

Article

# Performance Assessment of General Circulation Model in Simulating Daily Precipitation and Temperature Using Multiple Gridded Datasets

Najeebullah Khan <sup>1,2</sup>, Shamsuddin Shahid <sup>1</sup>, Kamal Ahmed <sup>1,2</sup>, Tarmizi Ismail <sup>1</sup>,  
Nadeem Nawaz <sup>2</sup> and Minwoo Son <sup>3,\*</sup>

<sup>1</sup> School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Malaysia; najeebmarri@gmail.com (N.K.); sshahid@utm.my (S.S.); kamal\_brc@hotmail.com (K.A.); tarmiziismail@utm.my (T.I.)

<sup>2</sup> Faculty of Water Resource Management, Lasbela University of Agriculture Water and Marine Sciences (LUAWMS), 90150 Uthal, Balochistan, Pakistan; nn.engr96@yahoo.com

<sup>3</sup> Department of Civil Engineering, Chungnam National University, Daejeon 34134, Korea

\* Correspondence: mson@cnu.ac.kr; Tel.: +82-42-821-5676

Received: 28 September 2018; Accepted: 4 December 2018; Published: 6 December 2018



**Abstract:** The performance of general circulation models (GCMs) in a region are generally assessed according to their capability to simulate historical temperature and precipitation of the region. The performance of 31 GCMs of the Coupled Model Intercomparison Project Phase 5 (CMIP5) is evaluated in this study to identify a suitable ensemble for daily maximum, minimum temperature and precipitation for Pakistan using multiple sets of gridded data, namely: Asian Precipitation–Highly-Resolved Observational Data Integration Towards Evaluation (APHRODITE), Berkeley Earth Surface Temperature (BEST), Princeton Global Meteorological Forcing (PGF) and Climate Prediction Centre (CPC) data. An entropy-based robust feature selection approach known as symmetrical uncertainty (SU) is used for the ranking of GCM. It is known from the results of this study that the spatial distribution of best-ranked GCMs varies for different sets of gridded data. The performance of GCMs is also found to vary for both temperatures and precipitation. The Commonwealth Scientific and Industrial Research Organization, Australia (CSIRO)-Mk3-6-0 and Max Planck Institute (MPI)-ESM-LR perform well for temperature while EC-Earth and MIROC5 perform well for precipitation. A trade-off is formulated to select the common GCMs for different climatic variables and gridded data sets, which identify six GCMs, namely: ACCESS1-3, CESM1-BGC, CMCC-CM, HadGEM2-CC, HadGEM2-ES and MIROC5 for the reliable projection of temperature and precipitation of Pakistan.

**Keywords:** general circulation model; symmetrical uncertainty; Pakistan; temperature; precipitation; gridded dataset

## 1. Introduction

Simulations of general circulation models (GCM) are generally used for the assessment of climate change. However, the GCM simulated climate is associated with large uncertainties due to model structure, parametrization, assumption, calibration processes and so on [1,2]. Therefore, all the GCMs cannot be directly used to project the future climate in certain regions. In general, a small group of GCMs is selected for a given region by excluding those which have no similarity to the climate in order to reduce the uncertainty associated with the GCMs [3–6].

Generally, GCMs are selected based on several approaches; first, the past performance-based approach which selects GCM based on its ability to replicate historical climate [7] and second,

the envelope approach where GCMs are selected according to their agreement in the projections of the future climate [8]. The past performance evaluation method does not take account of the future projection of the climate while the envelope-based evaluation does not consider the GCM ability to replicate the past climate. Therefore, the third approach known as the hybrid approach, which combines the past-performance approach with the envelope method is often used. In this approach, the GCMs are screened based on their past performance. Then, the final set of GCMs are selected based on envelope approach [3]. However, when a whole pool of GCMs is considered for the selection of an ensemble, the past-performance approach is more realistic as the pool of GCMs covers the whole range of projections. Most of the previous studies also suggest that past performance evaluation is one of the most suitable approaches because the GCMs that are best in simulating the past climatic conditions are more likely to predict the future climate [9,10].

The past-performance of GCMs is usually assessed by comparing the GCM simulated climatic variables with the gridded climatic variables over a historical period. It is assumed that all gridded data used for this purpose represent the seasonal and spatial variability of the climate of the desired region. However, the performance of gridded climate data varies dramatically with space and time. This is mainly due to the type of interpolation technique, the quality of observed data and the spatial distribution of the gauges [11]. Additionally, the temporal variation in the number of observed data also affects the performance of gridded data over time. Therefore, the past performance assessment of GCMs based on single gridded data over a longer period may not be reliable.

Feature selection approaches are generally used for the selection of GCMs. The feature selection algorithm can be broadly classified into filter and wrapper. In the case of a filter approach, the GCMs are selected based on the underlying characteristics of the data using certain criteria such as correlation, mean error and so on. A wrapper approach uses a learning process to select a subset of features that have a higher capability to predict the independent variable. In the case of a filter approach, each feature is considered individually without consideration of correlation among the features. For GCM selection, it is not necessary to consider the relation between GCM simulations or predictive capability of GCMs. Therefore, filters are generally used for the selection of GCMs [12].

Various attempts have been conducted to determine the performance of GCMs in respect to gridded data using different types of filters which include weighted skill score [13,14], clustering hierarchy approach [15,16], spectral analysis [17], Bayesian weighting [18], different performance indicators [19,20] and so on. Besides, various statistical indicators have been used for GCM evaluation, ranking and selection [21–27]. The performance matrices are mostly evaluated on the basis of the mean state of the climatic condition where temporal variability such as trends or seasonal variability of the climate is not given full attention [28].

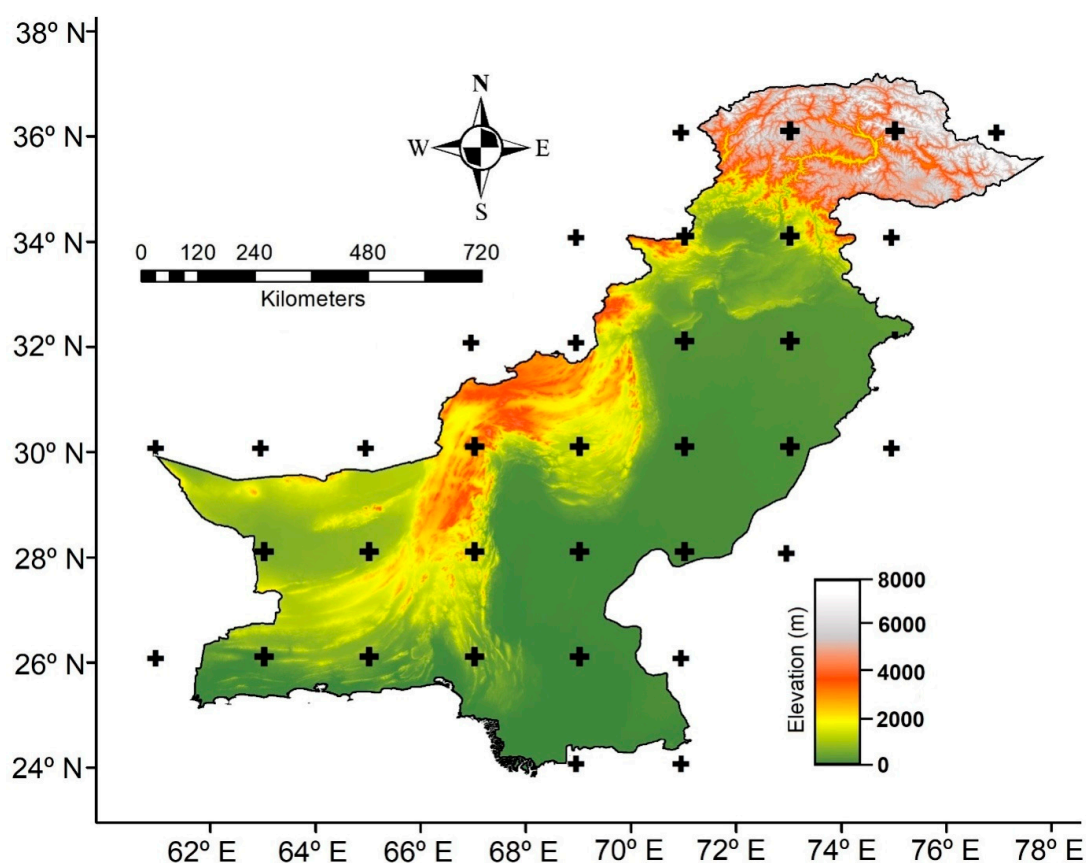
In recent years, one of the widely used algorithms for the variable selection is the symmetrical uncertainty (*SU*) algorithm. This filter-based tool [29] selects variables in an unbiased and reliable manner. The *SU* can track the similarity and dissimilarity between the time series of GCM simulation for historical period and observed climate variable and select the GCMs based on their ability to simulate observed climate variable. The major advantage of *SU* is that its performance does not depend on the underlying distribution and different conditional dependencies of data. It provides a general aspect of the association between the independent and the dependent variables. Therefore, it can be used to avoid multiple statistical metrics for performance assessment.

