

CMIP5 Decadal Precipitation over an Australian Catchment

Md Monowar Hossain ^{1,*} , A. H. M. Faisal Anwar ^{2,*} , Nikhil Garg ³, Mahesh Prakash ³
and Mohammed Abdul Bari ⁴ 

¹ Department of Civil Engineering, Dhaka University of Engineering & Technology, Gazipur 1700, Bangladesh

² School of Civil and Mechanical Engineering, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

³ Commonwealth Scientific and Industrial Research Organisation (CSIRO), Data61, Clayton, VIC 3168, Australia; nikhil.garg@data61.csiro.au (N.G.); mahesh.prakash@data61.csiro.au (M.P.)

⁴ Bureau of Meteorology, West Perth, WA 6872, Australia; mohammed.bari@bom.gov.au

* Correspondence: monowar84@duet.ac.bd (M.M.H.); f.anwar@curtin.edu.au (A.H.M.F.A.)

Abstract: The fidelity of the decadal experiment in Coupled Model Intercomparison Project Phase-5 (CMIP5) has been examined, over different climate variables for multiple temporal and spatial scales, in many previous studies. However, most of the studies were for the temperature and temperature-based climate indices. A quite limited study was conducted on precipitation of decadal experiment, and no attention was paid to the catchment level. This study evaluates the performances of eight GCMs (MIROC4h, EC-EARTH, MRI-CGCM3, MPI-ESM-MR, MPI-ESM-LR, MIROC5, CMCC-CM, and CanCM4) for the monthly hindcast precipitation of decadal experiment over the Brisbane River catchment in Queensland, Australia. First, the GCMs datasets were spatially interpolated onto a spatial resolution of $0.05 \times 0.05^\circ$ (5×5 km) matching with the grids of observed data and then were cut for the catchment. Next, model outputs were evaluated for temporal skills, dry and wet periods, and total precipitation (over time and space) based on the observed values. Skill test results revealed that model performances varied over the initialization years and showed comparatively higher scores from the initialization year 1990 and onward. Models with finer spatial resolutions showed comparatively better performances as opposed to the models of coarse spatial resolutions, where MIROC4h outperformed followed by EC-EARTH and MRI-CGCM3. Based on the performances, models were grouped into three categories, where models (MIROC4h, EC-EARTH, and MRI-CGCM3) with high performances fell in the first category, and middle (MPI-ESM-LR and MPI-ESM-MR) and comparatively low-performing models (MIROC5, CanCM4, and CMCC-CM) fell in the second and third categories, respectively. To compare the performances of multi-model ensembles' mean (MMEMs), three MMEMs were formed. The arithmetic mean of the first category formed MMEM1, the second and third categories formed MMEM2, and all eight models formed MMEM3. The performances of MMEMs were also assessed using the same skill tests, and MMEM2 performed best, which suggests that evaluation of models' performances is highly important before the formation of MMEM.

Keywords: CMIP5; decadal; precipitation; prediction; catchment; multi-model



Citation: Hossain, M.M.; Anwar, A.H.M.F.; Garg, N.; Prakash, M.; Bari, M.A. CMIP5 Decadal Precipitation over an Australian Catchment.

Hydrology **2024**, *11*, 24. <https://doi.org/10.3390/hydrology11020024>

Academic Editor: Ezio Todini

Received: 28 November 2023

Revised: 27 January 2024

Accepted: 30 January 2024

Published: 7 February 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The evaluation of General Circulation Models (GCMs) has become a very important task to measure the models' uncertainty in future prediction of climate variables. Comparison of models predicted historical data with their corresponding observed values determines how well the GCMs represent historical climate and thus forms an integral part of the confidence-building exercise for climate predictions. It is assumed that the better performance of models over the historical period leads to developing more confidence in their future predictions. As the GCMs are used to explore the future climate variabilities and potential impacts on the Earth, assessment of GCMs has been a growing need in the climate research community. However, depending on the requirements, available resources,

geographical locations, and variables considered to assess the model performances, assessment strategies become different. Since the change of climate and its potential impact on this planet varies from region to region, it is important to evaluate the models based on different regions and spatial scales though the evaluation of climate models and their ensembles is crucial in climate studies [1]. Research on regional or local climate variability and their potential impacts are high in demand for transferring research-based scientific knowledge to increase the resilience of society to climate change. This will help in planning the future development of the infrastructures of a region [2].

Coupled Model Intercomparison Project Phase-5 (CMIP5) provides an unprecedented collection of global climate data of different time scales including decadal experiments that were produced by a wide range of GCMs [3]. Evaluation of CMIP5 decadal prediction has been run so far based on different evaluation aspects such as different regions, different climate variables, and their different spatial and temporal resolutions. For instance, Choi et al. [4] investigated the prediction skill of CMIP5 decadal hindcast (“hindcast” is a methodological approach used to recreate and analyze past weather or ocean conditions, and model-generated historical data are termed as hindcast data) near-surface air temperature for the global scale, while other researchers have investigated other climate variables on continental or regional scales [5–7]. Lovino et al. [5] evaluated decadal hindcast precipitation and temperature over northern Argentina and reported higher skills of models to reproduce the temperature as opposed to precipitation, where precipitation skills were found remarkably lower. Mckeller et al. [6] investigated the decadal hindcast maximum and minimum temperature over the state of California and reported the best-performing model. Likewise, Gaetani and Mohino [7] evaluated model performances to reproduce Sahelian precipitation and reported better-performing models. However, these studies were for different geographical locations with coarser spatial resolutions for considered variables. For instance, the spatial resolution of models used by Kumar et al. [2] and Choi et al. [4] was 2.5° , Gaetani and Mohino [7] used models of more than 1.1° , and Lovino et al. [5] used precipitation data of 1.0° spatial resolution. At a regional level, Mehrotra et al. [8] assessed the multi-model decadal hindcast of precipitation for different hydrological regions over Australia using 0.5° spatial resolution and reported lower skills for precipitation as opposed to temperature and geopotential height. Climate data of 0.5° spatial resolution cover a ground area equivalent to a square of 50 km length over the Australian region. Comparatively, a 50×50 km area is very big where climate variabilities are also large and frequency and magnitude of precipitation vary in a few kilometers (such as in Australia). As the precipitation shows more spatial and temporal variability than temperature and the model performances vary from region to region, the model performances at the local level for finer spatial resolution are essential for precipitation.

Numerous studies have evaluated CMIP5 models over Australia [1,8–11], but studies on evaluating CMIP5 decadal precipitation at catchment scale can hardly be found. After Mehrotra et al. [8], who assessed the CMIP5 decadal hindcast precipitation over different hydrological regions ($0.5^\circ \times 0.5^\circ$) in Australia, Hossain et al. [12,13] recently used the CMIP5 decadal precipitation at a further finer resolution of $0.05^\circ \times 0.05^\circ$ (5×5 km) for the Brisbane River catchment in Australia for the first time. Hossain et al. [12,13] compared the model performances for investigating the model drift and their subsequent correction using alternative drift correction methods for both the monthly and seasonal mean precipitation. However, they compared the model performances at a single grid point within the Brisbane River catchment. On the contrary, Mehrotra et al. [8] used only a multi-model approach but did not consider individual models finer than 0.5° spatial resolution. Local climate variables of finer temporal and spatial resolution, especially for precipitation, are very important for water managers for planning and developing infrastructures, as well as decision-making for local businesses and agriculture. To maintain sustainable development with effective future planning based on the models’ projected precipitation, it is important to evaluate the performance of the CMIP5 models’ hindcast precipitation.

Many researchers have suggested using MMEM [14–17] while using GCM data to reduce the model biases. The use of MMEM may enhance the model performances [2,18] by reducing the biases to some extent, but there is no information available on the ranking of GCM models and, based on this, which and how many models should be considered to produce MMEM so that it could provide better outcome. This is essential for CMIP5 decadal precipitation because of its wide range in spatial and temporal variability in providing the model output 10 years ahead. That is why the objective of this paper is, first, to categorize the models based on their performances at the catchment level with a spatial resolution of 0.05° and, next, to identify the best combination of different models that would provide better performance. This would help the water managers and policymakers to sort out models depending on their specific needs while assessing the future water availability based on the GCMs-derived precipitation on a decadal scale through CMIP5.

2. Study Area

In this study, the Brisbane River catchment (Figure 1) in Queensland was selected as the study area, which is located in the eastern states of Australia between the latitudes 26.50°S – 28.150°S and the longitudes 151.70°E – 153.150°E . It has an area of 13,549 square kilometers and a sub-tropical climate where most of the precipitation occurs during summer (December–January–February) and minimum precipitation in winter (June–July–August). From the monthly observed gridded precipitation (1911–2015) over the Brisbane River catchment, it was found that the monthly precipitation varied from nil to 1360 mm with an annual average precipitation of 628 mm, and the number of upper and lower extremes were not quite small. The Brisbane catchment was selected because of its tropical climate nature with low to moderate yearly precipitation variability.

