

# Procedure for selecting GCM datasets for climate risk assessment

Chia-Yu Lin and Ching-Pin Tung\*

*Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan*

## Article history:

Received 18 November 2015

Revised 7 June 2016

Accepted 14 June 2016

## Keywords:

GCMs, Climate risk, Adaptation, Global warming, Climate change

## Citation:

Lin, C. Y. and C. P. Tung, 2017: Procedure for selecting GCM datasets for climate risk assessment. *Terr. Atmos. Ocean. Sci.*, 28, 43-55, doi: 10.3319/TAO.2016.06.14.01(CCA)

## ABSTRACT

General Circulation Models (GCMs) are indispensable tools to project future climate. It is not realistic or necessary to use all GCM datasets when assessing climate risks and building adaptive capacity. Thus, a rational procedure for selecting GCM datasets is needed. It is also required to classify weather stations into climate zones and then suggest a suitable list of GCM datasets to avoid weather stations with similar climate patterns but using different GCM datasets. The purpose of this study is to establish a process for selecting GCM datasets for a region. The process consists of climate zonation, applicability ranking, and a model similarity check. Principal component analysis (PCA) and cluster analysis are used to classify regional weather stations into climate zones. The weighted average ranking (WAR) method and merit point system (DPS) are then used to rank the GCM performance using CMIP5 (Coupled Model Intercomparison Project Phase 5) datasets. The GCM family tree is then applied to screen out highly similar GCMs before generating a GCM suggestion list. Taiwan is chosen as the study area for this investigation. Taiwan receives monthly mean precipitation data from 25 weather stations. The weather stations were clustered into ten climate zones with different GCM datasets suggested for each zone. The top five GCM datasets suggested for Taiwan by the WAR method are HadGEM2-AO, CESM1-CAM5, CCSM4, MIROC5, and GISS-E2-R while those suggested by the DPS method are CSIRO-Mk3-6-0, HadGEM2-AO, CESM1-CAM5, MIROC5, and CCSM4. The GCM selection process presented in this study is applicable to other regions to assist users in finding GCM datasets suitable for their research.

## 1. INTRODUCTION

The increasing number of disasters resulting in significant loss of life and wealth has drawn much attention in recent years. Climate change is predicted to bring more extreme weather events and natural disasters, making climate risk assessment, and adaptive capacity building crucial in reducing the adverse effects of future climate change. Climate risk assessment depends heavily on General Circulation Models (GCMs) predictions to predict future climate. There are many GCM datasets available; however, it is not realistic or necessary to use all available datasets in climate risk assessments.

Numerous GCMs have been developed, continuously improved and their outputs have been provided in CMIP5 (Coupled Model Intercomparison Project Phase 5). However, many unanswered questions such as “Which GCM is the

most suitable?” or “How should a proper GCM be selected?” remain as the complicated mechanism behind GCMs increases the understanding gap for the majority of end users. GCM outputs are widely used in climate risk assessment studies although using the outputs from inappropriate GCMs may not only lead to diverse results, but also increase the uncertainty in developing adaptation strategies. With the large number of datasets available in CMIP5, selecting appropriate GCM datasets has become an important task.

While GCMs are developed for global scale simulations and provide outputs for grid points, climate risk assessment studies on hydrology, agriculture, and environmental sectors, etc. are mostly at a basin scale and may have several weather stations present in a study site. Risk assessment often uses GCM projections to derive climate scenarios, after which weather data at a finer temporal scale is used for model assessment. Researchers often select GCM outputs from the nearest grid point for their study areas (Wilby and

\* Corresponding author  
E-mail: [cptung@ntu.edu.tw](mailto:cptung@ntu.edu.tw)

Wigley 1997; Schmidli et al. 2006). If weather stations in a site are within the same climate pattern but use GCM outputs from different grids, the rationality of the risk assessment may be questioned. Therefore, it is better to classify weather stations into climate zones ensuring that weather stations in the same zone use the same GCM output datasets.

The Köppen climate classification (Köppen 1936) was the first quantitative climate classification system. The Köppen climate classification divides the world's climate into five main clusters, with each cluster containing several subtypes (Peel et al. 2007). Kottek et al. (2006) presented a digital Köppen-Geiger world map, an update of the Köppen-Geiger world map released in 1961. The Trewartha method (Trewartha 1954) is also a climate classification system, based on a modified version of the Köppen climate classification. Wan (1973) classified Taiwan and its associated islands into 12 climate zones using the Köppen climate classification. Chiou et al. (2004) applied the Trewartha climate classification to Taiwan, also describing 12 climate zones. Baker et al. (2010) used the Trewartha classification to quantify the magnitude of change between historical and future ecoregions in China. Principal component analysis (PCA) is often used in climate classification research (Lorenz 1956; Kutzbach 1967) and was used by Wu and Chen (1993) to classify climate patterns for Taiwan. Monthly mean temperature and precipitation collected from 16 weather stations for the period 1942 - 1991 were used. According to Wu and Chen's research, the climate pattern in Taiwan can be divided into seven zones (Wu and Chen 1993). Liu (2010) also used PCA to determine climate zonation for Taiwan. With a smaller dataset of monthly mean precipitation from 18 weather stations for the period of 1961 - 1990, he divided Taiwan's climate into nine climate zones.

PCA is used in zonal classification and also in seasonal classification. Sadiq (2011) applied PCA to Karachi City in Pakistan. By analyzing diurnal data from 1981 - 2010 for precipitation, temperature, relative humidity, cloud amount, and wind speed, Sadiq classified this region's climate into six seasons. In this study, PCA is used to classify the climate patterns of a region.

Four criteria have been suggested for determining which GCM outputs should be used in climate risk assessment: vintage, resolution, validity, and results representations (Feenstra et al. 1998). The Intergovernmental Panel on Climate Change (IPCC 2001) further refined this definition:

- (1) Vintage: with new knowledge and more complete procedures and feedbacks, recent GCMs tend to have better reliability.
- (2) Resolution: enhanced GCM spatial resolution is likely to increase climate representation performance.
- (3) Validity: Based on the assumption that CGMs simulate the present-day climate most faithfully, a consistent exemplification of future climate can also be produced.
- (4) Results representation: choose GCMs with diverse key

variable output that can demonstrate a range of climate change.

Of the above-mentioned criteria, emphasis on vintage means that outputs new GCMs are often used as a reference for most climate risk assessment studies. In addition to this, new models tend to have higher resolution. For example, while half of the average atmospheric model resolution is finer than 1.3° in CMIP5, only one model in CMIP3 meets the same standard. In contrast, more than half of the CMIP3 models have a resolution coarser than 1° while only 2 CMIP5 models have a latitudinal resolution greater than this criterion (Taylor et al. 2012). Although there is processing cost associated with high-resolution climate simulation, these models are definitely ideal tools for climate system study (Gibelin and Déqué 2003).