The objective of this study is to evaluate the capability of 31 CMIP5 GCMs in reconstructing of temperature and precipitation from different gridded data using *SU* to identify the most suitable ensemble of GCMs for the reliable climate projection. This study is applied to the case of Pakistan where a diverse climate with large spatial and seasonal variability exist. The ensemble of GCMs selected in this study can be used for reliable projection of climate and mitigation planning for Pakistan which is highly vulnerable to climate change.

## 2. Materials and Methods

### 2.1. Geography and Climate of Pakistan

Pakistan (latitudes  $24^{\circ}00' \text{ N}$ – $37^{\circ}30' \text{ N}$  and longitudes  $61^{\circ} \text{ E}$ – $78^{\circ} \text{ E}$ ) covers an area of  $796,095 \text{ km}^2$  showing a predominantly arid to semi-arid climate [30]. The topography of Pakistan significantly varies from a high and elevated region in the north to the coastal flat plains in the south (Figure 1), depending on which huge spatial and temporal variation in the climatic condition is observed. The precipitation also varies widely with less than 100 mm in the western and northwestern area and more than 1100 mm in the northeastern and northern part of the country. The precipitation is observed to increase from the arid southern coastal regions to the northern and northeastern region of the country. However, the major parts of the country have a precipitation less than 500 mm [31]. The south eastern parts of the country have the highest rainfall during summer (June–September) and least rainfall during winter (December–February), while the northern parts have the highest rainfall during winter and the lowest rainfall in summer. The maximum temperature is often above  $50^{\circ} \text{ C}$  in the south in May, while the minimum temperature is below  $-20^{\circ} \text{ C}$  in the northern Himalaya and sub-Himalaya region in January [30].



**Figure 1.** Topographical map of the study area with the distributed points at  $2^{\circ} \times 2^{\circ}$ .

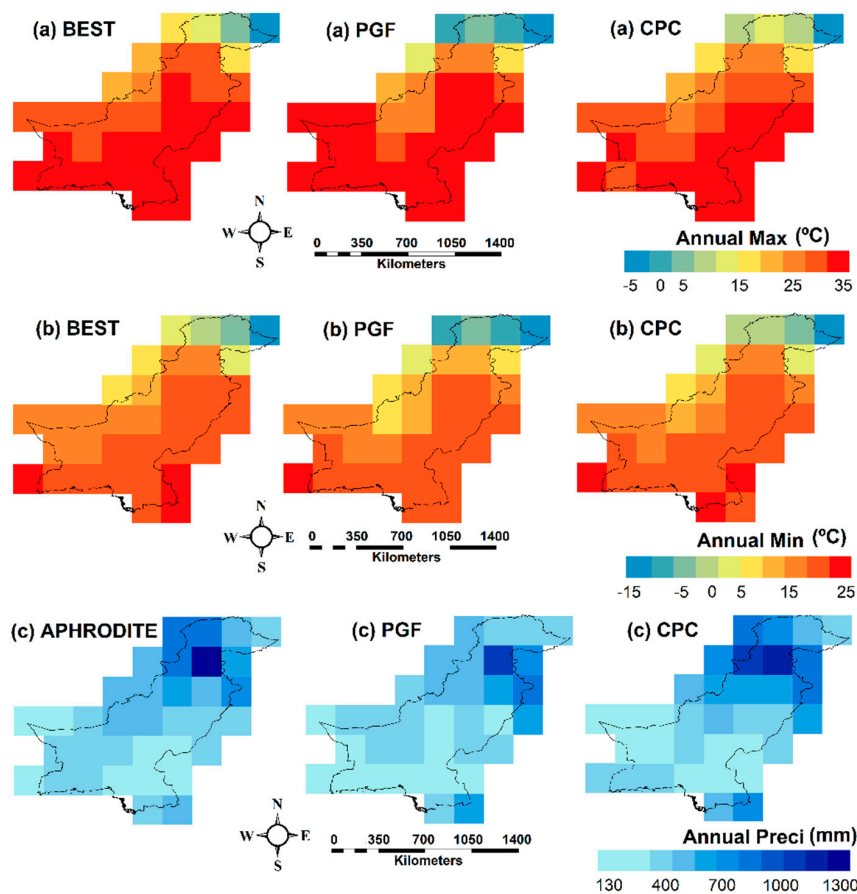
### 2.2. Data and Sources

#### 2.2.1. Gridded Datasets

Four gridded climate datasets are used in this study namely, Berkeley Earth Surface Temperature (BEST), Climate Prediction Centre (CPC), the Princeton Global Meteorological Forcing (PGF) and the Asian Precipitation Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE). The choice of these datasets is based on the availability of daily data for

a longer period. The gridded data products which have the daily resolution are used for the assessment of the performance of GCMs with an expectation that GCMs selected based on those gridded data can be used to assess the changes in different climatic extremes which are estimated by daily gridded data and selected by GCMs. Assessment of the changes of climate extremes due to global warming is important for Pakistan as most of the studies revealed that major impact of climate change in the country will be due to the changes of climatic extremes such as heat waves, hot days, dry spells and heavy rainfall [32]. GCMs selected from daily rainfall and temperature gridded data is suitable to this purpose. The major characteristics and sources of the gridded data used in this study are summarized in Table 1. The BEST dataset is obtained from data from 36,866 stations [33], while the CPC dataset is developed using 30,000 stations data [34]. The PGF dataset is developed by merging the reanalysis data with observed data collected across the globe [35,36]. On the other hand, APHRODITE is developed using 5000 to 12,000 valid stations in the Monsoon Asian region (60° E–150° E longitude, 15° S–55° N latitude) [37–39].

Among the four gridded datasets maximum and minimum temperature data of BEST, CPC and PGF are used to assess the GCMs’ temperature, while the precipitation of CPC, PGF and APHRODITE are used for the assessment of GCMs’ precipitation. The annual mean maximum, minimum and precipitation for each dataset are presented in Figure 2. The selected data products are available in different spatial resolution range from 0.25 to 1°. The gridded datasets are regridded to a common spatial resolution of 2° × 2° to be in accord with the agreed resolution of the GCM. The placement of these grid points across the study area is shown in Figure 1. Grid average technique is used to upscale the gridded data into the coarse resolution. In this approach, temperature and precipitation data of all the grids within each 2° × 2° domain are averaged to convert the higher resolution data to lower resolution.



**Figure 2.** Spatial distribution of annual average of daily (a) maximum temperature; (b) minimum temperature; and (c) annual total rainfall estimated using different gridded temperature and rainfall data.

**Table 1.** The description of the gridded datasets used in the study.

Products	Source	Period	Spatial Resolution	Climate Variable
BEST [33]	Berkley Earth Surface Temperature [40]	(1960–2013)	1° × 1°	Daily Temp.
CPC [34]	Climate Prediction Centre [41]	(1979–present)	0.5° × 0.5°	Daily Temp. and Precipitation
APHRODITE [37]	Asian Precipitation—Highly Resolved Observational Data Integration Toward Evaluation, (Monsoon Asia) [42]	(1951–2007)	0.5° × 0.5°	Daily Precipitation
PGF [35]	Princeton Global Meteorological Forcing [43]	(1948–2010)	0.25° × 0.25°	Daily Temp. and Precipitation

### 2.2.2. Coupled Model Intercomparison Project Phase 5 (CMIP5) General Circulation Model (GCM) Datasets

A total of 31 CMIP5 GCM simulations are used in this study. The name, modeling center and resolution of the GCMs used in the study are given in Table 2. The GCM simulations are downloaded from the Inter-governmental Panel on Climate Change (IPCC) portal [44]. The GCMs are selected on the basis of the availability of daily simulation for two representative concentration pathways (RCP) scenarios, RCP4.5 and RCP8.5. The RCP4.5 is an intermediate pathway scenario which shows a good agreement with the latest policy of lower greenhouse gas emission by the global community. Therefore, it is often considered as a very-good-case scenario in the context of recent policy directions [45]. On the other hand, RCP 8.5 is the business-as-usual scenario which provides the possible highest impact on climate change. Therefore, RCP 4.5 and RCP 8.5 were selected as these two scenarios can provide a possible complete range of impact. The GCMs are available in different resolutions. Therefore, all selected CMIP5 temperature and precipitation simulations are regridded into the common 2° × 2° grid for fair comparison using bilinear interpolation technique. This technique uses a minimum of four points from the domain and nearby areas to generate point output from each GCM at each grid point and thus provides a smooth interpolation. It is widely used for regridding of GCMs for comparison [4].

**Table 2.** The Coupled Model Intercomparison Project Phase 5 (CMIP5) general circulation models (GCMs) used in this study for the selection of an ensemble for Pakistan.

