CGCM2	0.4848	0.5152	−0.0304	7	0.4400	0.5600	−0.1200	9
PCMI	0.2793	0.7207	−0.4413	10	0.1400	0.8600	−0.7200	10
HADCM3	0.4041	0.5959	−0.1918	8	0.5400	0.4600	0.0800	5
HADGEM1	0.3955	0.6045	−0.2090	9	0.5200	0.4800	0.0400	6

0.736 + 0.736 + 0.736 + 0.736 + 0.736 + 1 + 0.736 + 0.736)/10 = 0.8152 (ϕ^+). Similarly, the average of the values in the column corresponding to MIROC3 (Eq. 12) in Table 3, i.e. the sum of all elements of that column/(number of elements − 1) = (0 + 7×0.264)/10 = 0.1848 (ϕ^-). The value of ϕ (Eq. 13) is the difference between ϕ^+ and ϕ^-, which is 0.6304 in this example. ϕ values of other GCMs are also presented in Table 4. The GCM having the highest ϕ value is considered best. Table 4 shows that MIROC3 has the highest ϕ value of 0.6304 and is therefore considered the best model (Rank 1), followed by GFDL2.1 with a ϕ value of 0.3848. IPSL is ranked lowest due to its low ϕ value of −0.7130. MIROC3, GFDL2.1 and GISS occupy the first 3 positions.

Similar efforts were also made by adopting equal weights (hereafter termed the equal weight, EW, scenario), i.e. 0.2 each to the 5 indicators. Table 5 presents the corresponding multicriterion preference index values. Table 4 presents the resultant values of ϕ^+, ϕ^-, ϕ and ranking pattern corresponding to the EW scenario. In this scenario, MIROC3, GISS and GFDL2.1 occupy the first 3 positions with ϕ values of 0.72, 0.32 and 0.28. The lowest positions are occupied by IPSL and PCMI, with a ϕ value of −0.72.

4.2.4. Application of the Spearman rank correlation

Ranks obtained for 11 GCMs using PROMETHEE-2 in the VW scenario and EW scenario are shown in Table 4. For computing the Spearman rank correlation, r_S (Eq. 14), the D_a values are −2, −2, 0, −1, 1, 0, 0, −2, 0, 3 and 3, and the D_a^2 values are 4, 4, 0, 1, 1, 0, 0, 4, 0, 9 and 9, respectively. The $\sum D_a^2$ value is 32, and the corresponding correlation coefficient value is 0.8545.

MIROC3, GISS and GFDL2.1 occupied the first 3 positions in both the scenarios. Hence these 3 GCMs were explored further for downscaling/hydrological modelling applications.

4.3. Case study of India

4.3.1. Application of the Entropy method

A similar process to that described in Section 4.2.2 was repeated for all 73 grid points. We observed (results not presented) that weights varied by indicator and by grid. R varied from 0.25 to 89.72%; NRMSE: 0.47–24.80%; ANMBE: 5.37–82.86%; AARE: 0.72–44.01% and SS: 0.24–10.61%. Of the 73 evaluations, the first position was occupied 63 times by ANMBE, 8

Table 5. Multicriterion preference index values for the Upper Malaprabha catchment (equal weight scenario) for 11 global climate models

	BCCR	ECHAM	GFDL2.0	GFDL2.1	GISS	IPSL	MIROC3	CGCM2	PCMI	HADCM3	HADGEM1
BCCR	0.0	0.4	0.4	0.2	0.4	0.8	0.0	0.8	0.8	0.4	0.4000
ECHAM	0.6	0.0	0.4	0.4	0.4	0.8	0.0	0.6	0.8	0.4	0.4000
GFDL2.0	0.6	0.6	0.0	0.4	0.4	1.0	0.2	0.8	1.0	0.6	0.6000
GFDL2.1	0.8	0.6	0.6	0.0	0.6	0.8	0.2	0.6	1.0	0.6	0.6000
GISS	0.6	0.6	0.6	0.4	0.0	1.0	0.2	0.8	1.0	0.8	0.6000
IPSL	0.2	0.2	0.0	0.2	0.0	0.0	0.2	0.0	0.6	0.0	0.0000
MIROC3	1.0	1.0	0.8	0.8	0.8	0.8	0.0	0.8	1.0	0.8	0.8000
CGCM2	0.2	0.4	0.2	0.4	0.2	1.0	0.2	0.0	0.8	0.4	0.6000
PCMI	0.2	0.2	0.0	0.0	0.0	0.4	0.0	0.2	0.0	0.2	0.2000
HADCM3	0.6	0.6	0.4	0.4	0.2	1.0	0.2	0.6	0.8	0.0	0.6000
HADGEM1	0.6	0.6	0.4	0.4	0.4	1.0	0.2	0.4	0.8	0.4	0.0000