Along with vintage and resolution, validity is commonly used as GCM dataset selection criterion. Suppiah et al. (2007) calculated the root mean square error (RMSE) and correlation coefficient between observations and model projections for the 1961 - 1990 period for temperature, rainfall, and sea level pressure. They also applied the demerit point system (DPS) as the basis for GCM selection (Whetton et al. 2005; Suppiah et al. 2007; Collier et al. 2011; Evans et al. 2014). In their study of the Xindian River Basin, Dahan River Basin, and Touchien River Basin in Taiwan, Lien et al. (2013) applied the DPS and weighted average ranking method (WAR) (Bentley and Wakefield 1998). Lien et al. (2013) also used the RMSE and the rainfall correlation coefficient as the evaluation criteria for suitable GCMs. However, as their research did not consider differences in climatic zonation, different weather stations within the same watershed may use different GCM datasets.

Masson and Knutti (2011) proposed the family tree concept for GCMs to ensure that they are representative. Their study indicated that some GCMs are similar for the following reasons: they originated from the same research institution; they consider the same atmospheric component, or they pass on source code in the process of producing a new model version. Their study used Kullback-Leibler divergence (KL divergence) (Kullback 1968) to quantify the divergence between each GCM. They generated the family tree from different models using hierarchical clustering analysis. Knutti et al. (2013) also used KL divergence to analyze CMIP3 and CMIP5 models and constructed a GCM family tree from the CMIP3 and CMIP5 datasets. They used the MSE (mean squared error) of the average temperature and average rainfall as a measure of similarity between the models. In the GCM family tree, models in the same branch are more similar than models on different branches. In this study, the family tree developed by Knutti et al. (2013) is used to screen out some GCMs with high similarity as a means of increasing the GCM suggestion list representation.

The purpose of this study is to propose a procedure for selecting GCM projection datasets based on four steps: (1)

Climate zonation - using the PCA and hierarchical cluster analysis to divide a region into several climate zones. (2) GCM ranking for a single weather station using the WAR method and DPS separately to rank all GCMs available for a given weather station. (3) GCM ranking for climate zones, which combines the results from steps 1 and 2 to calculate ranks of GCMs for each climate zone (zonal rank). (4) The zonal GCM suggestion list determination, which further uses the GCM family tree to remove highly similar models. Detailed methodology and results of a case study for Taiwan are described in the following sections.

## 2. MATERIALS AND METHODS

Figure 1 outlines the proposed procedure for selecting the GCM datasets. Step 1 is data collection, including historical weather records and GCM datasets. Step 2 is divided into 2 sections: (a) where a PCA and cluster analysis are used to classify climate zones; and (b) where the rank of the GCMs for each weather station is determined using the WAR method and DPS. Step 3 takes the climate zonation and GCM ranking results for a single weather station to determine the GCM rank for each climate zone. In step 4, the GCM family tree is used as a similarity check to generate the final suggestion list for each climate zone. Detailed descriptions of the application of these methods to our study are described sequentially below.

### Step 1. Data Collection

Historical observed weather data and GCM projections were collected for the baseline period of 1986 - 2005. All of the weather data for the 25 weather stations shown in Fig. 2 were obtained from the Taiwan Central Weather Bureau, including monthly mean precipitation data. GCM outputs from CMIP5 were provided by TCCIP (Taiwan Climate Change

Projection and Information Platform) with only GCMs that had outputs containing all four RCP (Representative Concentration Pathways) scenarios (Table 1) considered as suitable candidates for analysis.

### Step 2a. Climate Zonation

Step 2a is climate zonation, which classifies weather stations into several climate zones using PCA (Pearson 1901) and cluster analysis (Tryon 1939). Monthly mean precipitation from weather stations was used for PCA. Only the principal components where cumulative contribution rate exceeds 85% of each station were chosen. PC (principal component) scores of each station were then calculated. The PC scores of weather stations were used to determine the Euclidean distance between two stations. Cluster analysis was then applied to classify weather stations by distance. Weather stations were classified into several zones using a hierarchical clustering technique named Ward's method (Ward 1963). Ward's method uses a criterion based on the value of total within-cluster variance to select two clusters for merging.

### Step 2b. GCM Ranking for Single Weather Station

Evaluation of the applicability of 20 GCMs was done using three criteria, correlation coefficient, and normalized root mean square error (*NRMSE*) for wet season and dry season. Generally, the higher the correlation coefficient value between GCM outputs and observed weather data, the better the monthly mean precipitation trend simulation. On the other hand, *NRMSE* represents deviation between GCM outputs and observations where lower *NRMSE* also means smaller bias. Only GCM outputs from the nearest grid point to the weather station were used in the correlation coefficient and *NRMSE* calculations. During the evaluation process, *NRMSE* for the wet (from November to April) and

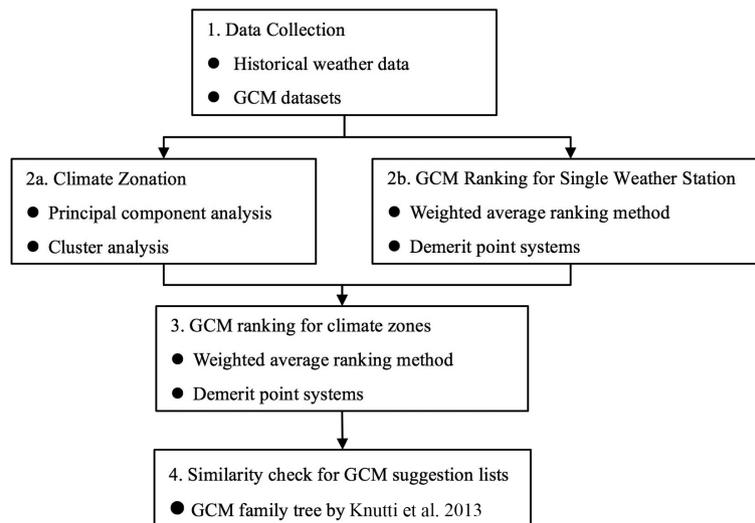


Fig. 1. Procedures to select GCM datasets.



Fig. 2. Study area and spatial distributions of weather stations. (Color online only)

Table 1. GCM datasets for analysis.