GCM No.	GCM Name	Developer	Resolution
1	ACCESS1-0	Commonwealth Scientific and Industrial Research Organisation—Bureau of Meteorology, Australia	1.9 × 1.2
2	ACCESS1-3		
3	bcc-csm1-1-m	Beijing Climate Centre, China	2.8 × 2.8
4	bcc-csm1-1		1.1 × 1.1
5	BNU-ESM	Beijing Normal University, China	2.8 × 2.8
6	CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	2.8 × 2.8
7	CCSM4	National Centre for Atmospheric Research USA	0.94 × 1.25
8	CESM1(BGC)		0.94 × 1.25
9	CMCC-CM	Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy	0.7 × 0.7
10	CMCC-CMS		1.9 × 1.9
11	CNRM-CM5	Centre National de Recherches Météorologiques, Centre, France	1.4 × 1.4
12	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization, Australia	1.9 × 1.9
13	EC-EARTH	EC-EARTH consortium published at the Irish Centre for High-End Computing, Netherlands/Ireland	1.1 × 1.1
14	FGOALS-g2	Institute of Atmospheric Physics, Chinese Academy of Sciences, China	2.8 × 2.8



Table 2. Cont.

GCM No.	GCM Name	Developer	Resolution
15	GFDL-CM3		2.5 × 2.0
16	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, USA	2.5 × 2.0
17	GFDL-ESM2M		2.5 × 2.0
18	GISS-E2-R	NASA/GISS (Goddard Institute for Space Studies), USA	2.5 × 2.0
19	HadGEM2-CC	Met Office Hadley Centre, UK	1.9 × 1.2
20	HadGEM2-ES		1.9 × 1.2
21	INMCM4.0	Institute of Numerical Mathematics, Russia	2.0 × 1.5
22	IPSL-CM5A-LR	Institut Pierre Simon Laplace, France	3.7 × 1.9
23	IPSL-CM5A-MR		2.5 × 1.3
24	IPSL-CM5B-LR		3.7 × 1.9
25	MIROC-ESM-CHEM	The University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	2.8 × 2.8
26	MIROC-ESM		2.8 × 2.8
27	MIROC5		1.4 × 1.4
28	MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	1.9 × 1.9
29	MPI-ESM-MR		1.9 × 1.9
30	MRI-CGCM3	Meteorological Research Institute, Japan	1.1 × 1.1
31	NorESM1-M	Meteorological Institute, Norway	2.5 × 1.9

### 3. Methodology

#### 3.1. Procedure

Temporal alignment does not exist between the daily simulation of GCMs and observed data. However, it is expected that GCMs are able to characterize the wet and dry seasons and warm and cold seasons reasonably well. Therefore, the daily gridded datasets are converted to monthly scale and the association of GCM data with gridded data is compared at a monthly scale to evaluate the performance of GCMs. The gridded data products which have daily simulations are used for the assessment of the performance of GCMs in order to facilitate the changes in different climatic extremes. The procedure for the evaluation of GCMs based on their historical simulation of temperature and precipitation is outlined below:

- (1) The GCMs are evaluated and ranked using *SU* at each grid point (total 36 grid points at  $2^\circ \times 2^\circ$  resolution to cover whole Pakistan, Figure 1) according to their capability to replicate gridded temperature and precipitation data for the period 1979–2005.
- (2) The GCMs are ranked based on their positions at different grid points using a weighting technique, where more weight is given to GCMs which archive a higher rank in most of the grid points. A separate list of rank is prepared for each climatic variable (maximum and minimum temperature and precipitation) and each gridded dataset.
- (3) The common GCMs appeared above the 50-th percentile of all GCMs in each list are identified as the most appropriate ensemble for the projection of temperature and precipitation of Pakistan.

The daily mean temperature can be used for the selection of GCMs instead of selecting GCMs separately for maximum and minimum temperature. For the projection of extremes, better simulation of minimum and maximum temperature instead of mean temperature is required. However, better simulation mean the temperature does not guarantee a better simulation of maximum and minimum temperature. Therefore, GCMs are selected in this study based on both maximum and minimum temperature. The ranking of GCMs at each grid point using *SU* and the ranking of GCMs for whole of Pakistan considering the *SU* rank at each grid points using a weighting technique are elaborated in the following sections.

### 3.2. Symmetric Uncertainty (SU)

It is expected that GCMs have a capability to replicate the annual and seasonal variations of observed rainfall and temperature for the historical period. *SU* is a filtering approach based on information theory which can be used to assess the similarity between two time series [46]. In this study, *SU* is used to measure the similarity between the time series of gridded data and GCMs historical simulation [4,5]. For example, the monthly rainfall time series of the gridded data product is compared with result of GCMs at each grid point to assess the capability of the GCMs to replicate rainfall at each grid point.

*SU* uses the concept of mutual information (*MI*) to assess the association between two variables as shown in Figure 3. *MI* estimates the amount of information common between two variables. Let *A* and *B* be gridded rainfall and GCMs historical rainfall simulation at a grid point. If  $p(A)$  and  $p(B)$  are the probability density functions and  $p(A, B)$  is the joint probability density function of *A* and *B*, then *MI* between *A* and *B* is [47,48],

$$MI(A, B) = \sum p(A, B) \log \frac{p(A, B)}{p(A) \cdot p(B)} \tag{1}$$

*MI* estimates common information between two variables as the difference between the sum of the entropies and their joint entropy:

$$MI(A, B) = H(A) - H(A, B) \tag{2}$$

where,  $H(A)$  and  $H(A, B)$  denotes Shannon’s entropy of *A* and the joint entropy of *A* and *B*, respectively. The *MI* estimated using Equation (1) indicates the amount of mutual information between gridded rainfall and GCMs historical rainfall simulation have. If the variables are independent to each other, the *MI* is zero while a higher value of *MI* indicates the GCMs simulated rainfall has higher similarity with the gridded rainfall.

The *MI* is biased toward the variable having higher values. This can be overcome using *SU* where the estimated *MI* is divided by the total entropies of GCMs simulation (*B*) and gridded data (*A*):

$$SU(A, B) = 2 \frac{MI(A, B)}{H(A) + H(B)} \tag{3}$$

*SU* value 1 means a complete agreement between gridded data and GCMs simulation, while 0 indicates no agreement [49]. Details of the selection of GCMs based on *SU* are found in [4,5].

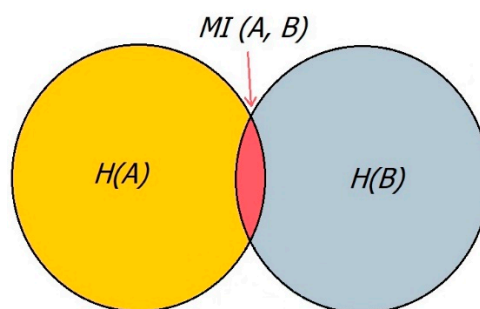


Figure 3. The concept of mutual information.

### 3.3. Ranking of GCMs using Weighting Method

*SU* provides a ranking of GCMs at each grid point. A simple multi-criteria decision-making approach is used in this study for the selection of GCMs for the whole study area based on the ranking of GCMs obtained at different grid points. If a GCM achieved 1st, 2nd, 3rd, 4th and 5th ranks at  $X_1, X_2, X_3, X_4$  and  $X_5$  number of grid points, the GCMs are assigned a weighted score:

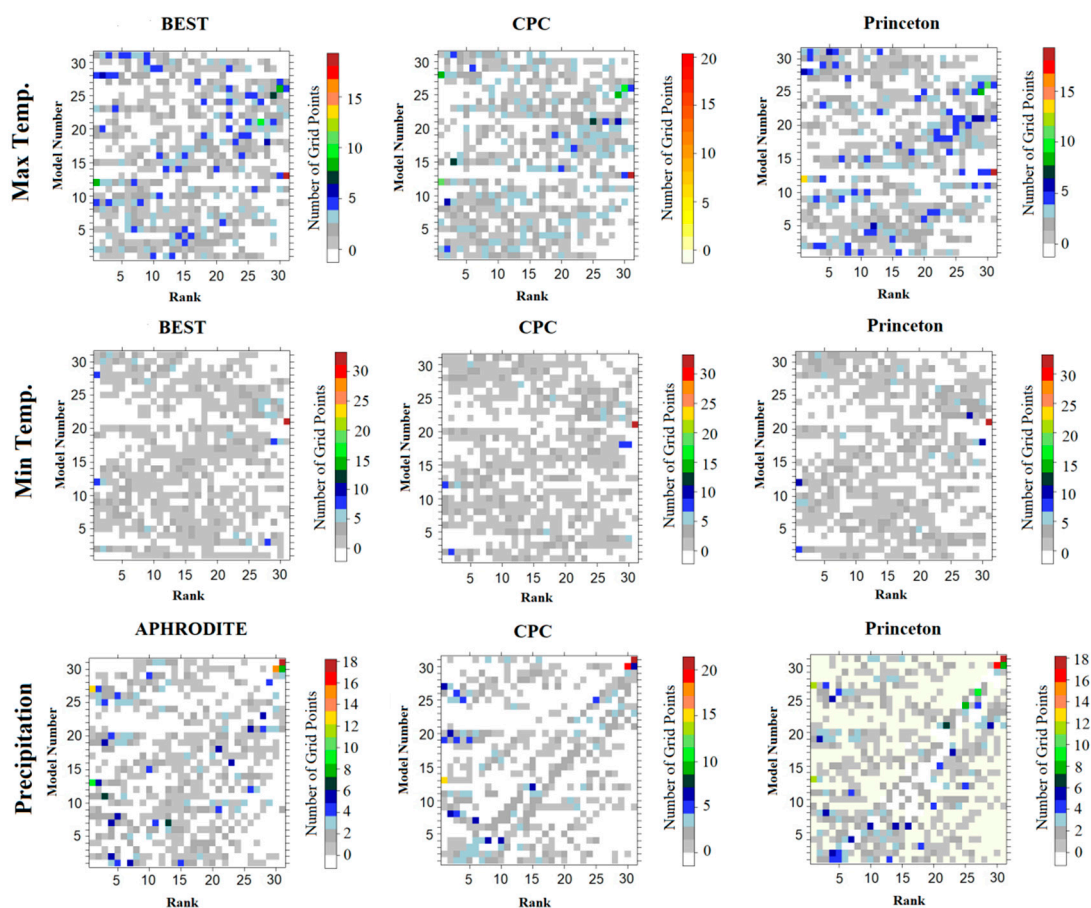
$$s = X_1 (1/1) + X_2 (1/2) + X_3 (1/3) + X_4 (1/4) + X_5 (1/5) \quad (4)$$