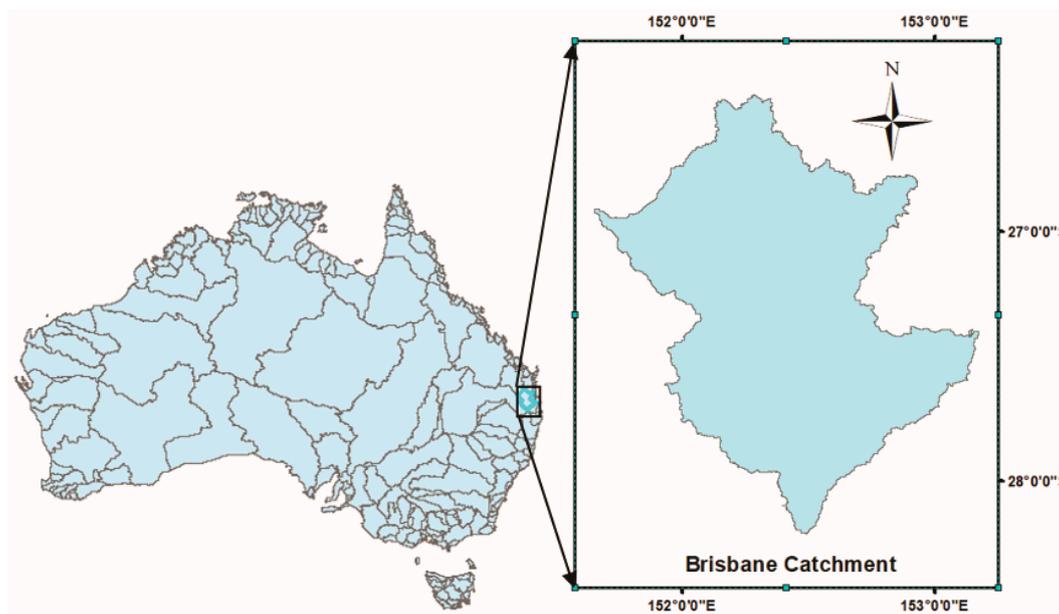


Figure 1. Brisbane River catchment.

3. Data Collection and Processing

3.1. Data Collection

The CMIP5 decadal experiment provides 10- and 30-year-long ensemble predictions from multiple modeling groups (henceforth mentioned as CMIP5 decadal hindcasts [19]). For monthly decadal hindcasts precipitation from eight GCMs (out of ten)—MIROC4h, EC-EARTH, MRI-CGCM3, MPI-ESM-MR, MPI-ESM-LR, MIROC5, CMCC-CM, and CanCM4—decadal hindcast precipitation were downloaded from the CMIP5 data portal (<https://esgf-node.llnl.gov/projects/cmip5/> accessed on 20 June 2018). The other two models, HadCM3 (spatial resolution $3.75 \times 2.5^\circ$) and IPSL-CM5A-LR (spatial resolution $3.75 \times 1.89^\circ$), were

not considered in this study because of their relatively coarser spatial resolution and different calendar system (HadCM3). For the initialized period 1960–2005, data simulated over 10 years, which were initialized every 5 years during this period, were selected for this study as they were found comparatively better than the 30-year simulation [20]. The details of the selected models are given in Table 1.

Table 1. Selected models with the initialization year 1960–2005.

Model Name (Modelling Center or Group) Resolutions: ° lon × ° lat	Initialization Year (1960–2005)										
	60	65	70	75	80	85	90	95	00	05	
EC-EARTH (EC-EARTH Consortium) 1.125 × 1.1215	14	14	14	14	14	14	14	14	14	10	18
MRI-CGCM3 (Meteorological Research Institute) 1.125 × 1.1215	06	08	09	09	06	09	09	09	09	09	06
MPI-ESM-LR (Max Planck Institute for Meteorology) 1.875 × 1.865	10	10	10	10	10	10	10	10	10	10	10
MPI-ESM-MR (Max Planck Institute for Meteorology) 1.875 × 1.865	03	03	03	03	03	03	03	03	03	03	03
MIROC4h (AORI-Tokyo University, NIES and JAMST) 0.5625 × 0.5616	03	03	03	06	06	06	06	06	06	06	06
MIROC5 (AORI-Tokyo University, NIES and JAMST) 1.4062 × 1.4007	06	06	06	06	04	06	06	06	06	06	06
CanCM4 (Canadian Centre for Climate Modelling and Analysis) 2.8125 × 2.7905	20	20	20	20	20	20	20	20	20	20	20
CMCC-CM (Centro Euro-Mediterraneo per I Cambiamenti Climatici) 0.75 × 0.748	03	03	03	03	03	03	03	03	03	03	03

The observed gridded monthly precipitation of 0.050×0.050 ($\approx 5 \times 5$ km) was collected from the Australian Bureau of Meteorology (Observed/Bureau). This data were produced using the Australian Water Resources Assessment Landscape model (AWRA-L V5) [21].

3.2. Data Processing

The GCMs' resolutions (100–250 km) were found inadequate for regional studies due to a lack of information at the catchment levels [22–24]. The regional climate model (RCM) was useful to transfer the coarse spatial GCMs' data to the local scale, but it needs a wide range of climate variables as well as rigorous efforts to develop. For this reason, GCMs data were spatially interpolated onto a $0.05 \times 0.05^\circ$ spatial resolution using the second-order conservative (SOC) method matching with the grids of observed data. For the gridded precipitation data, the SOC method was found comparatively better than other commonly used spatial interpolation methods [25]. Skelly and Henderson-Sellers [26] suggested GCM derive gridded precipitation to consider as areal quantities, and spatial interpolation will not create any new information except the spatial precision of the data. Skelly and Henderson-Sellers [26] also suggested that researchers could subdivide the grid box in

almost any manner until the original volume remains the same. On the contrary, Jones [27] suggested that precipitation flux must be remapped in a conservative manner to maintain the water budget of the coupled climate system. While sub-gridding the GCM data using the SOC method, it conserves precipitation flux from their native grids to subsequent grids [27]. For this reason, this study used the SOC method for spatial interpolation as it was followed in other research [13].

4. Methodology

A simple and direct approach for the model evaluation is to compare the model output with the observations and analyze the differences. In this study, models were evaluated for temporal skills, dry and wet periods, and total precipitation based on the observed values. The meaning of the temporal skills and their descriptions are given below. Here, CC, ACC, and IA are used to measure the temporal skills, FSS are used to measure the skills over dry and wet periods, and field-sum and total-sum are used to measure the skills for total precipitation. There are 496 grids in the Brisbane River catchment with a spatial resolution of 5.0×5.0 km.

4.1. Correlation Coefficient (CC)

CC measures the linear association and presents the scale of temporal agreement between predicted and observed values. Statistically, it measures how much closer the scatter plot points to a straight line. The CC ranges from -1 to 1 for no to perfect correlation, respectively.

$$CC = \frac{\sum (F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum (F_i - \bar{F})^2} \sqrt{\sum (O_i - \bar{O})^2}} \quad (1)$$

Here, F and \bar{F} , represent models' predicted and their mean value, whereas O and \bar{O} represent observed precipitation and their mean, respectively. In the following skill tests, these notations will remain the same. Note that the mean was calculated for every individual year.

4.2. Anomaly Correlation Coefficient (ACC)

ACC was suggested by Wilks [28] to measure the temporal correlation between anomalies of the observed and predicted values. For the verification of numerical weather models' prediction, ACC is frequently used. Its value ranges from -1 to 1 for no to perfect anomaly matching.

$$ACC = \frac{\sum \{(F_i - C) - (\bar{F} - C)\} * \{(O_i - C) - (\bar{O} - C)\}}{\sqrt{\sum (F_i - C)^2} \sqrt{\sum (O_i - C)^2}} \quad (2)$$

Here, C represents the mean of the entire timespan (10 years) of the observed (Bureau) data. The higher value of ACC will indicate the higher performance in reproducing the monthly anomalies.

4.3. Index of Agreement (IA)

Wilmot [29] suggested IA to measure the accuracy of predictions. The index of agreement can be calculated as follows.

$$IA = 1 - \frac{\sum_{i=1}^n (F_i - O_i)^2}{\sum_{i=1}^n (|F_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (3)$$

The index is bounded between 0 and 1 ($0 \leq IA \leq 1$). The value closer to 1 indicates the most efficient predicting of the models.