Modeling Center	Model	Institution
BCC	BCC-CSM1.1 BCC-CSM1.1(m)	Beijing Climate Center, China Meteorological Administration
CSIRO-QCCCE	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with the Queensland Climate Change Centre of Excellence
FIO	FIO-ESM	The First Institute of Oceanography, SOA, China
IPSL	IPSL-CM5A-LR IPSL-CM5A-MR	Institute Pierre-Simon Laplace
MIROC	MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
MIROC	MIROC-ESM MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MRI	MRI-CGCM3	Meteorological Research Institute
NASA GISS	GISS-E2-H GISS-E2-R	NASA Goddard Institute for Space Studies
NCAR	CCSM4	National Center for Atmospheric Research
NCC	NorESM1-M NorESM1-ME	Norwegian Climate Centre
NIMR/KMA	HadGEM2-AO	National Institute of Meteorological Research/Korea Meteorological Administration
NOAA GFDL	GFDL-CM3 GFDL-ESM2G GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory
NSF-DOE-NCAR	CESM1-CAM5	National Science Foundation, Department of Energy, National Center for Atmospheric Research

dry season (from May to October) precipitation data were calculated individually. The *NRMSE* can be calculated using Eq. (1) as Lien et al. (2013).

$$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{X_{sim,i} - X_{obs,i}}{X_{obs,i}} \right)^2} \quad (1)$$

Where  $X_{sim,i}$  is the  $i$ th normalized data for simulated monthly mean precipitation,  $X_{obs,i}$  is the  $i$ th-normalized data for the observed precipitation, and  $N$  is the number of data.

Two methods were employed to rank GCMs in this study, WAR method and the DPS. In the WAR method, each GCM was sorted on three criteria. An average rank for each GCM was then calculated. All GCMs were then sorted according to the average rank calculated using Eq. (2).

$$R_j = \frac{1}{N} \sum_{i=1}^N Rank_i \quad (2)$$

Where  $R_j$  is the average rank of the  $j$ th GCM,  $Rank_i$  is the rank corresponding to the  $i$ th criterion, and  $N$  is the number of criteria. In this study, the range of  $j$  is one to twenty five, and  $N$  is equal to three.

The second appraisal method used is the DPS where a threshold is set on whether to accept a given GCM. If the performance does not exceed the threshold, the GCM will incur demerit points. All models are sorted by total demerit points after demerit points are summed for each GCM where the lower the total demerit points of a GCM, the better the model's performance. In this study, total demerit points of each GCM were calculated using the following equations:

$$D_j = \sum_{i=1}^3 C_i(x) \quad (3)$$

$$C_i(x) = \begin{cases} 1, & x \leq \text{threshold of correlation coefficient} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

for  $i = 1$

$$C_i(x) = \begin{cases} 1, & x > \text{threshold of } NRMSE \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

for  $i = 2, 3$

Where  $D_j$  is total number of demerit points for the  $j$ th GCM and  $C_i(x)$  is the number of demerit points for the  $i$ th criterion. Where there is only a single weather station, a GCM is suggested only if its total number of demerit points is zero. However, the DPS analysis results are strongly influenced by the threshold settings, which may be adjusted according to the research of Whetton et al. (2005), Suppiah et al. (2007), and Lien et al. (2013).

### Step 3. GCM Ranking for Climate Zones

GCM ranking for climate zones can be seen as the joint application of climate zonation and GCM ranking for a single weather station. Using cluster analysis procedures, a region can be divided into different climate zones with each zone containing one or multiple weather stations. For climate zones having a single weather station, the zonal GCM ranking will be the same as the weather station ranking located within this zone. However, if multiple weather stations are situated within a climate zone, the average GCM rank for all weather stations in the zone will be calculated using Eq. (6). All GCMs are then sorted based on their average ranks to determine the zonal rank.

$$ZR_{j,k} = \frac{1}{N_k} \sum_{i=1}^{N_k} R_{j,i} \quad (6)$$

Where  $ZR_{j,k}$  is the average rank of the  $j$ th GCM in the  $k$ th climate zone,  $N_k$  is the number of weather stations in the  $k$ th climate zone,  $R_{j,i}$  is the rank of  $j$ th GCM corresponding to the  $i$ th weather station in the  $k$ th climate zonation.

DPS can also be used to directly determine the zonal GCM ranking by summing the number of demerits for all weather stations within a climate zone and then sorting using the average number of demerit points.

$$ZD_{j,k} = \frac{1}{N_k} \sum_{i=1}^{N_k} D_{j,i} \quad (7)$$

Where  $ZD_{j,k}$  is the average number of demerit points for the  $j$ th GCM in the  $k$ th climate zone,  $N_k$  is the number of weather stations in  $k$ th climate zone, and  $D_{j,i}$  is the number of demerit points for  $j$ th GCM at the  $i$ th weather station in the  $k$ th climate zonation. For a climate zone with multiple weather stations, the average number of demerits for a GCM is not likely to be zero. Thus, if the average number of demerit points for a GCM in a climate zone is less than 1, the GCM will be retained in the suggestion list.

### Step 4. GCM Similarity Check by Family Tree

The zonal GCM ranking results using the WAR or DPS analysis are used to produce a GCM suggestion list for each climate zone. However, a suggestion list may not be entirely representative as some of the GCMs may be similar in their projections or different GCMs may be derived from the same parent model. To check the similarity of GCMs in the initial suggestion list the GCM family tree in Fig. 3 (simplified and redrawn by this research based on Knutti et al. 2013) is applied to identify the relationship between GCMs in the list. Figure 3 shows that for each GCM, there is a value for "divergence" or "D number". A GCM located on a branch with a larger value for D means it may be a derivative version of the GCM located on a branch with a smaller D number. Additionally, a GCM may be removed

from the suggestion list if there is another GCM with a lower rank on the same branch of the family tree.

### 3. RESULTS AND DISCUSSION

This study suggests a procedure that consists of climate zonation, applicability ranking, and a model similarity check proposed to produce the GCM selection lists for a region. To test the applicability of this procedure, the technique was applied to data from the area of Taiwan. Taiwan is an island in East Asia, covering an area of 36197 km<sup>2</sup>. The average annual precipitation for Taiwan is around 2510 mm with 78% of the precipitation falling during the wet season from May to October. As an initial check, this study calculated the correlation coefficients of monthly mean temperature and precipitation between observed weather data and GCM baseline projections to determine the temperature and precipitation applicability as selection criteria. The baseline period was chosen as the years 1986 through 2005. The results showed that GCMs' performance at projecting temperature is much better than projecting precipitation. The correlation coefficients for monthly mean temperature are higher than 0.99, and the difference in performance between GCMs was quite small compared to the large differences in projecting mean precipitation. As we are seeking to separate GCMs based on predictive performance, the GCM projections for precipitation were regarded as the criterion for initial GCM selection.

#### 3.1 Climate Zonation of Taiwan

In the climate zonation analysis for Taiwan, the monthly mean precipitation for the period of 1986 - 2005 was used. Table 2 lists the eigenvalue, percent variance and cumulative percentage given by the first three principal components. The PCA result shows that the first three principal components explain 96% of the variation in the data. Table 3 lists the eigenvectors for the first three principal components of data from January to December. PC scores of the first three PCs of 25 weather possible stations can be calculated using the monthly mean precipitation and eigenvectors. The Euclidean distance between stations was calculated using the PC scores from the 25 weather stations. Figure 3 is a cluster dendrogram of the weather stations using Ward's method. Taiwan was classified into 7 to 9 climate zones by Wu and Chen (1993) and Liu (2010) but since the total number of weather stations used in this research is greater than those used in previous studies, the maximal number ( $z$ ) of 10 climate zones was considered.