The GCMs rank more than 5th at a grid point are not considered because it is considered that they are not fully capable to simulate temperature or precipitation at that grid point.

## 4. Results

### 4.1. Ranking of GCMs

The ranking of GCMs in respect of their capability to replicate maximum and minimum temperature of BEST, CPC and PGF and precipitation of APHRODITE, CPC and PGF obtained by *SU* are shown in Figure 4 (heatmap plots). The *x*-axis and *y*-axis of the plot represent GCMs and their ranks, respectively. The number in the *y*-axis corresponds to GCM number which can be found in Table 2. For example, GCM number 1 represents ACCESS1-0, number 2 represents ACCESS1-3 and so on as seen in Table 1. The color ramp in the plot indicates the number of grid points. For example, CSIRO-Mk3-6-0 (GCM 12) achieved rank 1 at 11 grid points, EC-EARTH (GCM 13) is not ranked top at any grid point.



**Figure 4.** Ranking of the GCM according to their performance in replicating maximum and minimum temperature and precipitation. The *x*-axis represents the rank of the GCM while the *y*-axis represents the GCM number as described in Table 2. The color bar shows the number of grid points.

Different GCMs are ranked top at a different number of grid points (Figure 4). However, there is a consistency among the plots which indicate a similar ranking of GCMs at each grid points obtained from different gridded datasets. For example, CSIRO-Mk3-6-0 (GCM 12) is found to be best at 11 grid points for CPC and BEST, while it is found best at 12 grid points for PGF. The MPI-ESM-LR (GCM 28)



is found best at 6 to 8 grid points for different gridded datasets. Overall, the BEST and CPC show better similarity in GCM ranks compared to PGF.

The similar consistency is observed in the case of minimum temperature. BEST and CPC produce similar rankings of GCMs in the case of minimum temperature. The GCM ranks using PGF is also found very close to that obtained by BEST and CPC. In the case of precipitation, better similarity is observed in the case of APHRODITE and PGF. A slight variation in the ranking is observed in the CPC. It is known from the results that rankings of GCMs vary when different gridded datasets are used although the intensity of variation is not significantly high.

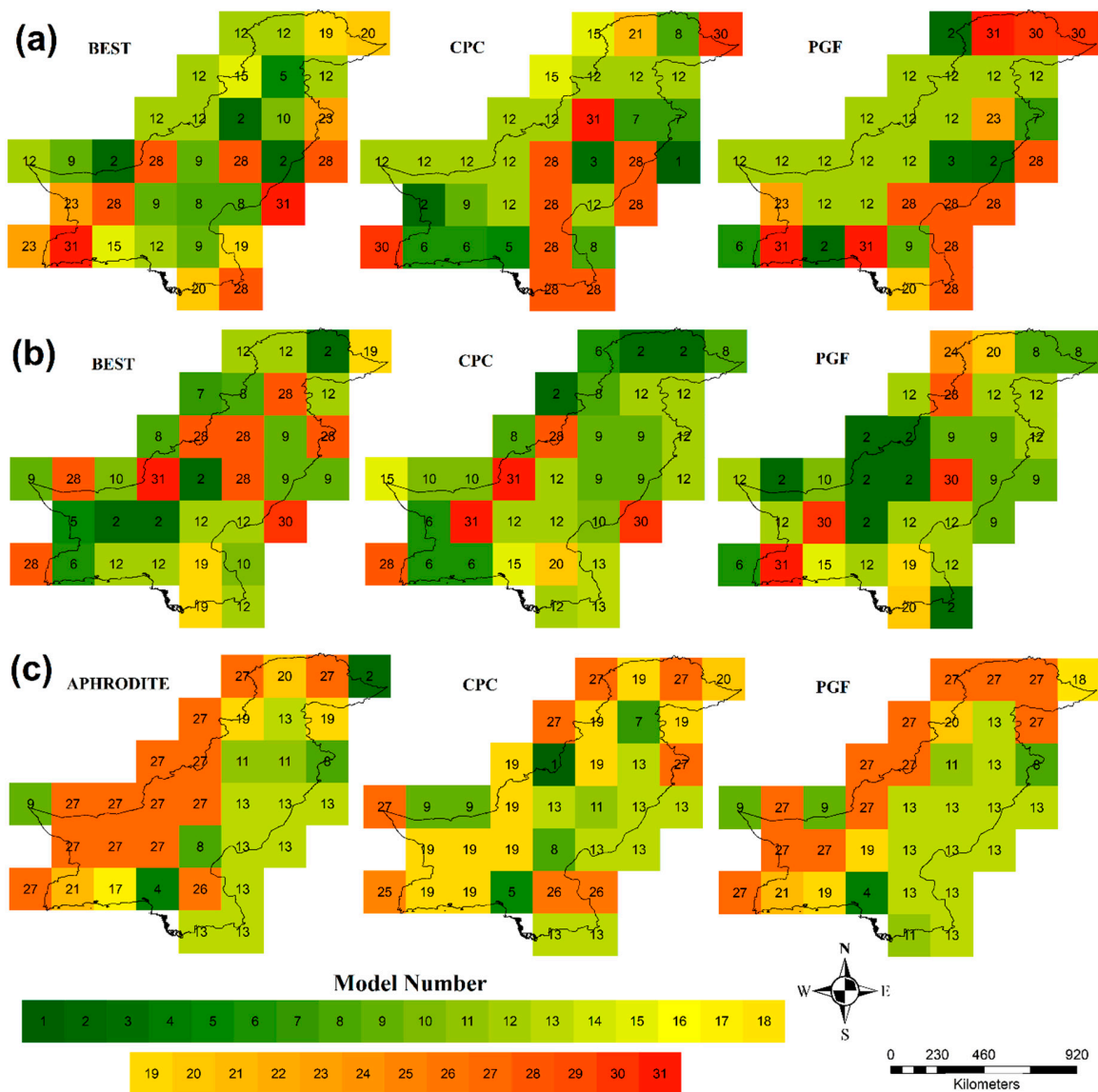
#### 4.2. Spatial Distribution of Top Ranked GCMs

Although the number of GCMs ranked top are found to vary less for different gridded datasets as shown in the previous section, it is still required to assess whether the spatial distribution of the ranking varies for different gridded data or not. For this purpose, the spatial distribution of the GCMs ranked top using different gridded data is prepared. Figure 5 shows the spatial distribution of top ranked GCMs in simulating maximum and minimum temperature and precipitation, respectively at different gridded data. They show a significant difference in the spatial distribution of the top-ranked GCMs for different gridded data for the same climatic variable.

Figure 5 shows best GCMs at different grid points in term of their capability to simulate gridded maximum and minimum temperatures. CSIRO-Mk3-6-0 (GCM 12) shows the best result in terms of simulating the maximum temperature at most of the grid points (8 to 14 grid points for different gridded datasets) followed by MPI-ESM-LR (5 to 7 grid points). The performance of CSIRO-Mk3-6-0 (GCM 12) is mostly noticed in the western and north-western mountainous regions of Pakistan, while MPI-ESM-LR (GCM 28) in the south-eastern parts. Few other GCMs are found best at 1 to 3 grid points at different locations. CPC and PGF show a similar spatial pattern in best performing GCMs, while it is found a bit different from BEST.

The best GCMs in simulating the minimum temperature are found similar to results for the maximum temperature at most of the grid points. CSIRO-Mk3-6-0 (GCM 12) performs best at most of the grid points (8 to 11 grid points for different gridded datasets) followed by CMCC-CM (GCM 9) and ACCESS1-3 (GCM 2). CSIRO-Mk3-6-0 (GCM 12) is best in simulating the minimum temperature in the northern coastal region and southern mountainous regions, while CMCC-CM (GCM 9) in the central-western part. Similar to the maximum temperature, CPC and PGF show a similar spatial pattern in best performing GCMs in simulating minimum temperature while the BEST showed a bit different pattern.