4.4. Fractional Skill Score (FSS)

FSS is a grid-box event that directly compares the fractional coverage of models' predicted and observed values for the entire catchment. It measures how the spatial variability of models' predicted values corresponds to the spatial variability of the observed values. FSS can be obtained as:

$$FSS = 1 - \frac{\frac{1}{N} \sum_N (P_{f,m} - P_{f,o})^2}{\frac{1}{N} [\sum_N P_{f,m}^2 + \sum_N P_{f,o}^2]} \quad (4)$$

where P_f and N refer to calculated fraction and number of years, respectively. The subscript m and o present the modeled and observed fractions, respectively. In this study, fractions were calculated according to Roberts and Lean [30], but the entire catchment was considered as a whole unit, and the temporal averages (for considered months) were taken instead of the spatial averages. For doing this, threshold values; ≥ 85 percentile for the months of wet seasons (December to February—DJF) and < 15 percentile for the months of the dry seasons (June to August—JJA) were considered. To obtain the fractions (say, for January), the number of grid points covered for a specified threshold value was counted and then divided by the total number of grids within the catchment. The differences between the predicted and observed fractions (the numerators of Equation (4)) were calculated for individual months. The FSS is a temporal average score for the catchment for each considered month. It ranges from 0 to 1 for no to perfect match respectively.

4.5. Field-Sum and Total-Sum

The models' ability to reproduce the total precipitation over the entire catchment was considered as the spatial skills of the models. Field-sum is the sum of precipitation over the entire catchment for individual time steps, and the total-sum is the field-sum over the total timespan. Field-sum and total-sum of the models' precipitation were compared with the corresponding observed values.

5. Results and Analysis

5.1. Temporal Skills

The temporal skills are computed at every individual grid (total 496 grids) of the catchment for all initialization years of each model. Spatial variations of models' temporal skills across the catchment for the initialization year 1990 (1991–2000) are presented in Figure 2. The models were evaluated from the spatial perspective by counting the number of grids covered by different models for different threshold values of CC, ACC, and IA, as shown in Figure 3. The higher number of grids represents the higher spatial skill of models across the catchment. From the comparison of temporal skills, it is evident that model performance varied over the initialization years and across the catchment. From the initialization year 1990 and onward, all models showed a comparatively higher number of grids for the same thresholds of CC, ACC, and IA, and the lowest skill was observed in 1980. With the increase of threshold values, the number of grids declined for all models in all three temporal skills except CMCC-CM and MIROC5 in ACC. Compared to other selected models, MIROC4h, EC-EARTH, and MRI-CGCM3 showed a higher number of grids for all thresholds in which MIROC4h was much ahead of EC-EARTH and MRI-CGCM3. This means that the temporal agreement, the resemblance of anomalies, and the prediction accuracy of MIROC4h and EC-EARTH were spatially higher than the other models. This study also checked the number of grids for the threshold ≥ 0.6 for CC and ACC but no model could reproduce CC and ACC ≥ 0.6 at any grid. However, MIROC4h, EC-EARTH, and MRI-CGCM3 showed a significant number of grids for the IA threshold ≥ 0.6 , where MIROC4h outperformed EC-EARTH and MRI-CGCM3 (Figure 3). Comparing the models, MIROC4h showed higher temporal skills from the spatial perspective, followed by EC-EARTH and MRI-CGCM3, while MPI-ESM-MR, MIROC5, and CMCC-CM showed the low

to lowest temporal skills. Over the catchment MIROC5, MPI-ESM-MR, CanCM4 showed comparatively better scores than CMCC-CM.

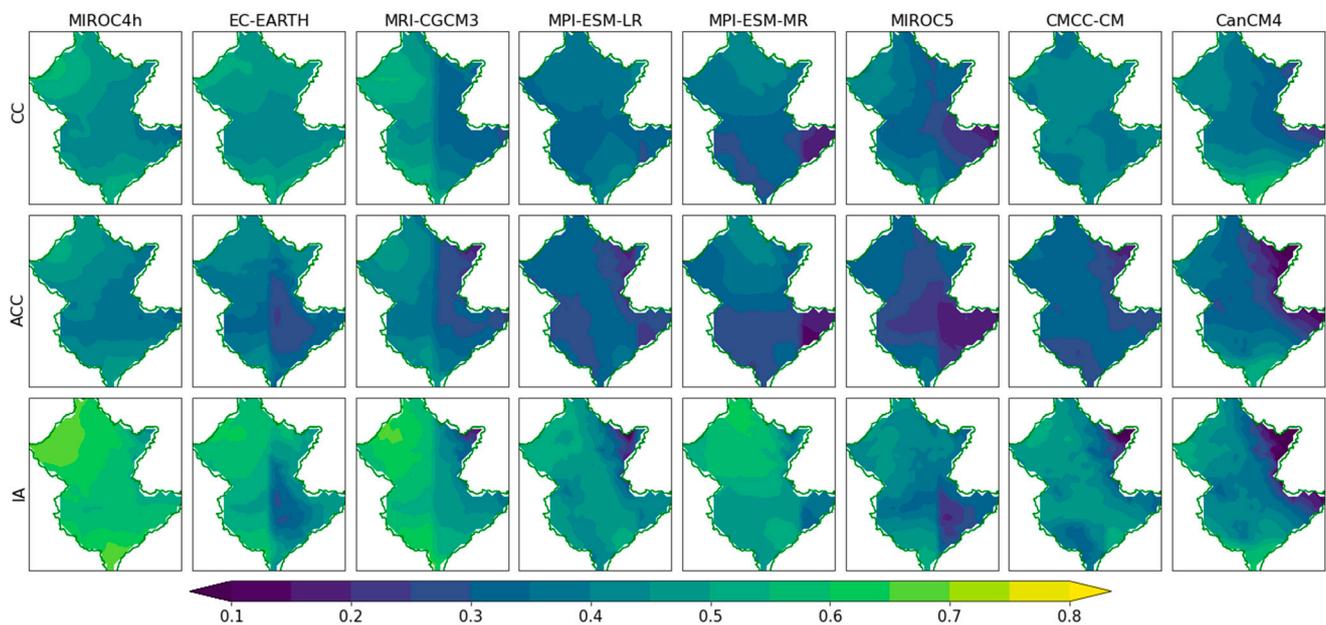


Figure 2. Spatial variations of temporal skills (CC, ACC, and IA) of the models initialized in 1990 (period; 1991–2000) over the Brisbane River catchment.

5.2. Evaluation for Dry and Wet Periods

Skills to reproduce the dry and wet events were assessed at the selected grid and over the entire catchment. For the selected grid, all months were considered against four different thresholds (25th, 50th, 75th, and 90th percentiles correspond to 25, 60, 110, and 175 mm, respectively), whereas for the entire catchment, FSS were used for the months of dry (JJA) and wet (DJF) periods only.

5.2.1. At the Selected Grid

A comparison to reproduce the dry and wet events based on the selected precipitation thresholds at the selected grid is presented in Figure 4. This comparison was based on the ratio of the number of months of respective precipitation thresholds (mentioned on the top of the individual plot in Figure 4) in the model data to the observed data. It was observed that EC-EARTH and MIROC5 could reproduce no dry events ($Pr \leq 25$ mm), while CMCC-CM overestimated the number of dry events to be almost double the dry events in the observed data. Meanwhile, MIROC4h performed better to produce dry events, showing 50th and 75th percentile values compared with the other models. However, MIROC4h was a little behind the MPI-ESM-MR for the extreme wet events ($Pr \geq 175$ mm). This means MPI-ESM-MR can reproduce extreme wet events better than the other models. EC-EARTH, MPI-ESM-LR, and MPI-ESM-MR underestimated the events of threshold $Pr \leq 60$ mm, whereas these models overestimated the wet events ($Pr \geq 110$ mm), which is an indication of the models' tendency to reproduce a higher number of wet events than opposed to dry. However, MRI-CGCM3 performed similarly to MIROC4h in reproducing the number of events for the threshold of ≤ 60 mm but underestimated the number of events thresholds of ≥ 110 mm. To reproduce the extreme wet events ($Pr \geq 175$ mm), all models showed an underestimation, in which MPI-ESM-MR and MIROC4h showed considerably better skills. The CMCC-CM and CanCM4 showed poorest and no skills, respectively, for extreme wet events.