The zonation results from  $z = 6$  to  $z = (10)$  are displayed in Fig. 4. Among the 25 weather stations, Keelung was always classified in a climate zone, as it is a unique station. Alishan is also a special station and was placed in the same zone as Yushan and Sun Moon Lake at  $z = 6$ . However, it

was moved alone in a new climate zone after  $z$  increased from 6 - 7. When the number of climate zones increased from 7 - 8, Dongjidao and Penghu were separated to form a new zone. After  $z$  increased to 9, Dawu and Hengchun were further classified to form a new zone to better explain the effects of their geographic locations on the southern side of the central mountain range. Finally, when  $z$  increased to 10, only Suao was moved from its original zone to form a new zone implying ten zones may not be necessary and nine climate zones may still be appropriate.

It was noticed that Lanyu weather station was classified as the same zone as Tamsui, Pengjiayu, Hsinchu, Wuqi, Taipei, and Taichung, which are all located in northern Taiwan, except for Lanyu (Fig. 2). Therefore, Lanyu station was manually moved to the tenth zone in this study. Table 4 shows the climate zonation results and lists the weather stations in each zone.

#### 3.2 GCM Ranking for Single Weather Station

The WAR method and DPS analysis were applied to evaluate the GCM applicability here based on three criteria, allowing the GCM suggestion list for a weather station to be produced. The ranks were further used to determine the zonal ranking in the next section.

##### 3.2.1 Analysis of WAR Method

After the WAR method analysis, each GCM was assigned a rank in the range of 1 - 20. It is envisioned that even if a GCM is applicable for one weather station it does not imply that it will also perform well for another weather station. Figure 5 gives an example of this situation, showing the ranking results from the WAR method for three weather stations, including Taipei, Tainan, and Taitung. The HadGEM-AO model earned rank 1 at Taipei weather station (located in northern Taiwan), and also got rank 1 at Tainan weather station (located in southern Taiwan) however it only ranked 11 at Taitung weather station (located in eastern Taiwan). The WAR method produce a GCM suggestion for a given weather station, but different weather stations may have different GCM suggestion lists.

The ranks for all weather stations can be further statistically analyzed to select GCMs for a region, but it does not necessarily make it easier to identify which GCM is more applicable. When a given GCM was marked as ranks 1, 2, and 3, since 25 weather stations have been analyzed in this research, a GCM can gain rank 1 up to 25 times, as shown in Fig. 6. However, among the 20 GCMs that were analyzed in this study, only seven GCMs have earned rank 1. HadGEM2-AO earned rank 1 10 times, which means it has better performance than the other GCMs at most weather stations. The GISS-E2-R model received rank 1 five times, but received no ranks 2 or 3. Based on the statistical analysis, the

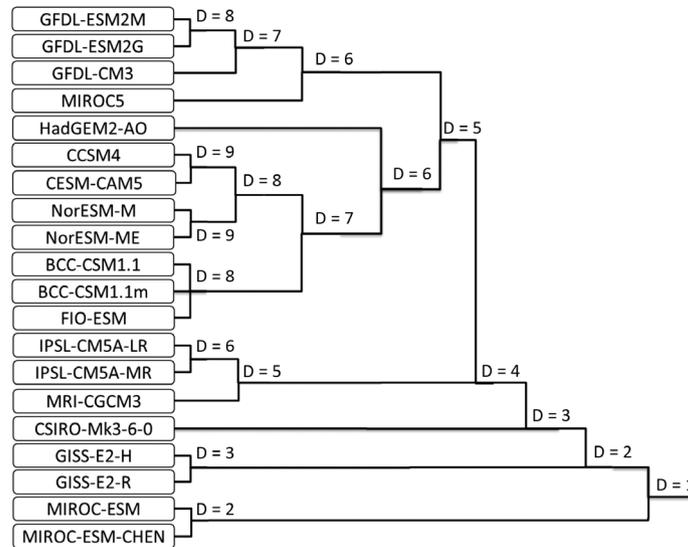


Fig. 3. GCM family tree from CMIP5 (simplified and redrawn in this research based on Knutti et al. 2013). A GCM located on a branch with a larger “D (number of divergence)” means that it might be the derivative version of GCM located on a branch with a smaller D.

Table 2. Eigenvalues and the proportion of variation explained by the first three principal components.

	Comp. 1	Comp. 2	Comp. 3
Eigenvalue	1356	427	130
Percentage of Variance	68%	21%	7%
Cumulative Percentage	68%	89%	96%

Table 3. Eigenvectors of the first three principal components.

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
Comp. 1	-0.25	-0.22	-0.15	-0.09	-0.08	0.05	0.09	0.00	-0.39	-0.58	-0.50	-0.32
Comp. 2	-0.02	-0.01	0.01	0.10	0.26	0.45	0.56	0.59	0.20	0.02	-0.07	-0.03
Comp. 3	-0.26	-0.37	-0.39	-0.39	-0.39	-0.10	0.08	0.06	0.42	0.31	-0.05	-0.16

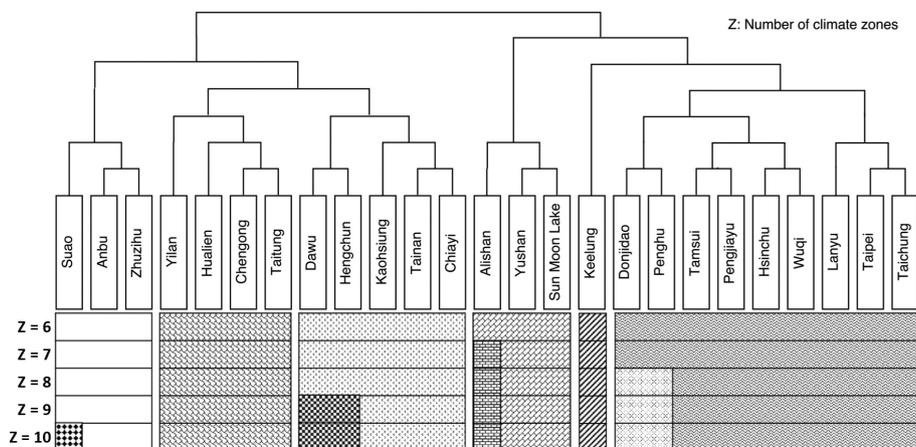


Fig. 4. Cluster dendrogram and zonation results of 25 weather stations. Under a number of climate zones (for example: k = 6), GCMs of the same shade belong to the same climate zone.