times by R and 2 times by AARE. NRMSE and SS never attained the first position. Of the 5 indicators, ANMBE is thus the preferred one. This may be due to its simple and effective deviation principle. In further studies, the 2 indicators NRMSE and SS could be eliminated or used as complimentary indicators so that only the relevant indicators need to be analysed. However, in such a situation, weights of indicators will significantly change and accordingly, the ranking of GCMs may also be affected. We also attempted to analyse the weight distribution in different ranges over the 73 grid points obtained by the Entropy method (Table 6). Table 6 shows that the number of grid points within 10% weight range are 38, 67, 2, 18 and 72 in the case of R, NRMSE, ANMBE, AARE and SS, respectively, although ANMBE and R are spread out across almost all weight ranges.

Table 6. Distribution of weights over 73 grid points obtained by the Entropy method (varying weight scenario) into various ranges. Performance indicators are R: correlation coefficient, NRMSE: normalised root mean square error, ANMBE: absolute normalised mean bias error, AARE: average absolute relative error, SS: skill score

Weight range (%)	R	NRMSE	ANMBE	AARE	SS
≤10	38	67	2	18	72
>10 and ≤20	16	4	–	21	1
>20 and ≤30	8	2	3	25	–
>30 and ≤40	3	–	4	8	–
>40 and ≤50	3	–	8	1	–
>50 and ≤60	2	–	16	–	–
>60 and ≤70	1	–	26	–	–
>70 and ≤80	–	–	11	–	–
>80 and ≤90	2	–	3	–	–
>90 and ≤100	–	–	–	–	–

To our knowledge, this is the first application of the Entropy method for determining the weights of performance indicators that can be used to rank GCMs.

4.3.2. Application of PROMETHEE-2

A similar process of application of PROMETHEE-2 as described in Section 4.2.3. was repeated for the 73 grid points, and an analysis of the results is presented in Table 7. Table 7 shows that in the VW scenario, GFDL2.0, MIROC3, BCCR and HADCM3 occupied the first position 17, 13, 10 and 8 times (totalling 48 among a possible 73). In the EW scenario, GFDL2.0, MIROC3, HADCM3 and BCCR occupied the first position 20, 15, 10 and 9 times (totalling 54). These 4 GCMs alone thus occupied the first position 65.75 and 73.97% of 73 grid points in the VW and EW sce-

Table 7. Number of times each global climate model (GCM) occupied the first and second positions in varying weight (VW) and equal weight (EW) scenarios

GCM	VW		EW	
	First	Second	First	Second
BCCR	10	7	9	7
ECHAM	5	11	3	11
GFDL2.0	17	7	20	15
GFDL2.1	4	6	2	3
GISS	2	1	–	4
IPSL	–	2	2	3
MIROC3	13	9	15	12
CGCM2	6	6	8	4
PCMI	4	4	3	1
HADCM3	8	12	10	9
HADGEM1	4	8	1	4

narios, respectively. The remaining 7 GCMs occupied the first position only 34.25 and 26.03%.

In the VW scenario, HADCM3, ECHAM, MIROC3 and HADGEM1 occupied the second position 12, 11, 9 and 8 times (totalling 40), and in the EW scenario, GFDL2.0, MIROC3, ECHAM and HADCM3 occupied the second position 15, 12, 11 and 9 times (totalling 47). These 4 GCMs alone occupied the second position 54.79 and 64.38% in the VW and EW scenarios, respectively, whereas the remaining 7 GCMs occupied this position only 45.21 and 35.62%. Fig. 3 presents the spatial distribution of GCMs occupying the first and second positions in the VW and EW scenarios.