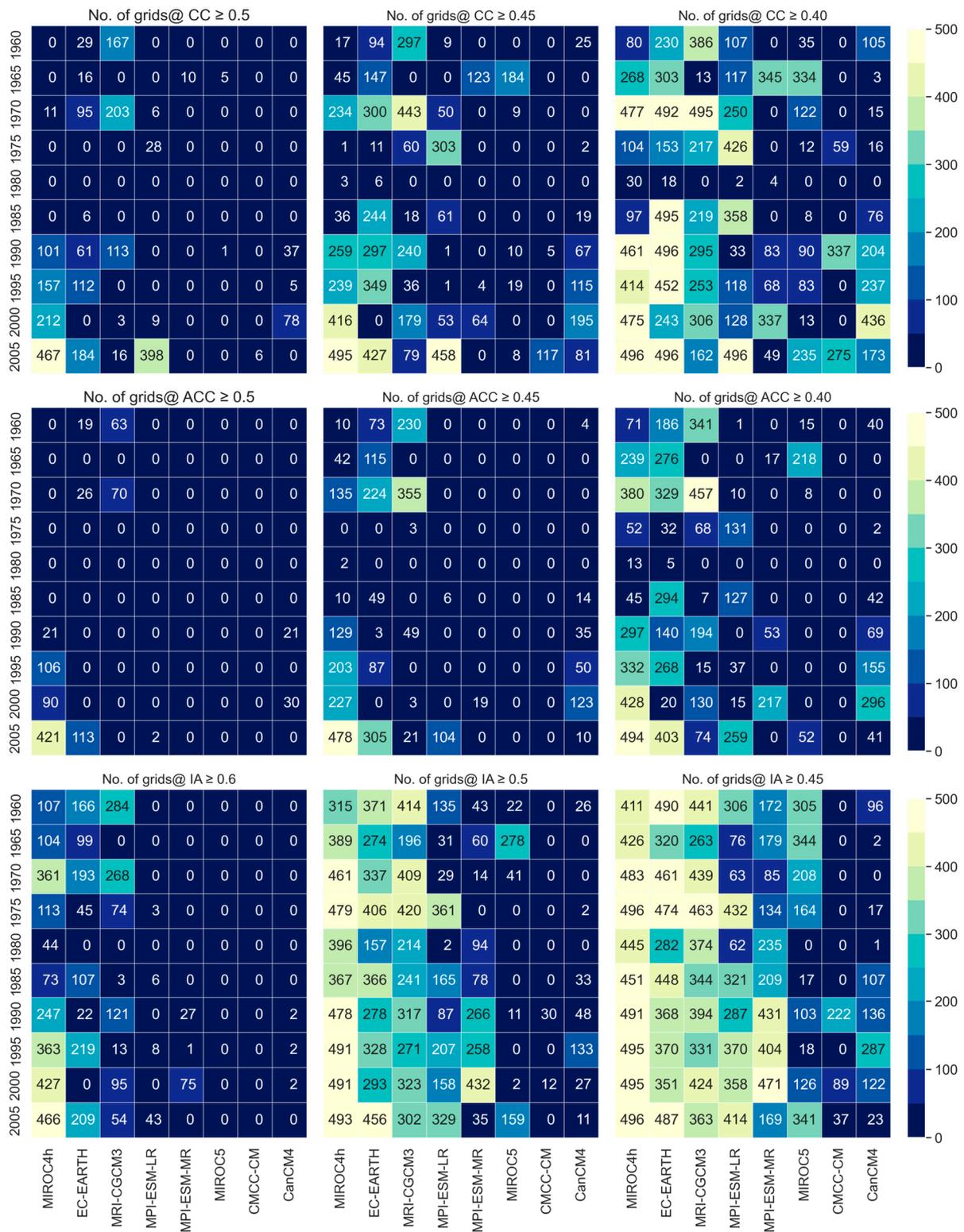


Figure 3. Number of grids covered by different models for different thresholds of CC, ACC, and IA. The vertical axis presents the initialization years, and the horizontal axis presents the model’s name. Threshold values are provided on the top of each subplot.

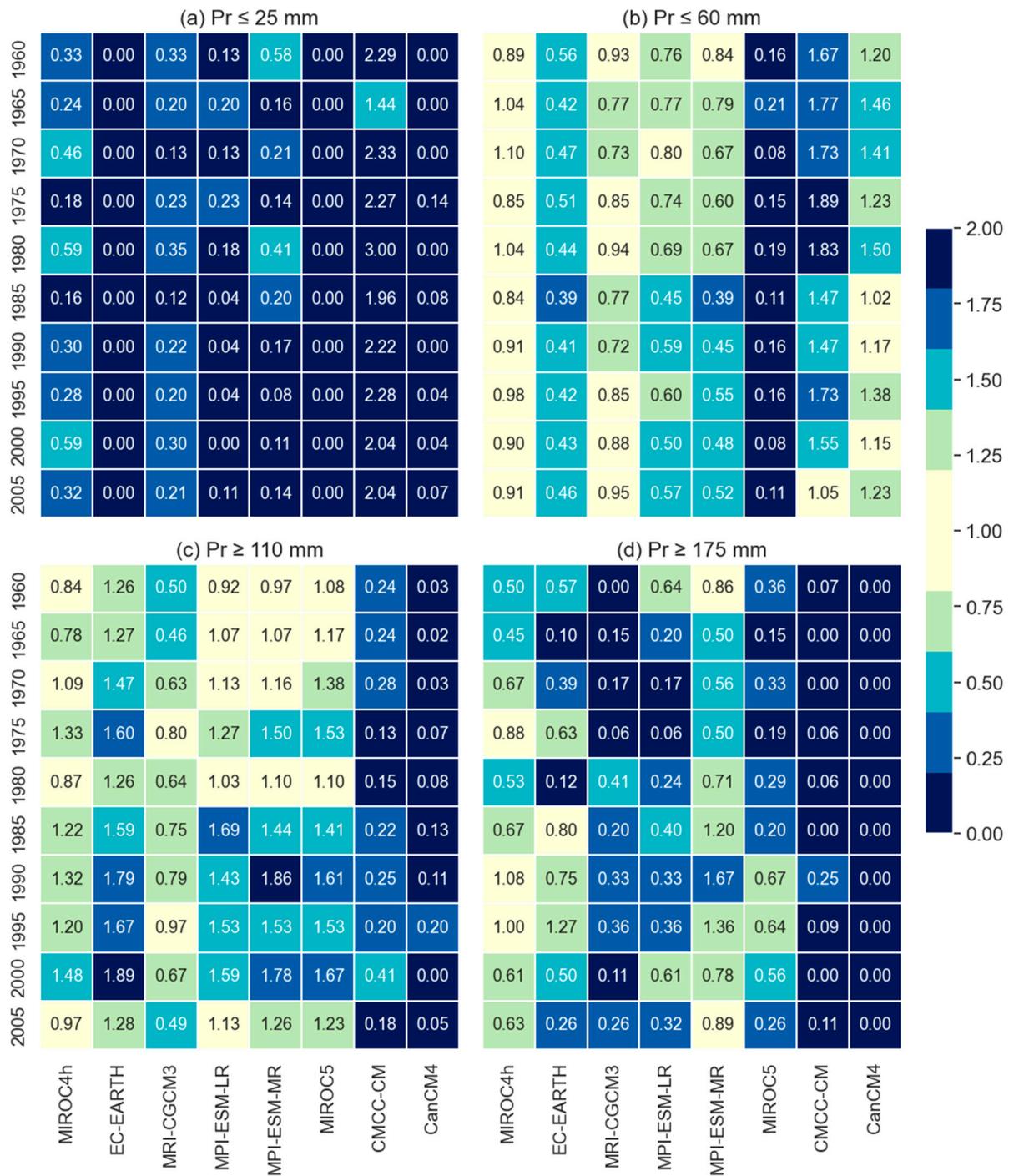


Figure 4. Comparison of model skills to reproduce dry and wet events at a selected grid point. Values 1.0 present perfect matching, while values below and above 1.0 present under- and over-prediction, respectively.

5.2.2. Over the Entire Catchment

FSSs were calculated for the months of winter (dry) and summer (wet) seasons only. FSS of all the initialization year of all models are shown in Figure 5. Results showed that for the months of summer seasons (DJF), MRI-CGCM3 showed higher skills in December and January but was a little behind EC-EARTH in February. On the contrary, CMCC-CM showed the lowest skill in December but showed similar skills with other models in January and February.