Table 4. Climate zonation of Taiwan

Climate Zone	Weather station(s)
North	Tamsui, Pengjiayu, Hsinchu, Wuqi, Taipei, Taichung
Northeast	Keelung
East	Yilan, Hualien, Chenggong, Taitung
Southwest	Chiayi, Tainan, Kaohsiung
South	Dawu, Hengchun
North mountain	Anbu, Zhuzihu, Suao
Central	Yushan, Sun Moon Lake
South mountain	Alishan
West Island	Dongjidao, Penghu
East Island	Lanyu

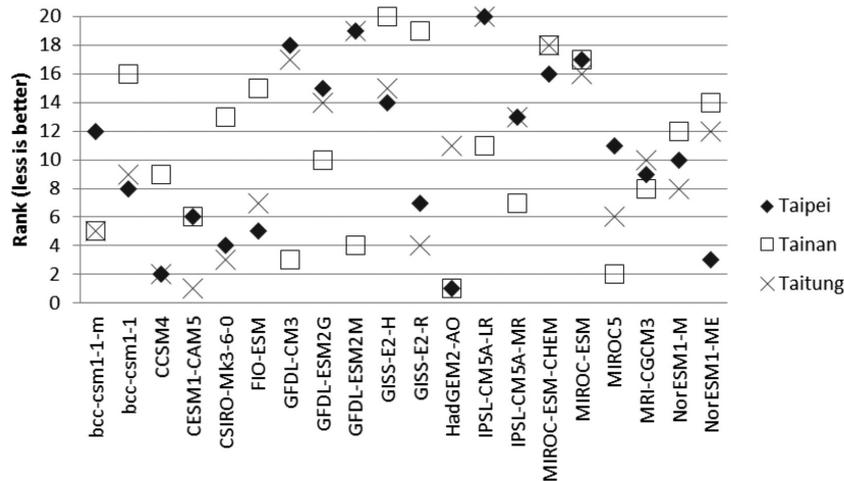


Fig. 5. GCMs ranking results of Taipei, Tainan, and Taitung weather stations by the WAR method.

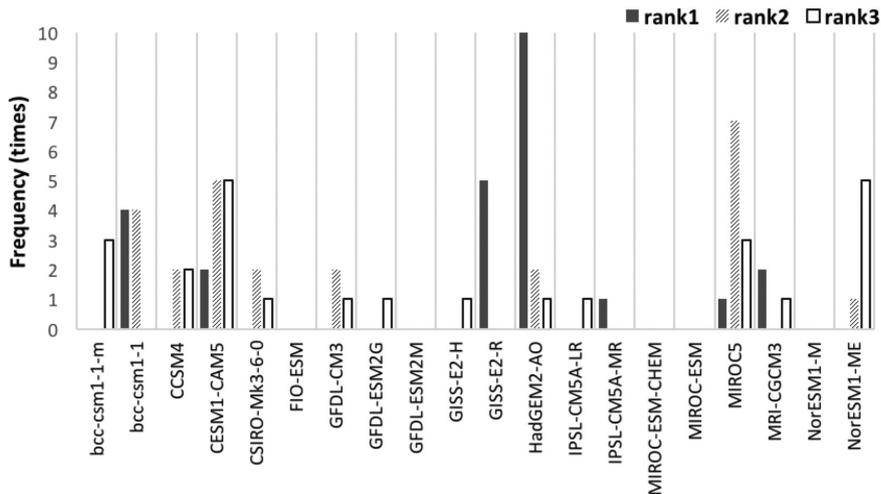


Fig. 6. Results of statistical analysis of frequency with GCMs marked as ranks 1, 2, and 3.

HadGEM2-AO model has better performance for the Taiwan area than other GCMs.

### 3.2.2 Analysis of DPS

In the DPS analysis, the correlation coefficient threshold was first set to 0.7, which means highly statistically correlated but after referring to previous studies, the *NRMSE* threshold was adjusted from 0.4 through 0.8. In the research of Whetton et al. (2005) and Suppiah et al. (2007), the RMSE threshold was set to 2.0 for mean sea-level pressure. Lien et al. (2013) set the precipitation threshold in the 0.4 - 0.7 range. The RMSE values are normalized. If the threshold is set to 2.0, most GCMs are able to easily meet the standard, and the DPS results will lack discrimination. In order to avoid this situation, the calculated *NRMSE* value between weather stations and GCMs should be considered. Since the *NRMSE* range in this study is about 0.3 - 0.9, the *NRMSE* threshold tested was from 0.4 - 0.8. Note that for the majority of GCMs there is a bottleneck when the threshold was set between 0.7 and 0.8. Setting the threshold to 0.7, most GCMs will not pass the test, making 0.8 the threshold to determine the number of GCMs that enter the suggestion list in this study. However, the threshold can be adjusted depending on the number of GCM datasets available or applicable to each climate zone.

While the WAR method produces a ranking for a single weather station, DPS will only exclude GCMs with a high number of demerit points. According to Eq. (3), if a GCM fails to pass the threshold for the three criteria, the model will incur demerit points. Figure 7 shows the demerit points for each GCM for Taipei, Tainan, and Taitung weather stations as examples. For Taipei station, only three GCMs (about 12%) (CSIRO-Mk3-6-0, MIROC5, NorESM1-M) passed the three thresholds. For Taitung and Tainan weather stations, 9 (about 36%) and 8 (about 32%) GCMs respectively could satisfy the three thresholds.

A GCM with zero demerit points is considered viable to become a candidate for a suggested model for a given weather station. However, with the thresholds of  $R = 0.7$  and  $NRMSE = 0.8$ , most of GCMs at Anbu, Zhuzihu, and Suao weather stations (all located in the North mountain zone) obtained demerit points and did not pass the DPS test. Figure 8 shows the correlation coefficients between the GCM projections and observations from the above weather stations. It is clear that only the CESM-CAM5 model passes the correlation coefficient threshold at Anbu and Zhuzihu stations and in some cases the correlation coefficients for some GCMs are negative. The same situation occurs at Keelung station, in the northeast zone. This situation shows that most GCMs cannot provide reasonable precipitation trends at some of the weather stations located in the mountainous northern and northeast area of Taiwan. On the other hand, stations such as Taitung and Tainan may have too many

possible model candidates. Different values for thresholds may be considered in different climate zones.

### 3.3 GCM Ranking for Climate Zone

With the GCM ranking result for a single weather station using the WAR method, the average rank of each GCM can be calculated using Eq. (6) to determine the zonal ranking and produce the zonal GCM suggestion list. The recommended GCMs for a climate zone should also reasonably reflect the characteristics and trend for the whole of Taiwan. Therefore, a GCM will be removed from the zonal GCM suggestion list if it has poor performance (rank greater than 10) for the whole of Taiwan area. Table 4 lists the Top 5 GCMs for each climate zone.