4.3.3. Spearman rank correlation

A similar process as described in Section 4.2.4 was repeated for the 73 grid points. The average R value over the 73 grid points is 0.7876, representing an important relationship (Raju & Nagesh Kumar 2010), whereas the minimum and maximum R values are

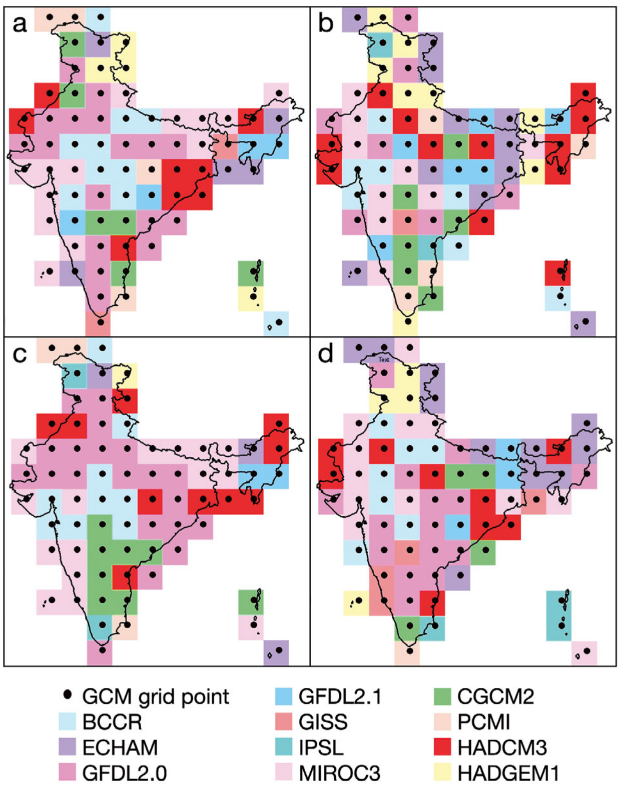

Fig. 3. Spatial distribution of global climate models (GCMs) occupying (a) the first position in the varying weight (VW) scenario, (b) the second position in the VW scenario, (c) the first position in the equal weight (EW) scenario and (d) the second position in the EW scenario

0.0818 and 0.9909. A low value of R occurred at grid point 7.5°N × 77.5°E. On the other hand, a high value of R was observed at 22.5° N × 75° E. Overall, 67 out of 73 (91.8%) grids points have R values >0.5, indicating substantial relationships (Raju & Nagesh Kumar 2010).

4.3.4. Ranking of GCMs for India

An effort was also made to rank the GCMs for all of India (over 73 grid points) for the precipitation rate. The average ranking method (i.e. average of all ranks corresponding to each GCM over 73 grid points) as suggested by Bui (1987) was used for this purpose. In the VW scenario, HADCM3, GFDL2.0 and MIROC3 occupied the first 3 positions, whereas in the EW scenario, GFDL2.0, MIROC3 and HADCM3 occupied the first 3 positions. In this regard, 3 GCMs (HADCM3, GFDL2.0, MIROC3) occur in both scenarios based on average rankings. PCMI, GISS and IPSL occupied the last 3 positions in both VW and EW scenarios.

From the extensive studies made (i.e. computing performance indicators for all GCMs for 73 grid points, analysing suitability of GCMs for 73 grid points individually for 2 weight scenarios, and computing r_s values between 2 weight scenarios for each grid point), we infer that no single GCM can be recommended for India as a whole, and so an ensemble of GCMs must be determined. The GCMs which occupied (1) the first 3 positions in average ranking perspective, i.e. HADCM3, GFDL2.0 and MIROC3, (2) the first position (48 of 73 grid points in VW and 54 of 73 grid points in EW perspective), namely, GFDL2.0, MIROC3, HADCM3 and BCCR, and (3) the second position (40 of 72 grid points in VW and 47 of 73 grid points in EW), namely, GFDL2.0, HADGEM1, HADCM3, ECHAM and MIROC3 (Section 4.3.2), were taken into consideration. Accordingly, the ensemble of GFDL2.0, MIROC3, BCCR, HADCM3, ECHAM and HADGEM1 among the available 11 GCMs is suggested for India for precipitation rate. However, further analysis, such as evaluating GCMs for more climatic variables (e.g. temperature), is required before adopting these models for downscaling.