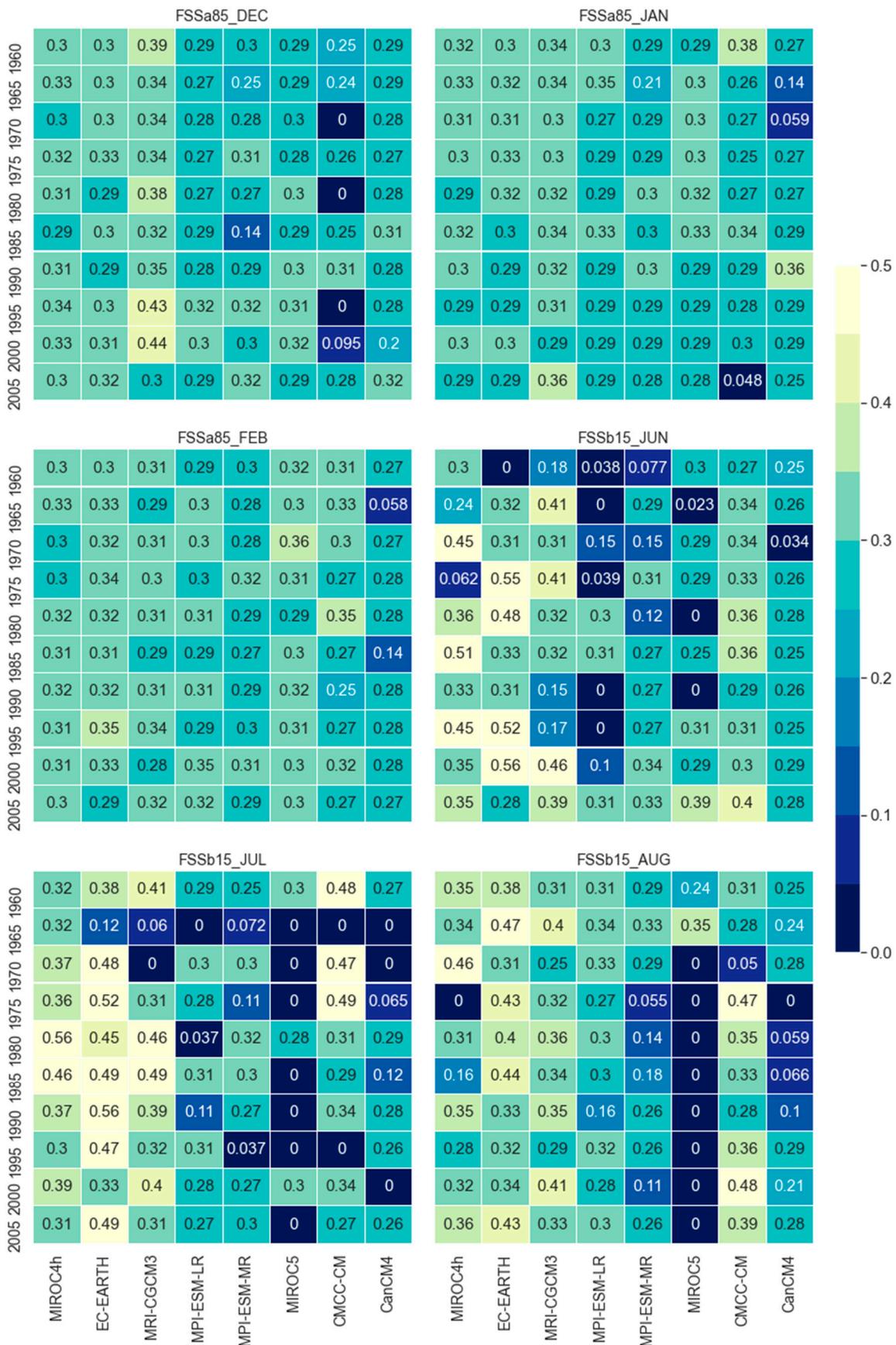


Figure 5. Fractional skill score for the months of winter and summer seasons.

However, except for the higher skill of MRI-CGCM3 and the lowest skill of CMCC-CM in December, all other models showed similar skill scores with few variations in the winter seasons. This indicates different models' skills are almost similar to reproducing wet events. In the dry season, MIROC5 showed the lowest skill, while EC-EARTH showed the higher skill, which was even higher than MIROC4h and MRI-CGCM3. The FSSb15 scores of EC-EARTH, MIROC4h, and MRI-CGCM3 were much better than the score obtained for FSSa85. This reveals that these models are better for reproducing dry events as opposed to wet events, and the reverse is true for MIROC5, MPI-ESM-MR, and CanCM4, respectively.

5.3. Evaluation for Total Precipitation

5.3.1. At the Selected Grid

To evaluate the model performances in reproducing the total precipitation, models' cumulative (over time) precipitation at several randomly selected grids (evenly distributed across the catchment) within the catchments and total precipitation over the entire catchment were compared. The cumulative sum of monthly precipitation of different models at the selected grid for different initialization years is presented in Figure 6. The models' skills showed both temporal and spatial variations in predicting accumulated precipitation, but no model could reproduce the accumulated precipitation as observed. However, only a few models (MIROC4h, MPI-ESM-LR, and MPI-ESM-MR) could reproduce the accumulated precipitation close to the observed accumulation. Nevertheless, CMCC-CM, CanCM4, and MRI-CGCM3 underestimated the accumulated precipitation, while EC-EARTH and MIROC5 overestimated the accumulated values. With the change of grid locations, model performances may change, but the relative performances among the models will remain the same.

5.3.2. Over the Entire Catchment

For comparing the model performances on total precipitation over the entire catchment, this study calculated the field-sum of the models and observed values then assessed through the temporal skills as shown in Figure 7. The total-sum of the models and observed values were also calculated and assessed through the ratio between model and observed values (Figure 7). From the comparison, it was observed that the field-sum of MIROC4h, EC-EARTH, and MRI-CGCM3 showed comparatively higher accuracy (IA), temporal agreement (CC), and the resemblance of anomalies (ACC) with the field-sum of the observed precipitation. The model performances on reproducing the total precipitation varied over the initialization years (Figure 7d).

Before and after 1985, MRI-CGCM3 and MPI-ESM-MR showed comparatively better resemblance with the observed total precipitation, followed by MIROC4h and EC-EARTH. On the contrary, CMCC-CM showed the lowest performance to reproduce total-sum precipitation throughout all initialization years. From the skill assessments, it was revealed that the MIROC4h surpassed other models in almost all performance indicators, followed by EC-EARTH and MRI-CGCM3, while MPI-ESM-LR and MPI-ESM-MR showed medium skill scores. Lower skill scores were observed for MIROC5, CanCM4, and CMCC-CM, respectively. MIROC4h was also marked as the best model to reproduce precipitation in other studies [5,31], though these studies did not use the decadal experiments' data. This may be due to the finer resolution of the atmospheric component of MIROC4h, which has enhanced its ability to capture the more realistic climate features [31,32] at the local level.

The overall skill assessment results revealed that all models showed comparatively lower skills in the initialization years from 1960 to 1985 and better skills observed from the initialization year 1990 and onward.