As shown in Table 5, HadGEM2-AO, CESM1-CAM5, and CCSM4 perform very well in most of the climate zones. According to the GCM ranks in different climate zones, it can be roughly observed that the behavior of the three GCMs may have some dissimilarity between areas in Taiwan. For instance, the HadGEM2-AO model performs well in the northern and southern zones, but does not have good performance in the eastern part of Taiwan. The CESM1-CAM5 model is good in the mountain and island areas of Taiwan, but performs poorly in the southern part of Taiwan. The overall performance of the CCSM4 model is worse than that of HadGEM2-AO and CESM1-CAM5, except in the central and southern part of Taiwan. In addition to the above three models, the MIROC5 model does not rank well in every climate zone. In the South and Central zones, the MIROC5 model produces more accurate projections than the other GCMs. Based on these results, a perfect GCM for all climate zones in Taiwan does not exist and the most applicable GCMs may be different for different climate zones. This does not impact the WAR method ability to easily find the most appropriate GCM datasets for each zone.

The DPS results for single weather stations can also be applied to zonal GCM selection. The average number of demerit points for GCMs within a climate zone are calculated using Eq. (7) and only a GCM with an average number of demerit points lower than 1 are retained in the suggestion list. This method does lead to quite a variety in the numbers of GCMs retained in each climate zone. In the Northeast and North mountain zones, all GCMs were excluded because no GCM could satisfy the 1 demerit point threshold. The reason for this is that no GCMs are able to reproduce the reasonable historical trend for precipitation for weather stations within the climate zone (see Fig. 8). Conversely, in some climate zones such as the South and West, the threshold number of demerit points seems to be too low and thus almost half of the GCMs are retained.

For the Northeast mountainous zones, the easiest way to allow more GCMs to be recommended for inclusion is to raise value of the average number of demerit points used

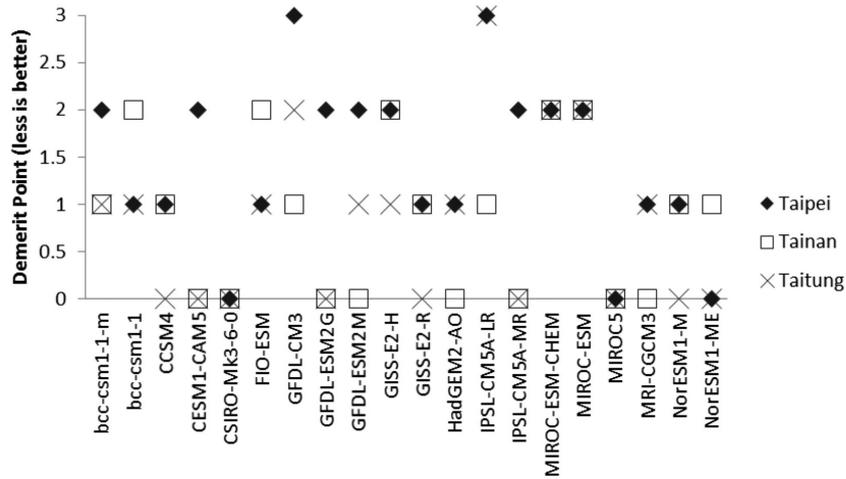


Fig. 7. Total demerit points of GCMs for Taipei, Tainan, and Taitung weather stations.

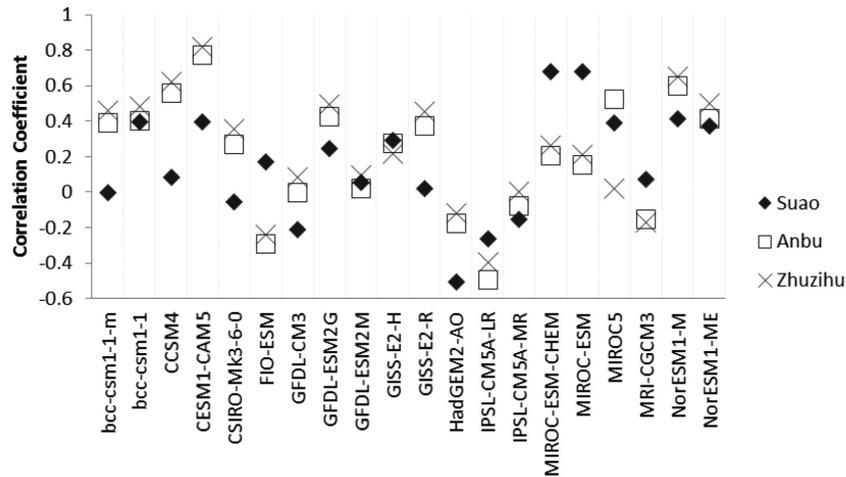


Fig. 8. Correlation analysis of GCMs for weather stations located in the North mountain zone. Only CESM-CAM5 model passes the threshold of correlation coefficient (0.7) at Anbu (0.77) and Zhuzihu (0.82) stations.

Table 5. Top five GCMs of each climate zonation of Taiwan by WAR method.

Rank Name	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
North	HadGEM2-AO	NorESM1-ME	CSIRO-Mk3-6-0	CCSM4	bcc-csm1.1m
Northeast	MRI-CGCM3	bcc-csm1.1	CESM1-CAM5	HadGEM2-AO	NorESM1-ME
East	CESM1-CAM5	GISS-E2-R	CCSM4	bcc-csm1.1	CSIRO-Mk3-6-0
Southwest	HadGEM2-AO	MIROC5	bcc-csm1.1m	CCSM4	CESM1-CAM5
South	MIROC5	GISS-E2-R	CCSM4	CSIRO-Mk3-6-0	HadGEM2-AO
North mountain	bcc-csm1.1	CESM1-CAM5	NorESM1-ME	HadGEM2-AO	MRI-CGCM3
Central	MIROC5	CCSM4	HadGEM2-AO	CESM1-CAM5	MRI-CGCM3
South mountain	HadGEM2-AO	CESM1-CAM5	MIROC5	MRI-CGCM3	CCSM4
West island	HadGEM2-AO	MIROC5	CESM1-CAM5	bcc-csm1.1m	CCSM4
East island	GISS-E2-R	CSIRO-Mk3-6-0	CESM1-CAM5	CCSM4	bcc-csm1.1m
Taiwan	HadGEM2-AO	CESM1-CAM5	CCSM4	MIROC5	GISS-E2-R

for exclusion. Even if a GCM passes into the recommendation list because of lowering the threshold standard, the real GCM performance does not change. On the other hand, for a climate zone with more than 5 recommended GCMs, an additional criterion is needed to maintain less than 5 GCMs in the suggestion list. The regional correlation coefficient for precipitation in this study was used as the additional criterion to screen out part of the GCMs. Table 6 lists the GCMs recommended for use in each climate zone based on the DPS.