4.4. Case study of selected river basins in India

The suitability of the GCMs for 4 river basins (Godavari, Krishna, Mahanadi and Cauvery) was also evaluated. Locations of these basins are presented in

Fig. 2. Details of these basins are available in the Integrated Hydrological Data Book (Central Water Commission 2012).

Table 8 presents the suitability of GCMs in VW and EW scenarios in terms of the first and second positions among the chosen grids (variable between river basins). The GCMs which occupied the first 3 positions in VW and EW scenarios (using average ranking) are also presented. These statistics indicate that for each basin, no single GCM can be recommended; thus an ensemble of GCMs is suggested. GCMs which occupy the first 3 positions (average ranking approach) and GCMs which occupy the first and second positions in both the VW and EW approaches (analogous to Section 4.3.2) are considered for determining the suitable ensembles.

The composition of the ensembles for India and the 4 river basins is presented in Table 9, which shows that 9 out of 11 GCMs are covered in the 5 ensembles leaving out IPSL and PCMI. GFDL2.0 is present for all 5 regions followed by CGCM2 and HADCM3 (4 times each). BCCR, GFDL2.0, MIROC3 and HADCM3 are present both for India and for the Godavari River basin. Similarly, BCCR, GFDL 2.0 and MIROC3 are present both for India and for the Krishna River basin; ECHAM, GFDL2.0 and HADCM3 are present both for India and the Mahanadi River basin, and GFDL2.0 and HADCM3 are present both for India and the Cauvery River basin.

It is interesting to note that even though the 4 river basins form part of India, different ensembles are proposed for the individual basins. However, if only a single GCM is to be used, GFDL2.0 can be considered.

Table 8. Suitability of global climate models (GCMs) for the selected river basins based on their positions in the varying weight (VW) and equal weight (EW) scenarios. The numbers of occurrences of GCMs are given in parentheses. GCMs appearing <2 times are not presented

River basin	Suitability (first position)	Suitability (second position)	First 3 positions for (average ranking)	Suggested ensemble both scenarios
Godavari				
VW	GFDL2.0(5) MIROC3 (5) BCCR (4) HADCM3 (3) GFDL2.1(2) CGCM2 (2)	BCCR (5) MIROC3 (4) GFDL2.1(3) CGCM2(3) ECHAM (2) GFDL2.0 (2)	GFDL2.0 BCCR HADCM3	GFDL2.0, MIROC3, BCCR, HADCM3, CGCM2, GFDL2.1
EW	BCCR (6) GFDL2.0(5) CGCM2(5) MIROC3(4) HADCM3(2)	GFDL2.0 (9) BCCR (3) MIROC3 (3) GISS (2) HADCM3 (2)	GFDL2.0 HADCM3 BCCR	
Krishna				
VW	GFDL2.0 (5) MIROC3 (3) CGCM2 (3) BCCR (2) GFDL2.1 (2)	MIROC3 (4) CGCM2 (4) BCCR (3) GFDL2.0 (2)	GFDL2.0 BCCR CGCM2	GFDL2.0, BCCR, CGCM2, MIROC3, GISS
EW	CGCM2 (7) MIROC3 (4) BCCR (3) GFDL2.0 (2)	GFDL2.0 (8) GISS (3) BCCR (2)	GFDL2.0, CGCM2, BCCR	
Mahanadi				
VW	GFDL2.0 (5) HADCM3 (4) BCCR (2)	ECHAM (4) HADCM3 (3) GFDL2.0 (2) GFDL2.1 (2) CGCM2 (2)	GFDL2.0 HADCM3 GFDL2.1	GFDL2.0, HADCM3, GFDL2.1, CGCM2, ECHAM
EW	GFDL2.0 (8) BCCR (2) CGCM2 (2) HADCM3 (2)	GFDL 2.0 (5) HADCM3 (4) CGCM2 (3) GFDL 2.1 (2)	GFDL 2.0 HADCM3 CGCM2	
Cauvery				
VW	GFDL2.0 (3)	CGCM2 (3) PCMI (2)	GFDL2.0 CGCM2 GISS	GFDL2.0, CGCM2, GISS, HADCM3
EW	CGCM2 (3) MIROC3 (2)	GFDL2.0 (3) GISS (2)	GFDL2.0 CGCM2 HADCM3	