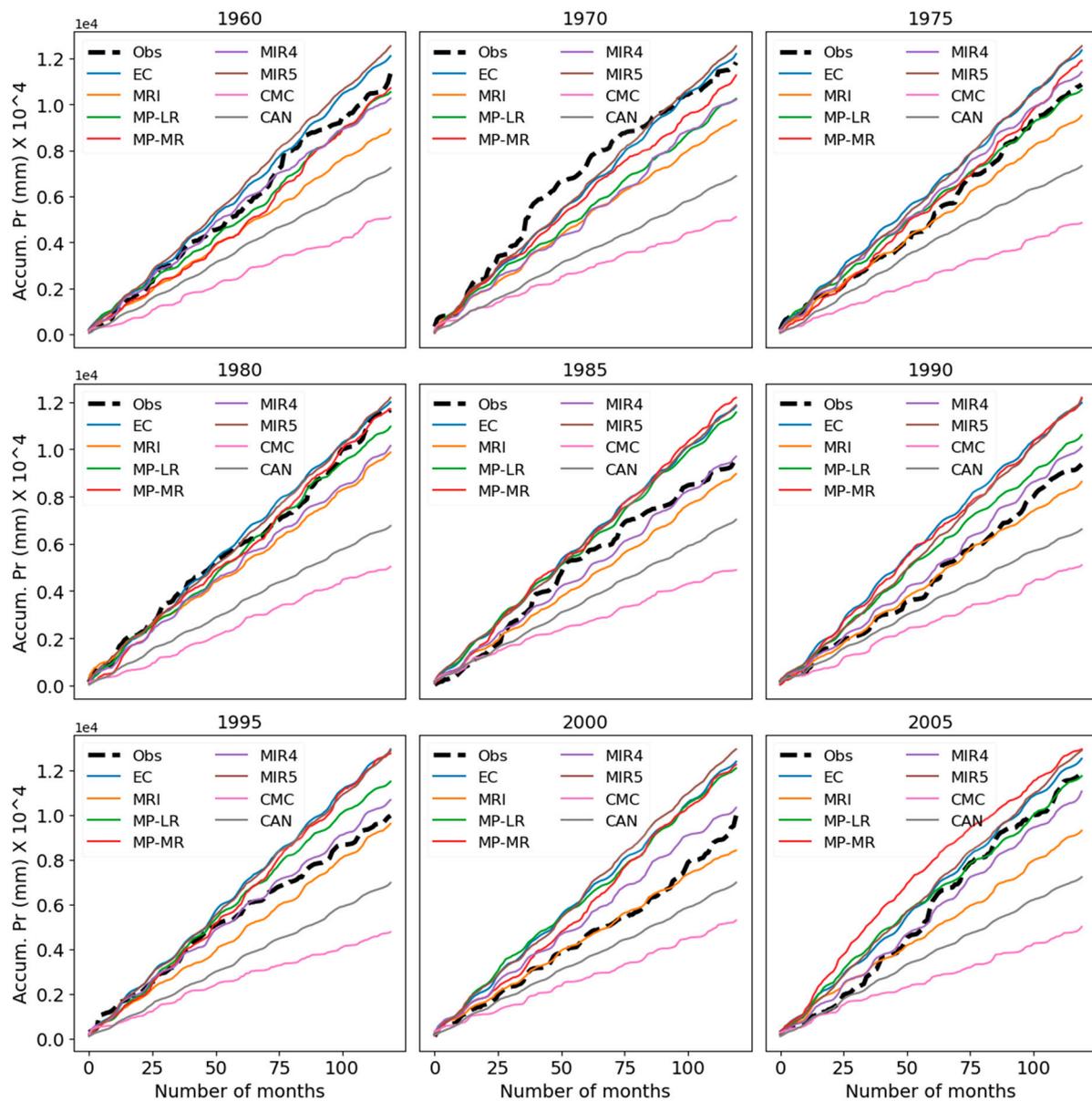


Figure 6. Cumulative sum of monthly precipitation of different models at the selected grid point in different initialization years. The vertical axis presents the accumulated precipitation, and the horizontal axis presents the number of months over the decade.

5.4. Model Categorisation and Formulation of MMEM

Based on the skill comparisons, this study divided the models into three different categories: Category-I, Category-II, and Category-III. While categorizing the models based on their skills at the selected grid and over the catchment, MIROC4h, EC-EARTH, and MRI-CGCM3 fell in the first category (Category-I), as they consistently performed in the top three, and their performance metrics were found to be very close to each other. Similarly, MPI-ESM-LR and MPI-ESM-MR were in the second (Category-II) category, as they showed medium skill scores in all skill tests over the initialization years. Lastly, MIROC5, CanCM4, and CMCC-CM fell in Category-III.

GCMs' outputs indeed contain uncertainties and biases, which will cause a lower skill score, but multi-model ensembles' mean (MMEM) may enhance the models' skills [2,13,17,18] by reducing uncertainties [13,15–17]. In this study, the skill tests were employed on the ensembles' mean of individual models' raw values (interpolated) only. Here, the arithmetic

mean of multiple models is referred to as the MMEM. The performances of MMEMs were also assessed based on the similar skill tests that were employed on individual models, and the results are summarized below. To form the MMEMs, three different combinations were considered. The arithmetic mean of Category-I models is referred to as first MMEM (MMEM1), the arithmetic mean of the Category-I and Category-II models is referred to as the second MMEM (MMEM2), and finally, the arithmetic mean of all models is referred to as the third MMEM (MMEM3).

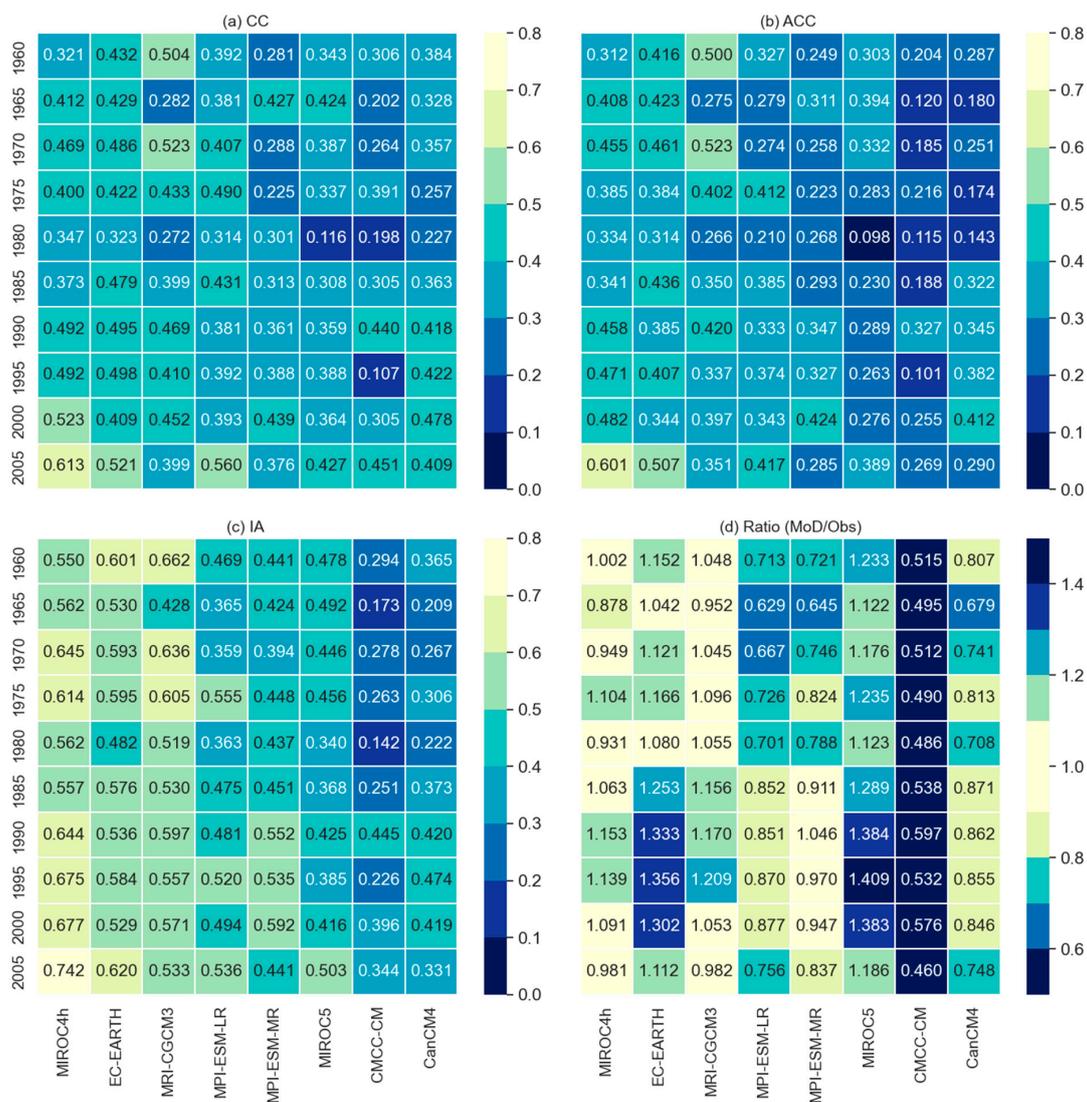


Figure 7. Performance indicators of the models to reproduce the total precipitation of the entire catchment.

5.5. Performance of MMEMs

The temporal skills at individual grids of the different thresholds, temporal skills along with the ratios of the field-sum, and skill on reproducing dry and wet events of different thresholds for MMEMs are presented in Figures 8–10, respectively. In general, MMEMs showed better performance than the individual models for comparatively lower thresholds of the performance metrics. For instance, the MIROC4h model showed the highest number of grids for CC and ACC at the threshold 0.5 (Figure 3), but no MMEMs could reproduce this number of grids at the same threshold (Figure 8). The same results were also observed for IA at the threshold 0.6 (see Figure 8i), but for the lower thresholds, MMEM2 showed better skill than MIROC4h in CC and ACC but not in IA. Among the three combinations, MMEM2 surpassed the other two combinations in reproducing CC and ACC.

Nevertheless, in the case of IA, MMEM2 was a little behind compared to MMEM1. Similar results were evident for performance indicators obtained from the field-sum of MMEM and the observed values (Figure 9), where MMEM2 showed the best performance for the CC and ACC, but both MMEM2 and MMEM1 showed similar skills for IA. However, to produce the dry and wet events, MMEMs showed lower performance as compared to the individual models. For instance, MIROC4h, MRI-CGCM3, and MPI-ESM-MR captured some dry events ($Pr \leq 25$ mm) at the selected grid point (Figure 4), but no combination could capture it (Figure 10), while for the wet events, MMEM showed very poor skills.