Comparison between the zonal GCMs suggested by the WAR method (Table 5) and DPS (Table 6) shows that most recommended GCMs by DPS also earn good ranks using the WAR analysis. The major advantage of the WAR method is that it identifies relatively better models, even if the absolute GCM performance may not be optimal. Conversely, the DPS analysis excludes unsuitable GCMs but it may leave no possible GCMs for the suggestion list. Setting an appropriate threshold for each criterion in DPS a critical task.

### 3.4 Similarity Check Between Zonal GCM in the Suggestion List

In the previous section, two zonal GCM suggestion lists were produced using the WAR method and the DPS

analysis, respectively. In this section, the criteria for “representativeness” proposed by Feenstra et al. (1998) are further considered by checking the similarity of suggested GCMs. The GCM family tree presented by Knutti et al. (2013) was employed to determine the GCM positions although the two GCMs used in this study were not included in the family tree allowing it to be simplified and presented as Fig. 3. Two GCMs are considered similar if they appear in the same branch of the family tree meaning that only one should be chosen as a GCM dataset for risk assessment study. Twenty GCM datasets were used in this study to generate the GCM suggestion list for weather stations in Taiwan. The preliminary suggestion lists containing 10 GCMs are shown in Tables 5 and 6. These 10 GCMs in the suggestion lists were further confirmed in the GCM family tree.

During the similarity check process, the numbers of divergence (D) in Fig. 3 is the index to classify similar GCMs in the family tree. For example, if  $D = 4$  is used, then GCMs where  $D > 4$  in Fig. 3 are considered to be a group with high similarity (Fig. 9). Table 7 shows the GCM similarity check results using divergence numbers (D) from 7 - 9. Note that the group number of similar GCMs increases with the layer number. Under  $D = 7$ , 10 GCMs listed in Table 5 are clustered into 5 groups, with the biggest group containing 5 GCMs.

Table 6. Recommended GCMs of each climate zonation of Taiwan by DPS.

Climate Zonation	Recommend GCMs
North	CSIRO-Mk3-6-0, HadGEM2-AO
Northeast	none
East	CESM1-CAM5
Southwest	CESM1-CAM5, HadGEM2-AO, MIROC5, MRI-CGCM3
South	HadGEM2-AO
North mountain	none
Central	MIROC5, HadGEM2-AO, GFDL-ESM2G, NorESM-ME, CCSM4
South mountain	MIROC5, GFDL-ESM2M, HadGEM2-AO, MRI-CGCM3, GFDL-ESM2G
West island	MIROC5, GFDL-ESM2G, HadGEM2-AO, GFDL-ESM2G, CESM1-CAM5
East island	CCSM4, CSIRO-Mk3-6-0
Taiwan	CSIRO-Mk3-6-0, HadGEM2-AO, CESM1-CAM5, MIROC5, CCSM4

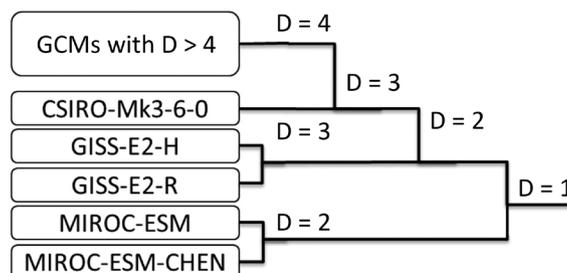


Fig. 9. Result of GCM similarity check with the GCM family tree in Fig. 3, and numbers of divergence (D) used as an index. In this example,  $D = 4$ , and GCMs with  $D > 4$  in Fig. 3 are considered as a group with high similarity.

Table 7. GCM similarity check results under different numbers of divergence (D).

D = 7	D = 8	D = 9
MIROC5	MIROC5	MIROC5
HadGEM2-AO	HadGEM2-AO	HadGEM2-AO
CCSM4 CESM-CAM5 NorESM-ME bcc-csm1.1 bcc-csm1.1(m)	CCSM4 CESM-CAM5, NorESM-ME bcc-csm1.1 bcc-csm1.1(m)	CCSM4 CESM-CAM5 NorESM-ME bcc-csm1.1 bcc-csm1.1(m)
MRI-CGCM3	MRI-CGCM3	MRI-CGCM3
CSIRO-Mk3-6-0	CSIRO-Mk3-6-0	CSIRO-Mk3-6-0
GISS-E2-R	GISS-E2-R	GISS-E2-R

The group number reaches maximum (8) under D = 9 leaving only two groups containing more than one GCM.

Based on these results, if both CCSM4 and CESM-CAM5 are recommended in a specific climate zone it is possible to screen out GCMs with lower ranks leaving the GCM with the highest rank in the suggestion list. The results shown in Table 7 can also be used to reduce the number of GCMs suggested by DPS. Of the five GCMs suggested in the West island zone, two GCMs in the suggestion lists (CESM1-CAM5 and CCSM4) are marked as having high similarity. Under this situation, the user can determine to use only CESM1-CAM5 or CCSM4.

#### 4. CONCLUSION

Climate risk assessment often requires GCM projections to derive climate scenarios. However, it is not realistic to use all datasets although many GCM datasets are available. It may also not be reasonable to randomly choose GCM datasets. This study proposed a procedure using climate zonation, applicability ranking and a model similarity check to select required GCM datasets. Taiwan was chosen as the study area to test the proposed procedure.

PCA and hierarchical cluster analysis are suitable tools to classify weather stations into several climate zones showing similar results as previous studies. However, as Fig. 4 shows, by placing the Lanyu station in the North zone, the combination of PCA and hierarchical cluster analysis cannot properly take into account the weather station location. To overcome this problem, other weather parameters, such as temperature, may be used to improve climate zonation.

The WAR method and DPS analysis were used to rank the applicability of GCMs. The results show that the WAR method can produce a GCM suggestion list for each climate zone. Nevertheless, for some specific weather stations located in Northeast Taiwan, some GCMs receive good ranks even if they cannot reasonably reproduce the historical monthly mean precipitation trend. On the other hand, the

DPS analysis removes non-applicable GCMs, but may result in an empty GCM suggestion list for a given climate zone. Using the same threshold of 1 demerit point for exclusion, some zones have 6 to 7 suggested GCMs while other zones have none. It is not easy to assign different demerit point thresholds for different climate zones. Thus, both suggestion lists from the WAR method and DPS analysis may be considered together in the future.

The procedure proposed in this study can be easily applied to other regions that contain a sufficient number of weather stations with adequate historical observations. The proposed procedure can help users' select suitable GCM datasets for their study areas. Additionally, the GCM family tree is a useful tool to check the similarity of GCMs and choose GCM datasets that are more representative. Unfortunately, new GCM versions may not be included in the GCM family tree developed by Knutti et al. (2013), indicating that an updated family tree would be very helpful for future study.