5. SUMMARY AND CONCLUSIONS

We used 5 performance indicators (R, NRMSE, ANMBE, AARE and SS) to rank 11 GCMs (BCCR, ECHAM 4, GFDL2.0, GFDL2.1, GISS, IPSL, MIROC3, CGCM2, PCMI, HADCM3 and HADGEM1) for India for a given climate variable, viz. precipitation rate. The Entropy method was employed to determine the weights of the indicators, and an outranking-based multicriterion decision-making (MCDM) method,

Table 9. Composition of the ensembles suggested for all of India and for each of 4 river basins. GCM: global climate model

	BCCR	ECHAM	GFDL2.0	GFDL2.1	GISS	IPSL	MIROC3	CGCM2	PCMI	HADCM3	HADGEM1	Total number of GCMs
India	✓	✓	✓				✓			✓	✓	6
Godavari River basin	✓		✓	✓			✓	✓		✓		6
Krishna River basin	✓		✓		✓		✓	✓				5
Mahanadi River basin		✓	✓	✓				✓		✓		5
Cauvery River basin			✓		✓			✓		✓		4
Number of occurrences	3	2	5	2	2		3	4		4	1	

PROMETHEE-2, was used to rank the 11 GCMs. Our study provided an opportunity to assess the relevance of the Entropy and PROMETHEE-2 methods. To our knowledge, this is the first MCDM application for India, along with a simple but effective weight estimation procedure, i.e. the Entropy method. The resulting ranking patterns may change with the addition of more indicators, more variables, different GCMs and changes in indicator weights. However, the same methodology can be used in similar situations.

The following conclusions can be drawn from our study.

(1) A quantitative evaluation of performance indicators for each GCM is useful and provides an opportunity for improved evaluation of the GCMs.

(2) Weights obtained by the Entropy method over 73 grid points vary from 0.25 to 89.72% for R; from 0.47 to 24.80% for NRMSE; from 5.37 to 82.86% for ANMBE; from 0.72 to 44.01% for AARE; and from 0.24 to 10.61% for SS.

(3) Out of the 73 cases, the first position was occupied 63 times by ANMBE, 8 times by R and 2 times by AARE, whereas NRMSE and SS never attained the first position. Thus ANMBE is considered to be the preferred indicator.

(4) Among the 5 indicators, ANMBE was given high importance (0.5481) for the case study of the Upper Malaprabha catchment, which means that its effect on ranking of GCMs is significant. The combined contribution of R, NMSE and SS was <0.20.

(5) MIROC3, GISS and GFDL2.1 occupied the first 3 positions in the VW and EW scenarios for the Upper Malaprabha catchment. The Spearman rank correlation coefficient, r_S, was found to be a good indicator to assess the association between the ranking patterns obtained with the VW and EW scenarios. The average R value was 0.7876, whereas minimum and maximum R values over 73 grid points were 0.0818 and 0.9909. Low values of R are mainly due to considerable differences in ranking between VW and EW scenarios, and this aspect may be analysed in depth with more data in future analyses.

(6) As no single GCM can be recommended, we recommend the ensemble of GFDL2.0, MIROC3, BCCR, HADCM3, ECHAM and HADGEM1 for India as a whole.

(7) With regard to individual river basins, the ensemble of GFDL2.0, MIROC3, BCCR, HADCM3, CGCM2 and GFDL2.1 is suggested for the Godavari River basin; GFDL2.0, BCCR, CGCM2, MIROC3 and GISS for the Krishna River basin; GFDL2.0, HADCM3, GFDL2.1, CGCM2, and ECHAM for the Mahanadi River basin; and GFDL2.0, CGCM2, GISS and HADCM3 for the Cauvery River basin.