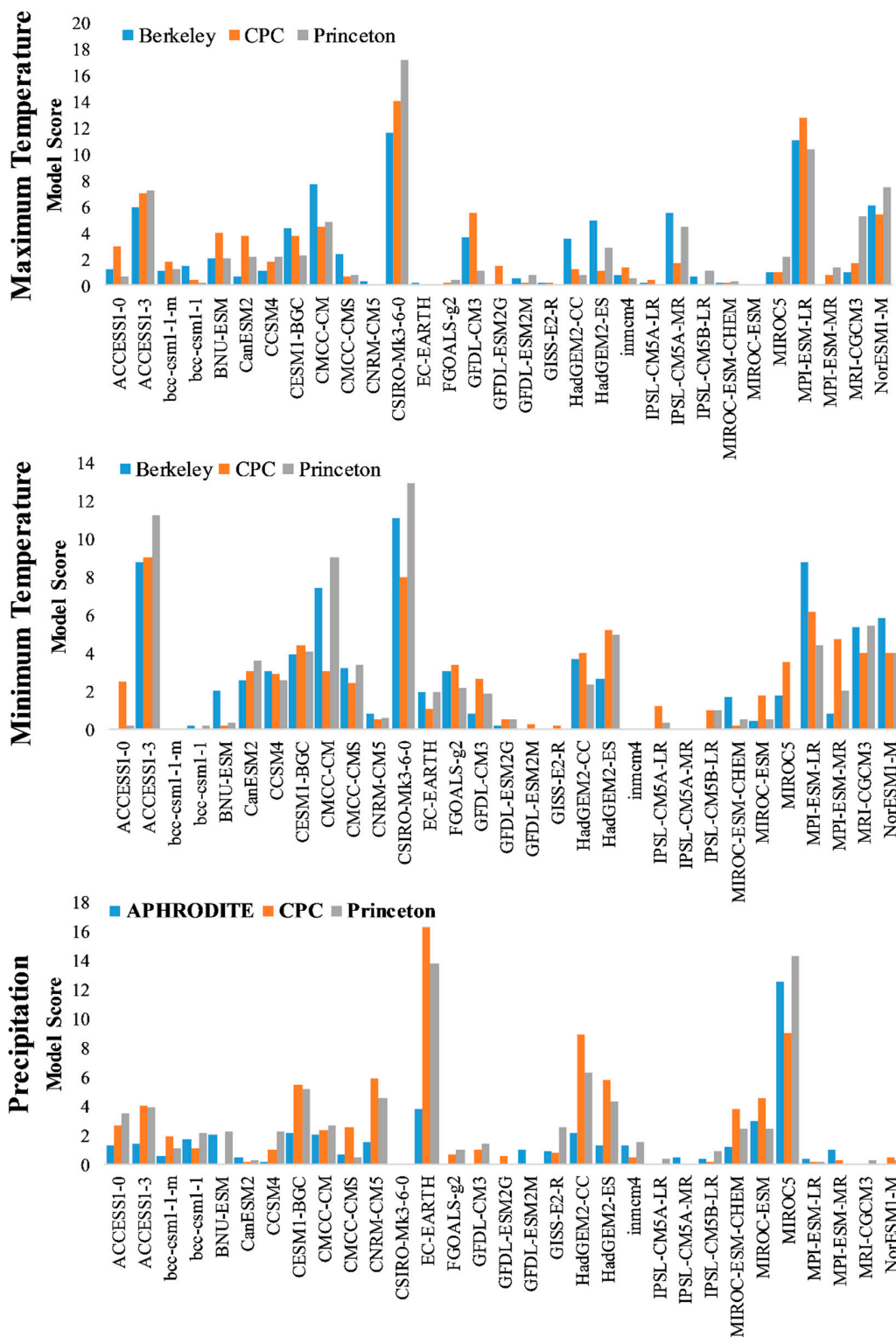
In the case of precipitation, the significant difference in the spatial pattern of best performing GCM is observed for different gridded data. MIROC5 (GCM 27) is found as best performing GCM in simulating precipitation for APHRODITE (13 grid points) followed by EC-EARTH (GCM 13) (9 grid points). For CPC the HadGEM2-CC (GCM 19) is found as the best performing (11 grid points) followed by EC-EARTH (GCM 13) (8 grid points). In the case of PGF, HadGEM2-CC (GCM 19) is found best at most of the grid points (12 grid points) followed by HadGEM2-CC (GCM 19) (7 grid points). Overall, HadGEM2-CC (GCM 19) is found to perform best in the southwest part of Pakistan for all the gridded datasets, while no consistency in GCM is observed in other parts.



**Figure 5.** The geographical distribution of best performing GCMs in simulating: (a) maximum temperature; (b) minimum temperature; and (c) precipitation over Pakistan for different gridded datasets. The number in the figure represent the GCM as mentioned in Table 2.

#### 4.3. Ranking of GCMs for Pakistan

The rankings of GCMs are estimated by Equation (4). The estimated scores for each GCMs in simulating the temperature and the precipitation over Pakistan are shown in Figure 6. Based on the scores the GCMs are ranked for different dataset, the ranks for each of the three datasets in the maximum, the minimum and the precipitation are averaged to calculate the score of the GCMs. CSIRO-Mk3-6-0 is found to receive the highest score for both the maximum and the minimum temperature for all the gridded data. MPI-ESM-LR is found to receive the second highest score for the maximum temperature, while ACCESS1-3 for the minimum temperature. Different GCMs are found to receive the higher score for precipitation. EC-EARTH is found to receive the highest score for CPC and PGF while MIROC5 for APHRODITE. EC-Earth is found to receive the second highest score for APHRODITE while MIROC5 for CPC and PGF.



**Figure 6.** Score estimated for different GCMs in simulating the maximum temperature (upper panel), the minimum temperature (middle panel) and the precipitation (lower panel) for Pakistan.

To assess the performance of *SU* in the selection of GCMs, a ranking of GCMs obtained from *SU* for the whole of Pakistan is compared with that obtained from conventional statistics. Two statistical metrics namely coefficient of determination ( $R^2$ ) and normalized root mean square error (*NRMSE*) are used for this purpose. The two metrics are estimated for all the GCMs at each grid point and then Equation (4) is used to estimate the score for each GCM for whole Pakistan. The scores obtained

using  $SU$ ,  $R^2$  and  $NRMSE$  for whole Pakistan are given in Table 3. The scores for  $R^2$  and  $NRMSE$  are averaged. GCMs are then ranked according to the combined score. The ranking obtained by different GCMs using  $SU$  and a combination of conventional statistical metrics shows very similar results, particularly for the top positions. For example, CSIRO-Mk3-6-0 is ranked top by  $SU$  while combined statistics ranked it as 3; MPI-ESM-LR is ranked as 2 by  $SU$  and 1 by combined statistics. This indicates that  $SU$  can be used for the ranking of GCMs by avoiding the use of a number of conventional statistical metrics.

**Table 3.** The scores obtained from different GCMs according to their performance estimated using symmetric uncertainty and conventional statistics.

GCM No.	Model Name	Score (SU)	Rank (SU)	Score ( $R^2$ )	Score (NRMSE)	Average ( $R^2 + NRMSE$ )	Rank ( $R^2 + NRMSE$ )
1	ACCESS1-0	1.66	15	3.01	3.77	3.39	15
2	ACCESS1-3	6.77	3	3.07	5.98	4.02	6
3	bcc-csm1-1	0.00	18	3.30	1.90	2.60	21
4	bcc-csm1-1-m	0.00	18	2.84	3.53	3.19	18
5	BNU-ESM	2.69	10	3.06	3.53	3.29	16
6	CanESM2	2.21	12	3.98	3.04	3.51	12
7	CCSM4	1.70	14	4.50	2.49	3.49	13
8	CESM1-BGC	3.48	7	5.48	1.58	3.53	11
9	CMCC-CM	5.64	5	2.84	4.92	3.88	8
10	CMCC-CMS	0.00	18	5.07	2.40	3.73	9
11	CNRM-CM5	0.00	18	3.16	1.51	2.33	24
12	CSIRO-Mk3-6-0	14.25	1	3.88	5.69	4.79	3
13	EC-EARTH	0.00	18	1.94	2.22	2.08	28
14	FGOALS-g2	0.00	18	3.29	3.10	3.20	17
15	GFDL-CM3	3.43	8	4.01	2.77	3.39	14
16	GFDL-ESM2G	0.00	18	1.49	1.06	1.28	30
17	GFDL-ESM2M	0.00	18	1.83	2.99	2.41	23
18	GISS-E2-R	0.00	18	4.93	3.47	4.20	4
19	HadGEM2-CC	1.83	13	4.24	2.87	3.56	10
20	HadGEM2-ES	2.99	9	2.52	3.38	2.95	20
21	INMCM4	0.00	18	1.76	3.26	2.51	22
22	IPSL-CM5A-LR	0.00	18	1.92	2.59	2.26	25
23	IPSL-CM5A-MR	3.93	6	2.98	5.13	4.05	5
24	IPSL-CM5B-LR	0.00	18	1.08	3.38	2.23	26
25	MIROC5	1.41	17	3.50	2.42	2.96	19
26	MIROC-ESM	0.00	18	0.76	2.11	1.44	29
27	MIROC-ESM-CHEM	1.41	16	0.92	3.47	2.19	27
28	MPI-ESM-LR	11.37	2	4.69	5.33	5.01	1
29	MPI-ESM-MR	0.00	18	0.00	1.68	0.84	31
30	MRI-CGCM3	2.66	11	3.17	4.75	3.96	7
31	NorESM1-M	6.30	4	5.10	4.50	4.80	2