**Acknowledgements** This study is a partial outcome of the TaiCCAT (Taiwan Integrated Research Program on Climate Change Adaptation Technology) project. The authors would like to thank the financial support from the Ministry of Science and Technology, Taiwan (MOST 104-2621-M-002-002).

#### REFERENCES

- Baker, B., H. Diaz, W. Hargrove, and F. Hoffman, 2010: Use of the Köppen-Trewartha climate classification to evaluate climatic refugia in statistically derived ecoregions for the People's Republic of China. *Clim. Change*, **98**, 113-131, doi: 10.1007/s10584-009-9622-2. [Link]
- Bentley, P. J. and J. P. Wakefield, 1998: Finding acceptable solutions in the Pareto-optimal range using multiobjective genetic algorithms. In: Chawdhry, P. K., R. Roy, and R. K. Pant (Eds.), *Soft Computing in Engineering Design and Manufacturing*, Springer, London, 231-240, doi: 10.1007/978-1-4471-0427-8\_25. [Link]
- Chiou, C. R., Y. C. Liang, Y. J. Lai, and M. Y. Huang, 2004: A study of delineation and application of the climatic zones in Taiwan. *J. Taiwan Geograph. Inform. Sci.*, **1**, 41-62.
- Collier, M. A., S. J. Jeffrey, L. D. Rotstain, K. K. H. Wong, S. M. Dravitzki, C. Moeseneder, C. Hamalainen, J. I. Syktus, R. Suppiah, J. Antony, A. E. Zein, and M. Atif, 2011: The CSIRO-Mk3.6.0 Atmosphere-Ocean GCM: Participation in CMIP5 and data publication. 19<sup>th</sup> International Congress on Modelling and Simulation, Perth, Australia, 2691-2697. Available at <http://mssanz.org.au/modsim2011>.
- Evans, J. P., F. Ji, C. Lee, P. Smith, D. Argüeso, and L. Fita, 2014: Design of a regional climate modelling projection

- ensemble experiment – NARClIM. *Geosci. Model Dev.*, **7**, 621-629, doi: 10.5194/gmd-7-621-2014. [[Link](#)]
- Feenstra, J. F., I. Burton, J. B. Smith, and R. S. J. Tol, 1998: Handbook on Methods for Climate Change Impact Assessment and Adaptation Strategies, Version 2.0, UNEP, 464 pp.
- Gibelin, A. L. and M. Déqué, 2003: Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Clim. Dyn.*, **20**, 327-339, doi: 10.1007/s00382-002-0277-1. [[Link](#)]
- IPCC, 2001: Climate Change 2001: Impacts, Adaptation, and Vulnerability, Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1032 pp.
- Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.*, **40**, 1194-1199, doi: 10.1002/grl.50256. [[Link](#)]
- Köppen, W., 1936: Das Geographische System der Klimate, Gebrüder Borntraeger, Berlin, 44 pp.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.*, **15**, 259-263, doi: 10.1127/0941-2948/2006/0130. [[Link](#)]
- Kullback, S., 1968: Information Theory and Statistics, Dover Publications, 432 pp.
- Kutzbach, J. E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *J. Appl. Meteorol. Climatol.*, **6**, 791-802, doi: 10.1175/1520-0450(1967)006<0791:EEOSLP>2.0.CO;2. [[Link](#)]
- Lien, W. Y., C. P. Tung, Y. H. Ho, C. H. Tai, and L. H. Chuang, 2013: Establish methods to evaluate the projection ability of general circulation models. *J. Taiwan Agric. Eng.*, **59**, 1-14. (in Chinese)
- Liu, T. M., 2010: Study on integrated climate change impact assessment system for regional water resources. Ph.D. Thesis, Department of Bioenvironmental Systems Engineering, College of Bioresources and Agriculture, National Taiwan University, Taipei, Taiwan, 134 pp. (in Chinese)
- Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Scientific Report No. 1, Statistical Forecasting Project, Department of Meteorology, Massachusetts Institute of Technology, 98 pp.
- Masson, D. and R. Knutti, 2011: Climate model genealogy. *Geophys. Res. Lett.*, **38**, 1-4, doi: 10.1029/2011GL046864. [[Link](#)]
- Pearson, K., 1901: LIII. On lines and planes of closest fit to systems of points in space. *Phil. Mag.*, **2**, 559-572, doi: 10.1080/14786440109462720. [[Link](#)]
- Peel, M. C., B. L. Finlayson, and T. A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, **4**, 439-473, doi: 10.5194/hessd-4-439-2007. [[Link](#)]
- Sadiq, N., 2011: Principal component and clustering analyses for seasonal classification of Karachi. *J. Geogr. Geol.*, **7**, 9-16.
- Schmidli, J., C. Frei, and P. L. Vidale, 2006: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.*, **26**, 679-689, doi: 10.1002/joc.1287. [[Link](#)]
- Suppiah, R., K. J. Hennessy, and P. H. Whetton, 2007: A comparison of Australian climate change projections based on IPCC TAR and AR4 climate model simulations. MODSIM07 International Congress on Modelling and Simulation: Land, Water & Environmental Management, Integrated Systems for Sustainability, Modelling and Simulation Society of Australia and New Zealand, Christchurch, 518-524.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.*, **93**, 485-498, doi: 10.1175/BAMS-D-11-00094.1. [[Link](#)]
- Trewartha, G. T., 1954: An Introduction to Climate, McGraw-Hill, New York, 402 pp.
- Tryon, R. C., 1939: Cluster Analysis: Correlation Profile and Orthometric (Factor) Analysis for the Isolation of Unities in Mind and Personality, Edwards's Brother, Incorporated, Lithotrippers and Publishers, 122 pp.
- Wan, P. C., 1973: A study of the regional climate and weather in Taiwan. *Meteorol. Bull.*, **19**, 1-19. (in Chinese)
- Ward, J. H., 1963: Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.*, **58**, 236-244, doi: 10.2307/2282967. [[Link](#)]
- Whetton, P. H., K. L. McInnes, R. N. Jones, K. J. Hennessy, R. Suppiah, C. M. Page, J. M. Bathols, and P. J. Durack, 2005: Australian climate change projections for impact assessment and policy application: A review. CSIRO Marine and Atmospheric Research Paper 001, CSIRO Marine and Atmospheric Research, Aspendale, Vic, 34 pp, doi: 10.4225/08/58615c38e8447. [[Link](#)]
- Wilby, R. L. and T. M. L. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations. *Progr. Phys. Geogr.*, **21**, 530-548, doi: 10.1177/030913339702100403. [[Link](#)]
- Wu, M. C. and Y. L. Chen, 1993: The classification of the climate in Taiwan. *Atmos. Sci.*, **21**, 55-66. (in Chinese)