Acknowledgements. We thank Richa Ojha, Research Scholar, School of Civil Engineering, Purdue University, USA, and H. R. Shwetha (Research Scholar, Department of Civil Engineering, Indian Institute of Science, Bangalore, India) for their valuable inputs. This work was supported by Council of Scientific and Industrial Research, New Delhi, through project no. 23(0023)/12/EMR-II dated 15.10.2012. We acknowledge the modelling groups for making their simulations available for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the CMIP3 model output and the WCRP's Working Group on Coupled Modelling (WGCM) for organising the model data analysis activity.

LITERATURE CITED

Anandhi A, Frei A, Pradhanang SM, Zion MS, Pierson DC, Schneiderman EM (2011) AR4 climate model performance in simulating snow water equivalent over Catskill Mountain watersheds, New York, USA. Hydrol Process 25:3302–3311

Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (2008) Climate change and water. Technical Paper VI. IPCC, Geneva

Brans JP, Vincke Ph, Mareschal B (1986) How to select and how to rank projects: the PROMETHEE method. Eur J Oper Res 24:228–238

Bui TX (1987) Coop: a group decision support system for cooperative multiple criteria group decision making. Springer-Verlag, Berlin

Central Water Commission (2012) Integrated Hydrological Data Book: non-classified river basins. Central Water Commission, New Delhi

Dessai S, Lu X, Hulme M (2005) Limited sensitivity analysis of regional climate change probabilities for the 21st century. J Geophys Res 110:D19108, doi:10.1029/2005JD005919

Fordham DA, Wigley TL, Brook BW (2011) Multi-model climate projections for biodiversity risk assessments. Ecol Appl 21:3316–3330

Gibbons JD (1971) Nonparametric statistical inference, McGraw-Hill, New York, NY

Giorgi F, Mearns LO (2002) Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the reliability ensemble averaging (REA) method. J Clim 15:1141–1158

Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. J Geophys Res 113:D06104, doi:10.1029/2007JD008972

Hashino T, Bradley AA, Schwartz SS (2007) Evaluation of bias-correction methods for ensemble streamflow forecasts. Hydrol Earth Syst Sci 11:939–950

Johnson FM, Sharma A (2009) Measurement of GCM skill in predicting variables relevant for hydro climatological assessments. J Clim 22:4373–4382

Johnson FM, Sharma A (2011) Accounting for interannual variability: a comparison of options for water resources climate change impact assessments. Water Resour Res 47:W04508, doi:10.1029/2010WR009272

Johnson F, Sharma A (2012) A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. Water Resour Res 48:W01504, doi:10.1029/2011WR010464

Kalnay E, Kanamitsu M, Kistler R, Collins W and others (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bull Am Meteorol Soc 77:437–471

Kistler R, Kalnay E, Collins W, Saha S and others (2001) The NCEP–NCAR 50-Year Reanalysis: monthly means CDROM and documentation. Bull Am Meteorol Soc 82: 247–267

Knutti R, Abramowitz G, Collins M, Eyring V, Gleckler PJ, Hewitson B, Mearns L (2010) Good practice guidance paper on assessing and combining multi model climate projections. In: Stocker TF, Qin D, Plattner GK, Tignor M, Midgley PM (eds) Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections. IPCC Working Group I Technical Support Unit, University of Bern, Bern, p 1–13

Macadam I, Pitman AJ, Whetton PH, Abramowitz G (2010) Ranking climate models by performance using actual values and anomalies: implications for climate change impact assessments. Geophys Res Lett 37:L16704, doi:10.1029/2010GL043877

Maxino CC, McAvaney BJ, Pitman AJ, Perkins SE (2008) Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation. Int J Climatol 28:1097–1112

Meehl GA, Stocker TF, Collins WD, Friedlingstein P and others (2007) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z and others (eds) Climate Change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, p 747–846

Mehrotra R, Sharma A (2010) Development and application of a multisite rainfall stochastic downscaling framework for climate change impact assessment. Water Resour Res 46:W07526, doi:10.1029/2009WR008423

Mehrotra R, Sharma A, Nagesh Kumar D, Reshmidevi TV (2013) Assessing future rainfall projections using multiple GCMs and a multi-site stochastic downscaling model. J Hydrol (Amst) 488:84–100