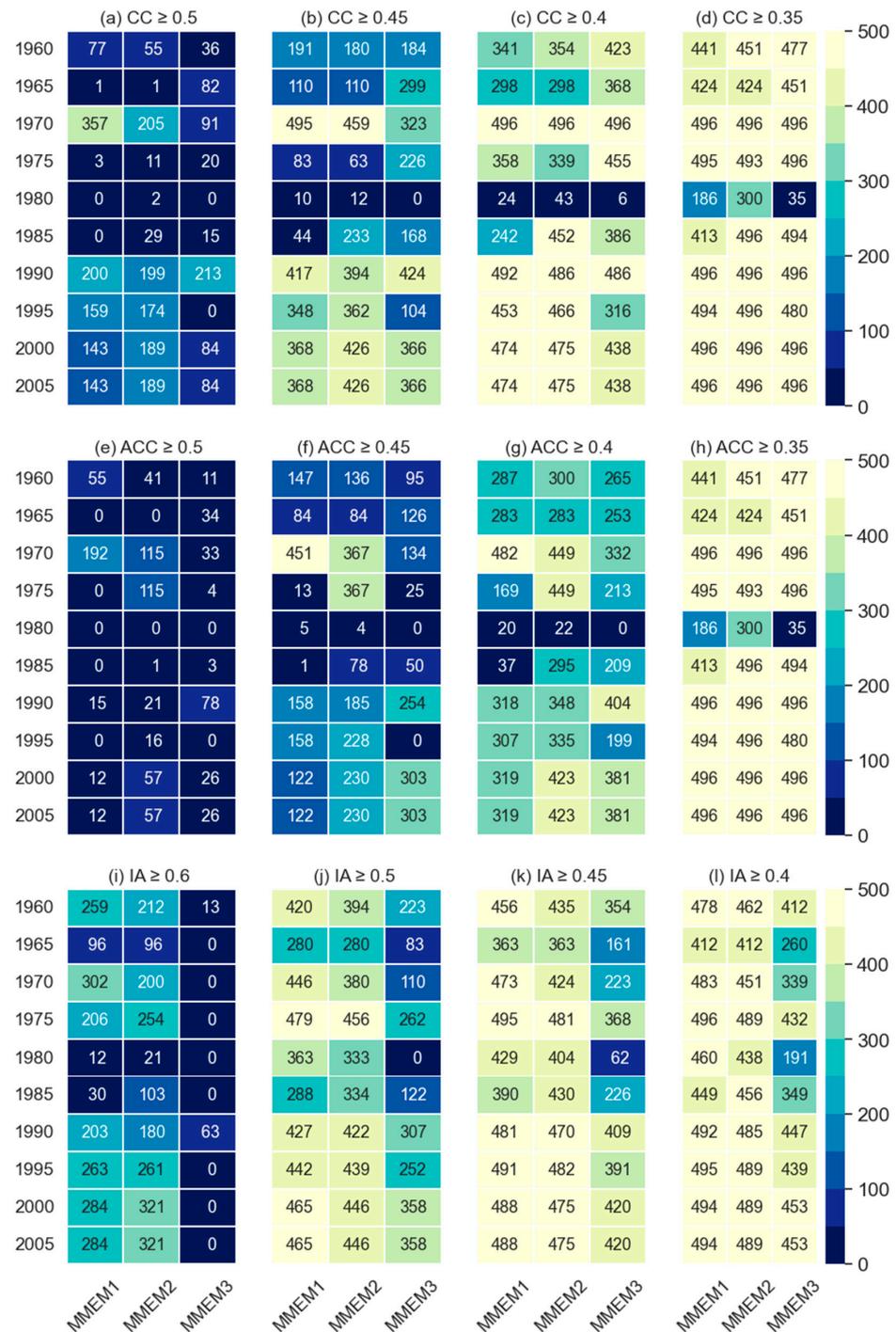


Figure 8. Number of grids covered by different combinations of models for different threshold values of performance metrics. Thresholds and the performance indicators are mentioned on the top of the individual blocks.

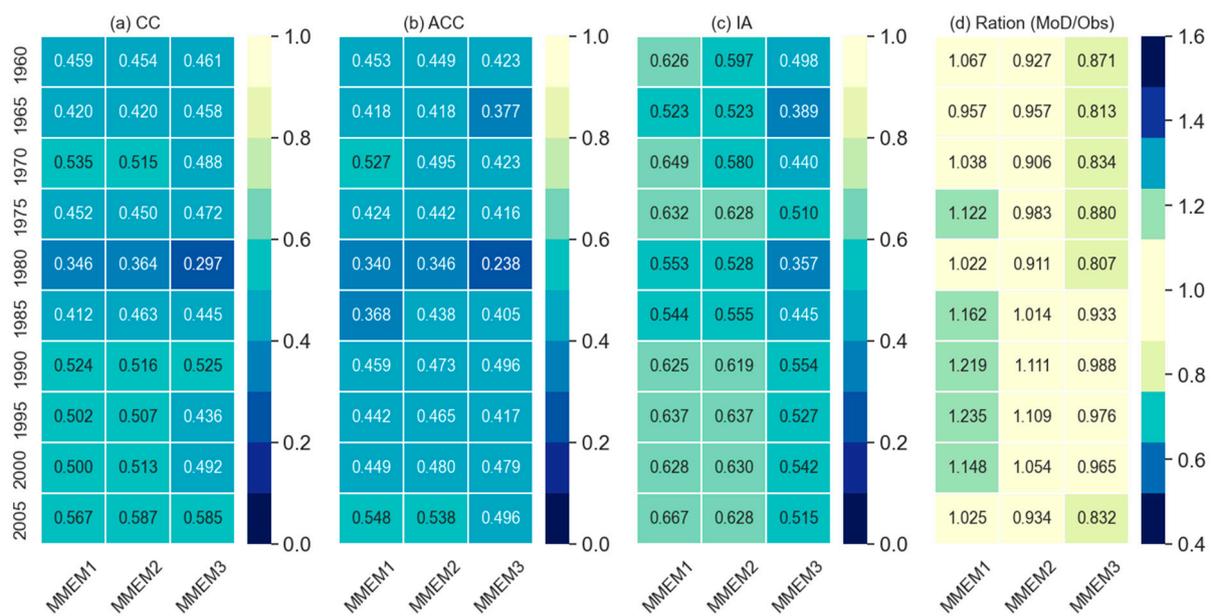


Figure 9. Performance indicators obtained from the field-sum of different MMEMs and corresponding observed values.

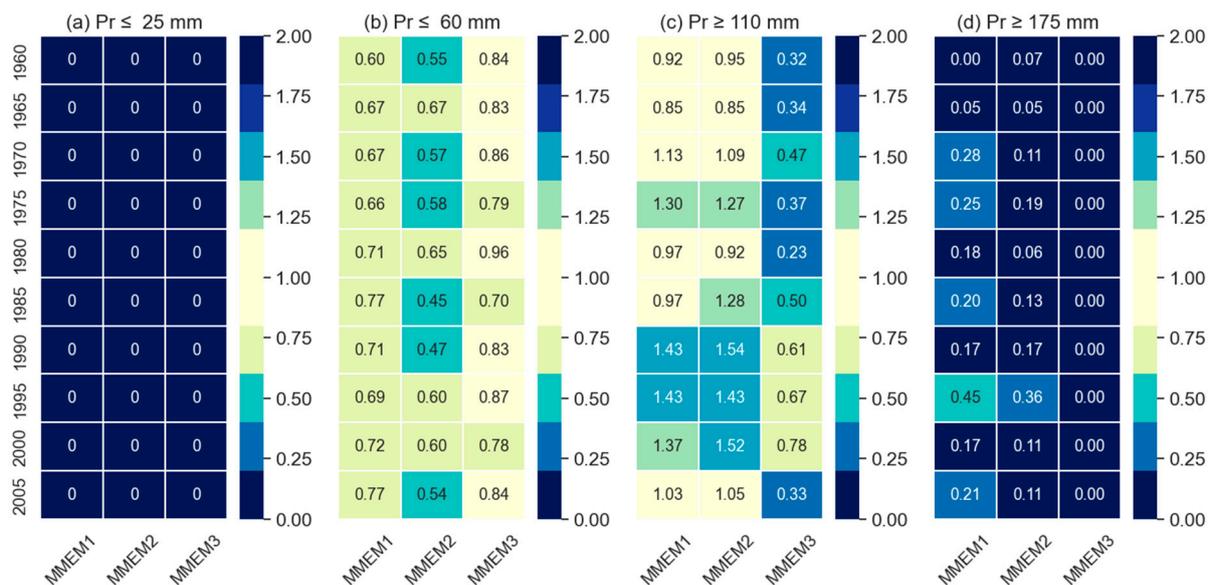


Figure 10. Skill comparison of three MMEMs to reproduce dry and wet events at the selected grid point. This comparison was based on the ratio, obtained from the number of months of the respective precipitation thresholds (mentioned on the top of the individual plot) in the model data to the number of months of observed values for different initialization years (Y-axis).

Meanwhile, MMEMs showed better performance indicators (CC, ACC, and IA) for the total precipitation of the entire catchment (field-sum), which was even better than the individual models. Nevertheless, MMEM was a little behind the MIROC4h and MRI-CGCM3 for the ratio of total-sum (sum over total time span and catchment) model combinations over the corresponding observed values (see Figure 6).