In the next step, GCMs are ranked based on their performance scores obtained for different sets of gridded data using  $SU$  and Equation (4). The average of scores obtained by each GCM for different sets of gridded data is calculated for the ranking of GCMs. The average scores obtained from different GCMs in simulating temperature and precipitation are summarized in Table 4. CSIRO-Mk3-6-0, MPI-ESM-LR, ACCESS1-3, NorESM1-M and CMCC-CM estimate the highest average score for the maximum temperature, CSIRO-Mk3-6-0, ACCESS1-3, CMCC-CM, MPI-ESM-LR, MRI-CGCM3 for the minimum temperature, MIROC5, EC-EARTH, HadGEM2-CC, CESM1-BGC, CNRM-CM5 for the precipitation.

**Table 4.** The scores obtained by different GCMs according to their performance based on different gridded dataset.

Model	Max Temperature				Min Temperature				Precipitation			
	BEST	CPC	PGF	Score	BEST	CPC	PGF	Score	APH	CPC	PGF	Score
ACCESS1-0	1.3	3.0	0.7	1.7	0.0	2.5	0.2	0.9	1.3	2.7	3.5	2.5
ACCESS1-3	6.0	7.0	7.3	6.8	8.7	9.0	11.2	9.7	1.5	4.0	3.9	3.1
bcc-csm1-1-m	1.2	1.8	1.2	1.4	0.0	0.0	0.0	0.0	0.5	1.9	1.1	1.2
bcc-csm1-1	1.5	0.5	0.3	0.7	0.2	0.0	0.2	0.1	1.7	1.1	2.2	1.6
BNU-ESM	2.1	4.0	2.0	2.7	2.0	0.2	0.3	0.9	2.0	0.0	2.2	1.4
CanESM2	0.7	3.8	2.2	2.2	2.5	3.1	3.6	3.1	0.5	0.2	0.3	0.3
CCSM4	1.1	1.8	2.2	1.7	3.0	2.9	2.6	2.8	0.2	1.0	2.3	1.2
CESM1-BGC	4.3	3.8	2.3	3.5	3.9	4.4	4.1	4.1	2.2	5.4	5.1	4.2
CMCC-CM	7.7	4.5	4.8	5.6	7.4	3.1	9.0	6.5	2.0	2.3	2.7	2.3
CMCC-CMS	2.4	0.7	0.7	1.3	3.2	2.4	3.4	3.0	0.7	2.6	0.5	1.2
CNRM-CM5	0.3	0.0	0.0	0.1	0.9	0.5	0.6	0.7	1.5	5.9	4.5	4.0
CSIRO-Mk3-6-0	11.6	14.0	17.2	14.3	11.1	8.0	12.9	10.7	0.0	0.0	0.0	0.0
EC-EARTH	0.2	0.0	0.0	0.1	2.0	1.0	1.9	1.6	3.8	16.3	13.8	11.3
FGOALS-g2	0.0	0.2	0.5	0.2	3.0	3.4	2.2	2.9	0.0	0.7	1.0	0.6
GFDL-CM3	3.7	5.5	1.1	3.4	0.8	2.7	1.9	1.8	0.0	1.0	1.4	0.8
GFDL-ESM2G	0.0	1.5	0.0	0.5	0.2	0.5	0.5	0.4	0.0	0.5	0.0	0.2
GFDL-ESM2M	0.5	0.2	0.8	0.5	0.0	0.3	0.0	0.1	1.0	0.0	0.0	0.3
GISS-E2-R	0.3	0.2	0.0	0.2	0.0	0.2	0.0	0.1	0.9	0.8	2.5	1.4
HadGEM2-CC	3.5	1.2	0.8	1.8	3.7	4.0	2.3	3.3	2.2	8.9	6.3	5.8
HadGEM2-ES	5.0	1.1	2.9	3.0	2.6	5.2	4.9	4.3	1.3	5.8	4.4	3.8
INMCM4	0.8	1.3	0.5	0.9	0.0	0.0	0.0	0.0	1.3	0.5	1.5	1.1
IPSL-CM5A-LR	0.3	0.4	0.0	0.2	0.0	1.3	0.3	0.5	0.0	0.0	0.3	0.1
IPSL-CM5A-MR	5.5	1.7	4.5	3.9	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.2
IPSL-CM5B-LR	0.7	0.0	1.2	0.6	0.0	1.0	1.0	0.7	0.3	0.2	0.9	0.5
MIROC-ESM-CHEM	0.2	0.2	0.3	0.2	1.7	0.2	0.5	0.8	1.2	3.8	2.4	2.4
MIROC-ESM	0.0	0.0	0.0	0.0	0.5	1.8	0.5	0.9	3.0	4.5	2.5	3.3
MIROC5	1.1	1.0	2.1	1.4	1.8	3.5	0.0	1.8	12.5	9.0	14.2	11.9
MPI-ESM-LR	11.0	12.7	10.4	11.4	8.7	6.2	4.4	6.4	0.3	0.2	0.2	0.2
MPI-ESM-MR	0.0	0.8	1.4	0.7	0.8	4.7	2.0	2.5	1.0	0.3	0.0	0.4
MRI-CGCM3	1.0	1.7	5.2	2.7	5.4	4.0	5.4	4.9	0.0	0.0	0.3	0.1
NorESM1-M	6.1	5.4	7.5	6.3	5.8	4.0	4.0	4.6	0.0	0.5	0.3	0.3

#### 4.4. Selection of GCM Ensemble

Both temperature and precipitation projections are required for most of the climate change impact assessments. Therefore, GCMs that can simulate both temperature and precipitation properly are desirable for climate change impact analysis. In this study, the final ensemble of GCMs is selected based on their simulation capability.

For this purpose, GCMs for each climate variable are first ranked based on their average score obtained using different sets of gridded data. If the average score of a GCM for a certain climate variable is high, it can be considered that the GCM is capable of simulating that climate variable and vice-versa. In this study, it is assumed that GCM is not capable of simulating the climate variable when GCMs do not achieve a position above 50-th percentile of all GCMs for the climate variable. Therefore, those GCMs which received a position less than the 50-th percentile of all GCMs for a climate variable were given a zero score for that specific climate variable. For example, CMCC-CMS receives an average score of only 1.3 and thus a rank of 18 for the maximum temperature below 50-th percentile position (rank 16) of all GCMs. Thus, it is considered that CMCC-CMS is not well capable to simulate the maximum temperature for Pakistan and the zero score is assigned for this GCM.

The scores of GCMs after applying this criterion are given in Table 5. It shows only six GCMs namely; ACCESS1-3, CESM1-BGC, CMCC-CM, HadGEM2-CC, HadGEM2-ES and MIROC5 receive the scores for all climate variables. Some of the GCMs which receive a higher average score for one climate variable are found to receive the zero score for other climate variables. For example, EC-Earth receives the highest average score for precipitation, but the zero score for both the maximum and minimum temperature. ACCESS1-3 received the high score for the maximum temperature and precipitation, but the zero score for the minimum temperature. Therefore, those GCMs which cannot