Milton JS, Arnold JC (2007) Introduction to probability and statistics. Tata McGraw Hill, New Delhi

Mujumdar PP, Ghosh S (2008) Modeling GCM and scenario uncertainty using a possibilistic approach: application to the Mahanadi River, India. Water Resour Res 44:W06407, doi:10.1029/2007WR006137

Mujumdar PP, Nagesh Kumar D (2012) Floods in a changing climate: hydrologic modeling. International Hydrology Series. Cambridge University Press, Cambridge

Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430:768–772

Ojha R, Kumar DN, Sharma A, Mehrotra R (2013) Assessing GCM convergence for the Indian region using the variable convergence score. J Hydrol Eng doi:10.1061/(ASCE)HE.1943-5584.0000888

Perkins SE, Pitman AJ, Holbrook NJ, McAveney J (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature and precipitation over Australia using probability density functions. J Clim 20:4356–4376

Perkins SE, Pitman AJ, Sisson SA (2009) Smaller projected increases in 20-year temperature returns over Australia in skill-selected climate models. Geophys Res Lett 36: L06710, doi:10.1029/2009GL037293

Perkins SE, Pitman AJ, Sissonb SA (2013) Systematic differences in future 20 year temperature extremes in AR4 model projections over Australia as a function of model skill. Int J Climatol 33:1153–1167

Pitman AJ, Arneth A, Ganzeveld L (2012) Regionalizing global climatic models. Int J Climatol 32:321–337

Pomerol JCh, Romero SB (2000) Multicriterion decision in management: principles and practice. Kluwer Academic, Dordrecht

Raje D, Mujumdar PP (2010) Constraining uncertainty in regional hydrologic impacts of climate change: nonstationarity in downscaling. Water Resour Res 46: W07543, doi:10.1029/2009WR008425

Raju KS, Nagesh Kumar D (2010) Multicriterion analysis in engineering and management. Prentice Hall of India, New Delhi

Randall DA, Wood RA, Bony S, Colman R and others (2007) Climate models and their evaluation. In: Solomon S, Qin D, Manning M, Chen Z and others (eds) Climate Change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC. Cambridge University Press, Cambridge, p 589–662

Reichler T, Kim J (2008) How well do coupled models simulate today's climate? Bull Am Meteorol Soc 89: 303–311

Reshmi Devi TV, Nagesh Kumar D (2010) Inter comparison of general circulation models for hydrologic predictability. In: Raju KS, Vasan A (eds) Proc Nat Conf on Sustainable Water Resources Management and Impact of Climate Change. BS Publications, Hyderabad, p 33–43

Smith IN, Chandler E (2009) Refining rainfall projections for

the Murray Darling basin of south-east Australia: the effect of sampling model results based on performance. Clim Change 102:377−393
Smith I, Chiew F (2010) Document and assess methods for generating inputs to hydrological models and extend delivery of projections across Victoria. Final report for Project 2.2.5P. South Eastern Australian Climate Initiative, CSIRO Land and Water, Acton
Srikanthan R (2009) A nested multisite daily rainfall stochastic generation model. J Hydrol (Amst) 371:142−153
Suppiah R, Hennessy KJ, Whetton PH, McInnes K and others (2007) Australian climate change projections derived from simulations performed for the IPCC 4th assessment report. Aust Meteorol Mag 56:131−152
Tang Y, Lin H, Moore AM (2008) Measuring the potential predictability of ensemble climate predictions. J Geophys Res 113:D04108, doi:10.1029/2007JD008804
Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philos Trans R Soc Lond A Math Phys Eng Sci 365:2053−2075
Wilby RL, Harris I (2006) A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. Water Resour Res 42:W02419, doi:10.1029/2005WR004065
Wilks DS (2011) Statistical methods in the atmospheric sciences. International Geophysics Series, Academic Press, San Diego, CA
Xu CY (1999) Climate change and hydrologic models: a review of existing gaps and recent research developments. Water Resour Manag 13:369−382

Editorial responsibility: Mikhail Semenov, Harpenden, UK

Submitted: May 21, 2013; Accepted: February 1, 2014
Proofs received from author(s): May 31, 2014