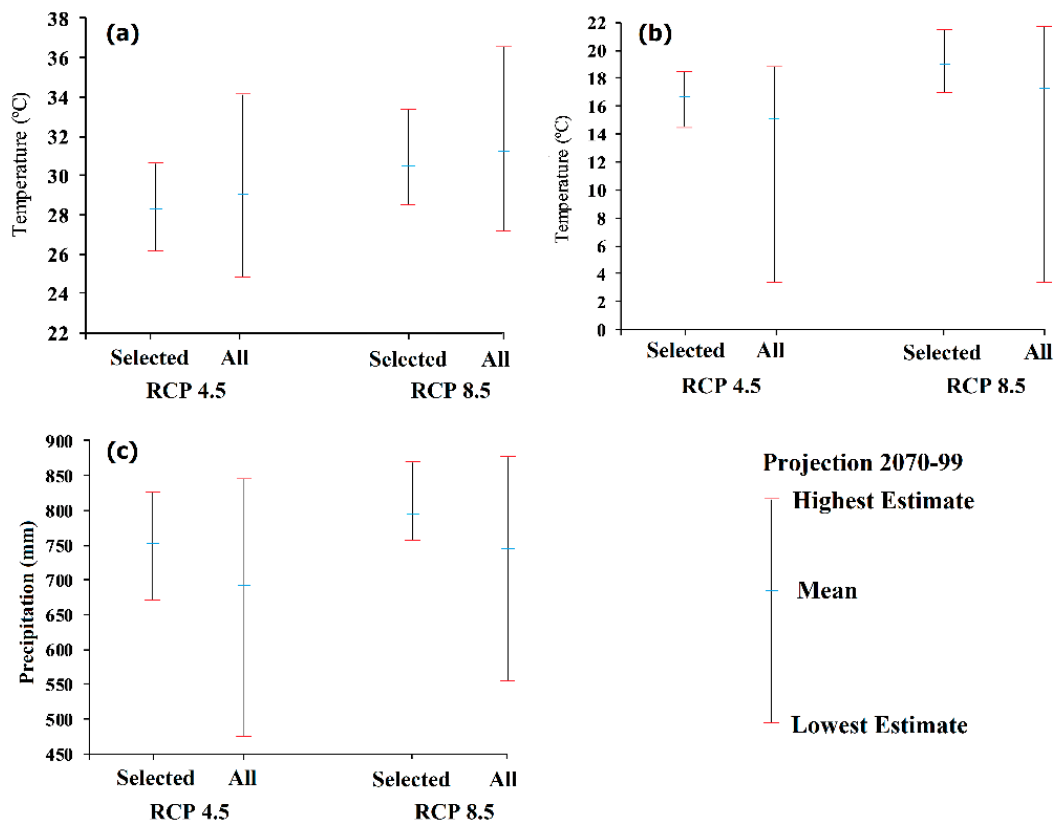
score for a certain climate variable, models are considered incapable and undesirable to project climate change in Pakistan.

**Table 5.** The GCMs received the highest scores in replicating temperature and precipitation at all the grid points over Pakistan. The selected GCMs are shown in bold fonts.

GCM	Max Temperature	Min Temperature	Precipitation
ACCESS1-0	1.7	0.0	2.5
<b>ACCESS1-3</b>	<b>6.8</b>	<b>9.7</b>	<b>3.1</b>
bcc-csm1-1	0.0	0.0	1.6
bcc-csm1-1-m	0.0	0.0	1.2
BNU-ESM	2.7	0.0	1.4
CanESM2	2.2	3.1	0.0
CCSM4	1.7	2.8	0.0
<b>CESM1-BGC</b>	<b>3.5</b>	<b>4.1</b>	<b>4.2</b>
<b>CMCC-CM</b>	<b>5.6</b>	<b>6.5</b>	<b>2.3</b>
CMCC-CMS	0.0	3.0	1.2
CNRM-CM5	0.0	0.0	4.0
CSIRO-Mk3-6-0	14.3	10.7	0.0
EC-EARTH	0.0	0.0	11.3
FGOALS-g2	0.0	2.9	0.0
GFDL-CM3	3.4	1.8	0.0
GFDL-ESM2G	0.0	0.0	0.0
GFDL-ESM2M	0.0	0.0	0.0
GISS-E2-R	0.0	0.0	1.4
<b>HadGEM2-CC</b>	<b>1.8</b>	<b>3.3</b>	<b>5.8</b>
<b>HadGEM2-ES</b>	<b>3.0</b>	<b>4.3</b>	<b>3.8</b>
INMCM4	0.0	0.0	0.0
IPSL-CM5A-LR	0.0	0.0	0.0
IPSL-CM5A-MR	3.9	0.0	0.0
IPSL-CM5B-LR	0.0	0.0	0.0
<b>MIROC5</b>	<b>1.4</b>	<b>1.8</b>	<b>11.9</b>
MIROC-ESM	0.0	0.0	3.3
MIROC-ESM-CHEM	0.0	0.0	2.4
MPI-ESM-LR	11.4	6.4	0.0
MPI-ESM-MR	0.0	2.5	0.0
MRI-CGCM3	2.7	4.9	0.0
NorESM1-M	6.3	4.6	0.0

#### 4.5. Future Projections using GCM Ensemble

The selected GCMs are used to project the future changes of temperature and precipitation in Pakistan. The projections obtained from the selected GCMs ensemble are compared with the ensemble of all the 31 GCMs. The projections of the maximum and minimum values of temperature and precipitation during the last part of this century (2070–2099) using the selected GCMs and all 31 GCMs are shown in Figure 7 based on selected RCP scenarios (RCP4.5 and RCP8.5). A significant increase of projection range estimated by all GCMs for all the cases is found in Figure 7. The ensemble mean of the projected temperature changes shifts systematically to the value higher when compared to the full ensemble mean values. Furthermore, the selected best models indicate that the maximum temperatures and precipitation (Figure 7b,c) do not include well-below average outlier models of the full ensemble. On the other hand, a higher increase of both temperature and precipitation is expected under the condition of RCP8.5 compared to RCP4.5.



**Figure 7.** The projections of (a) maximum temperature; (b) minimum temperature; and (c) precipitation during 2070–2099 obtained using the selected GCMs and all GCMs.

## 5. Conclusions

The performance of CMIP5 GCMs in simulating temperature and precipitation is assessed using multiple gridded datasets. In order to understand the influence of gridded data in the selection of GCMs and to identify the common GCMs that perform well for different gridded data and climatic variables. It is known from the results that the performance of GCMs varies according to different gridded data and climatic variables. A trade-off is made for the selection of GCMs based on their performance for different gridded datasets and climatic variables. Six GCMs namely, ACCESS1-3, CESM1-BGC, CMCC-CM, HadGEM2-CC, HadGEM2-ES and MIROC5 are found to be the most suitable for projection of climate of Pakistan according to their performance in reconstructing the historical climate for the period (1961–2005). The selected GCMs are found to be different from those found by [10]. This is due to the selection of GCMs based on different set of gridded data. The results emphasize the use of different gridded data in selection of GCMs to avoid uncertainty of selection. The GCMs are selected in this study solely based on their performances in simulating historical temperature and precipitation. Future projections of GCMs can be considered along with their past performance in the selection of GCMs.

**Author Contributions:** N.K. and S.S. designed the research and wrote the manuscript. M.S. and K.A. provided technical help with the model design and important analysis. T.I. proposed important additions to the manuscript and N.N. gave critical views on the manuscript for further improvement.

**Funding:** This research was supported by NRF-2015R1D1A1A02059944 through the National Research Foundation of Korea (NRF), and funded by the Ministry of Education, Science and Technology.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yip, S.; Ferro, C.A.; Stephenson, D.B.; Hawkins, E. A simple, coherent framework for partitioning uncertainty in climate predictions. *J. Clim.* **2011**, *24*, 4634–4643. [[CrossRef](#)]
2. Maraun, D.; Wetterhall, F.; Ireson, A.; Chandler, R.; Kendon, E.; Widmann, M.; Brienen, S.; Rust, H.; Sauter, T.; Themeßl, M. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* **2010**, *48*. [[CrossRef](#)]
3. Lutz, A.F.; ter Maat, H.W.; Biemans, H.; Shrestha, A.B.; Wester, P.; Immerzeel, W.W. Selecting representative climate models for climate change impact studies: An advanced envelope-based selection approach. *Int. J. Climatol.* **2016**, *36*, 3988–4005. [[CrossRef](#)]
4. Salman, S.A.; Shahid, S.; Ismail, T.; Ahmed, K.; Wang, X.-J. Selection of climate models for projection of spatiotemporal changes in temperature of Iraq with uncertainties. *Atmos. Res.* **2018**, *213*, 509–522. [[CrossRef](#)]
5. Pour, S.H.; Shahid, S.; Chung, E.-S.; Wang, X.-J. Model output statistics downscaling using support vector machine for the projection of spatial and temporal changes in rainfall of Bangladesh. *Atmos. Res.* **2018**, *213*, 149–162. [[CrossRef](#)]
6. Mote, P.; Brekke, L.; Duffy, P.B.; Maurer, E. Guidelines for constructing climate scenarios. *Eos Trans. Am. Geophys. Union* **2011**, *92*, 257–258. [[CrossRef](#)]
7. Srinivasa Raju, K.; Nagesh Kumar, D. Ranking general circulation models for India using TOPSIS. *J. Water Clim. Chang.* **2015**, *6*, 288–299. [[CrossRef](#)]
8. Warszawski, L.; Frieler, K.; Huber, V.; Piontek, F.; Serdeczny, O.; Schewe, J. The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3228–3232. [[CrossRef](#)]
9. Hidalgo, H.G.; Das, T.; Dettinger, M.; Cayan, D.; Pierce, D.; Barnett, T.; Bala, G.; Mirin, A.; Wood, A.; Bonfils, C. Detection and attribution of streamflow timing changes to climate change in the western United States. *J. Clim.* **2009**, *22*, 3838–3855. [[CrossRef](#)]
10. Raju, K.S.; Sonali, P.; Kumar, D.N. Ranking of CMIP5-based global climate models for India using compromise programming. *Theor. Appl. Climatol.* **2017**, *128*, 563–574. [[CrossRef](#)]
11. Ahmed, K.; Shahid, S.; Ali, R.; Harun, S.; Wang, X. Evaluation of the performance of gridded precipitation products over Balochistan Province, Pakistan. *Désalin. Water Treat.* **2017**, *79*, 73–86. [[CrossRef](#)]
12. Talavera, L. An evaluation of filter and wrapper methods for feature selection in categorical clustering. In *Advances in Intelligent Data Analysis VI*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 440–451.
13. Maxino, C.; McAvaney, B.; Pitman, A.; Perkins, S. Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation. *Int. J. Climatol. J. R. Meteorol. Soc.* **2008**, *28*, 1097–1112. [[CrossRef](#)]
14. Perkins, S.; Pitman, A.; Holbrook, N.; McAnaney, J. Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Clim.* **2007**, *20*, 4356–4376. [[CrossRef](#)]
15. Knutti, R.; Masson, D.; Gettelman, A. Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.* **2013**, *40*, 1194–1199. [[CrossRef](#)]
16. Yokoi, S.; Takayabu, Y.N.; Nishii, K.; Nakamura, H.; Endo, H.; Ichikawa, H.; Inoue, T.; Kimoto, M.; Kosaka, Y.; Miyasaka, T. Application of cluster analysis to climate model performance metrics. *J. Appl. Meteorol. Climatol.* **2011**, *50*, 1666–1675. [[CrossRef](#)]
17. Jiang, X.; Waliser, D.E.; Xavier, P.K.; Petch, J.; Klingaman, N.P.; Woolnough, S.J.; Guan, B.; Bellon, G.; Crueger, T.; DeMott, C. Vertical structure and physical processes of the Madden–Julian oscillation: Exploring key model physics in climate simulations. *J. Geophys. Res. Atmos.* **2015**, *120*, 4718–4748. [[CrossRef](#)]
18. Min, S.K.; Hense, A. 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.* **2006**, *33*. [[CrossRef](#)]
19. Afshar, A.A.; Hasanzadeh, Y.; Besalatpour, A.; Pourreza-Bilondi, M. Climate change forecasting in a mountainous data scarce watershed using CMIP5 models under representative concentration pathways. *Theor. Appl. Climatol.* **2017**, *129*, 683–699. [[CrossRef](#)]
20. Xuan, W.; Ma, C.; Kang, L.; Gu, H.; Pan, S.; Xu, Y.-P. Evaluating historical simulations of CMIP5 GCMs for key climatic variables in Zhejiang Province, China. *Theor. Appl. Climatol.* **2017**, *128*, 207–222. [[CrossRef](#)]

21. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [[CrossRef](#)]
22. Kharin, V.V.; Zwiers, F.W.; Zhang, X. Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations. *J. Clim.* **2005**, *18*, 5201–5223. [[CrossRef](#)]
23. Gleckler, P.J.; Taylor, K.E.; Doutriaux, C. Performance metrics for climate models. *J. Geophys. Res. Atmos.* **2008**, *113*. [[CrossRef](#)]
24. Johnson, F.; Sharma, A. Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments. *J. Clim.* **2009**, *22*, 4373–4382. [[CrossRef](#)]
25. Fu, G.; Charles, S.P.; Kirshner, S. Daily rainfall projections from general circulation models with a downscaling nonhomogeneous hidden Markov model (NHMM) for south-eastern Australia. *Hydrol. Process.* **2013**, *27*, 3663–3673. [[CrossRef](#)]
26. Rupp, D.E.; Abatzoglou, J.T.; Hegewisch, K.C.; Mote, P.W. Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. *J. Geophys. Res. Atmos.* **2013**, *118*, 10884–10906. [[CrossRef](#)]
27. Gu, H.; Yu, Z.; Wang, J.; Wang, G.; Yang, T.; Ju, Q.; Yang, C.; Xu, F.; Fan, C. Assessing CMIP5 general circulation model simulations of precipitation and temperature over China. *Int. J. Climatol.* **2015**, *35*, 2431–2440. [[CrossRef](#)]
28. Reichler, T.; Kim, J. How well do coupled models simulate today's climate? *Bull. Am. Meteorol. Soc.* **2008**, *89*, 303–312. [[CrossRef](#)]
29. Witten, I.H.; Paynter, G.W.; Frank, E.; Gutwin, C.; Nevill-Manning, C.G. KEA: Practical Automated Keyphrase Extraction. In *Design and Usability of Digital Libraries: Case Studies in the Asia Pacific*; IGI Global: Hershey, PA, USA, 2005; pp. 129–152.
30. Khan, N.; Shahid, S.; Ismail, T.; Ahmed, K.; Nawaz, N. Trends in heat wave related indices in Pakistan. *Stoch. Environ. Res. Risk Assess.* **2018**. [[CrossRef](#)]
31. Ahmed, K.; Shahid, S.; Nawaz, N. Impacts of climate variability and change on seasonal drought characteristics of Pakistan. *Atmos. Res.* **2018**, *214*, 364–374. [[CrossRef](#)]
32. Khan, N.; Shahid, S.; bin Ismail, T.; Wang, X.-J. Spatial distribution of unidirectional trends in temperature and temperature extremes in Pakistan. *Theor. Appl. Climatol.* **2018**, 1–15. [[CrossRef](#)]
33. Rohde, R.; Muller, R.; Jacobsen, R.; Muller, E.; Perlmutter, S.; Rosenfeld, A.; Wurtele, J.; Groom, D.; Wickham, C. A new estimate of the average Earth surface land temperature spanning 1753 to 2011. *Geoinform. Geostat. Overv.* **2013**. [[CrossRef](#)]
34. Xie, P.; Chen, M.; Shi, W. CPC unified gauge-based analysis of global daily precipitation. In Proceedings of the Preprints, 24th Conference on Hydrology, Atlanta, GA, USA, 17–21 January 2010.
35. Sheffield, J.; Goteti, G.; Wood, E.F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Clim.* **2006**, *19*, 3088–3111. [[CrossRef](#)]
36. Chaney, N.; Herman, J.; Reed, P.; Wood, E. Flood and drought hydrologic monitoring: The role of model parameter uncertainty. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 3239–3251. [[CrossRef](#)]
37. Yatagai, A.; Arakawa, O.; Kamiguchi, K.; Kawamoto, H.; Nodzu, M.I.; Hamada, A. A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Sola* **2009**, *5*, 137–140. [[CrossRef](#)]
38. Yatagai, A.; Kamiguchi, K.; Arakawa, O.; Hamada, A.; Yasutomi, N.; Kitoh, A. APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1401–1415. [[CrossRef](#)]
39. Palazzi, E.; Von Hardenberg, J.; Provenzale, A. Precipitation in the Hindu–Kush Karakoram Himalaya: Observations and future scenarios. *J. Geophys. Res. Atmos.* **2013**, *118*, 85–100. [[CrossRef](#)]
40. Berkeley Earth Surface Temperature. Available online: <http://www.berkeleyearth.org> (accessed on 1 June 2018).
41. Climate Prediction Centre. Available online: <http://www.cpc.ncep.noaa.gov/> (accessed on 1 June 2018).
42. Asian Precipitation—Highly Resolved Observational Data Integration toward Evaluation, (Monsoon Asia). Available online: <http://www.chikyu.ac.jp/precip/english/products.html> (accessed on 1 June 2018).
43. Princeton Global Meteorological Forcing Dataset. Available online: <http://www.hydrology.princeton.edu> (accessed on 1 June 2018).
44. Intergovernmental Panel on Climate Change. Available online: [http://www.ipcc-data.org/sim/gcm\\_monthly/AR5/Reference-Archive.html](http://www.ipcc-data.org/sim/gcm_monthly/AR5/Reference-Archive.html) (accessed on 1 June 2018).
45. Wang, L.; Ranasinghe, R.; Maskey, S.; van Gelder, P.M.; Vrijling, J. Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: A case study of the Huai River Basin, China. *Int. J. Climatol.* **2016**, *36*, 145–164. [[CrossRef](#)]

46. Novaković, J. Toward optimal feature selection using ranking methods and classification algorithms. *Yugosl. J. Oper. Res.* **2016**, *21*, 119–135. [[CrossRef](#)]
47. Singh, B.; Kushwaha, N.; Vyas, O.P. A feature subset selection technique for high dimensional data using symmetric uncertainty. *J. Data Anal. Inf. Process.* **2014**, *2*, 95–105. [[CrossRef](#)]
48. Kannan, S.S.; Ramaraj, N. A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm. *Knowl.-Based Syst.* **2010**, *23*, 580–585. [[CrossRef](#)]
49. Shreem, S.S.; Abdullah, S.; Nazri, M.Z.A. Hybrid feature selection algorithm using symmetrical uncertainty and a harmony search algorithm. *Int. J. Syst. Sci.* **2016**, *47*, 1312–1329. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).