6. Discussion

This study evaluated the performance of eight selected GCMs, contributed to CMIP5 decadal precipitation prediction, over Brisbane River catchment at 0.05-degree spatial

resolution. For the evaluation, different skill metrics were employed from both temporal and spatial perspectives. The models showed a wide range of performance scores over the initialization years as well as across the catchments. This may have been due to the differences in understanding of models on local climate features, the precipitation data of finer temporal and spatial resolutions, or the combination of both.

Indeed, the model performances are dependent on the model assumptions or basic principle on understanding the Earth's climate system, its processes, and interactions among atmosphere, oceans, land, and ice-covered regions of the planet. In addition to them, decadal prediction skill also depends on the method of model initialization, and the quality and coverage of the ocean observations [3]. Different initializations also may cause models' internal variability, which is still open for further discussion. For the decadal prediction, one of the most important aspects is the model drift and its correction [8]. However, to evaluate the performance of models derived raw data, neither the drifts were investigated, and no any drift correction methods were employed. The reason is that the drift correction method itself may introduce additional errors that may not reflect the real performance of the models [12,13]. Based on the understanding of physical, chemical, and biological mechanisms of Earth's systems, different modeling groups have come up with different models with reproducing capabilities of climate variables that may vary over different regions [4,33,34] and climate variables [2,35,36]. For instance, Kumar et al. [2] analyzed the precipitation and temperature trends of the 20th century from 19 CMIP5 models and reported that the models' relative performances were better for temperature as opposed to precipitation trends. Generally, models showed lower skill to simulate precipitation than they did for temperature. This is because that the temperature was obtained from a thermodynamic balance, while precipitation results were from simplified parameterizations approximating actual processes (Flato et al. [1] and references therein). In addition, the temporal and spatial scale (considered area) of the considered variables, including the seasons of the year [18,37], may also be the reason for varying model performances. For instance, few models can reproduce winter precipitation very well but others may not, and vice versa. Likewise, Lovino et al. [5] evaluated CMIP5 model performances for decadal simulation and concluded that both were the best models. They also suggested that the MEM could reproduce large-scale features very well but fail to replicate the smaller-scale spatial variability of the observed annual precipitation pattern. These show clear evidence that there is a spatial variation in the climate model performances across the globe as they are developed by different organizations [38]. This study noticed the higher skills in the initialization year of 1990 and onward, whereas lower skills were observed in the initialization years from 1960 to 1985, but the reason behind the higher and lower skills remains unknown. However, Meehl et al. [39] reported that the consequences of Fuego (in 1974) and Pinatubo (1991) eruption degraded the decadal hindcasts skill of Pacific Sea surface temperature in the mid-1970s and in the mid-1990s, respectively. Fuego was smaller than Mount Pinatubo, and a lower degrade of skill in the mid-1970s and higher degrade of skill in the mid-1990s were evident, but no degradation on the hindcast skill was evident from Agung (erupted in 1963) and El Chichón (1982) [39]. In this study, models' higher and lower skills of initialization 1990s and 1980s seem neither relevant to volcanic eruption nor the post-eruption sequences. Nevertheless, the observed precipitation or coverage of the ocean observed state to initialize the models have been affected.

The CC and ACC values of all the selected models in all initialization years remained under the threshold ≥ 0.6 , which was marked as the threshold of significant level in previous studies [4,5], though those studies were for coarser spatial resolutions and one of them for different climate variables. Lovino et al. [5] compared CMIP5 model performances over two variables at the local level and reported higher skill scores for the temperature than precipitation of the same models where the skill scores for precipitation were remarkably lower than the scores for temperature. Similar results were also reported by Jain et al. [31]. In this sense, it seems precipitation data with higher spatial resolution may be the reason for not capturing the significant level of skills on linear association (CC) and phase differences

or anomalies (ACC). However, few models have shown that the level of significance (threshold ≥ 0.6 , for example) for the performance metric IA, which is a measure of the predicting accuracy that seems promising predictive skill of the models. However, the studies that mentioned 0.6 as the level of significance for CC and ACC used either coarser resolution data [5] or different climate variables [4]. For the local or regional level, as well as models' raw precipitation data of higher spatial and temporal resolution, 0.50 seems a significant score, which is also the same for the similar performance metrics for the case of total precipitation.

This study also investigated the model performances to reproduce the summer and winter precipitation. Upon comparing the model skills to reproduce the extreme wet (≥ 85 percentile of the observed values) and dry events (< 15 percentile of the observed values) across the catchment and at the selected grid, this study revealed that except CMCC-CM, all models showed almost similar skills to reproduce the summer precipitation but exhibited some variations to produce the winter precipitation. Similar skills were also noted for other intermediate thresholds. This was due to the maximum and minimum precipitation occurring in Brisbane River catchment during summer and winter, respectively. This means that models' responses to reproduce summer precipitation were better than the winter with the tendency to overestimate higher-precipitation events. However, the Category-I model comparatively performed better to capture the dry events (Figure 5) than the wet events, but this may vary for different regions around the globe. For instance, MRI-CGCM3 showed very good skills and has been marked as the first category model in this study but to reproduce the Sahelian precipitation, MRI-CGCM3 showed insignificant or no skills, while MPI-ESM-LR and MIROC5 were categorized as the second- and third-category models but were marked as improved skilled models for Sahelian precipitation [7].

Previous studies [5,31] have reported that MMEM improves the models' skills to reproduce climate variables, but the selection of models to form MMEM is very challenging as the arithmetic means of the models' output may further lead to loss of individual ensembles' signal [15]. This study also examined the performance of MMEM and revealed that MMEM improves the performance metrics to some extent, but not always, and the performances are highly dependent on models' combination to form MMEM. For instance, MMEM2 showed better performance metrics than the other two combinations in reproducing the extremely dry and wet events where MMEM3 showed worse performance (Figure 10). On contrary, for the highest thresholds of individual metrics, few individual models were found better than MMEM3. Similar results have also been reported in some other studies [2,6] where individual models were found better to some extent than the MMEM. However, lower skills of CMIP5 models for decadal precipitation as compared to temperature is also true for the MMEM, which was also reported by Mehrotra et al. [8].

In addition to understanding the climate system, the models' configuration, as well as structuring spatial and temporal resolutions of the simulating variables, also play a vital role in determining the model performance [32]. In this study, except for CMCC-CM, models with finer atmospheric resolutions performed better than the coarser resolutions' models (see Table 1 Category-I model). In other words, models of finer atmospheric resolutions can reproduce local climate features better than the models of coarser spatial resolutions, and similar results have also been reported in previous studies [5,31]. However, the lower skill of CMCC-CM may be due to the difference in understanding or geographical locations. However, for different climate variables like temperature, the performance of CMCC-CM may be different [5]. This study will help the water manager, infrastructure developers, and agricultural stakeholders to sort out the models before taking any decision in planning and developing infrastructures based on the models' predicted future precipitation. The findings of this study will also help the researchers for hydrological modeling, as well as other relevant stakeholders to increase the resilience of the society to climate change in relation to future water availability and uncertainty.

7. Conclusions

Evaluation of models' performance is important to check the uncertainty of their future projections. Eight models (GCMs), contributed to CMIP5 decadal prediction, were assessed here for monthly hindcast precipitation over the Brisbane River catchment, Australia. For the decadal hindcast, this was the first attempt that assessed the CMIP5 models at a catchment level with finer spatial resolution where the performance of individual models was reported based on a wide range of skill tests. Models were categorized based on their performance for temporal skills, dry and wet periods, and total precipitation (over time and space) at a selected grid and also over the entire catchment. In addition, this study assessed the performance of different MMEMs formed from the combinations of different model categories. Considering a wide range of skill tests from both the temporal and spatial perspectives, the following conclusions are drawn.

- Models with higher atmospheric resolutions showed comparatively better performances as opposed to the models of coarse spatial resolutions.
- Model performances varied over the initialization years and across the catchment. From 1990 onward, the skills of all models improved across the catchment, where MIROC4h showed the highest skills followed by EC-EARTH and MRI-CGCM3, respectively. The internal structure of high atmospheric resolutions may be the main reason for MIROC4h reproducing the local climate variables comparatively better than the other.
- To reproduce the dry events and total precipitation over the entire catchment, EC-EARTH and MRI-CGCM3, respectively, outperformed all models, while CMCC-CM showed the lowest scores in all forms of skills. For capturing the wet periods, all models showed almost similar skills with little exceptions for CMCC-CM and CanCM4 but for the dry periods, models showed a range of skill scores.
- Based on the performance skills, the GCM models were ranked into three categories in ascending order: Category-I (MIROC4h, EC-EARTH, and MRI-CGCM3), Category-II (MPI-ESM-LR and MPI-ESM-MR), and Category-III (MIROC5, CanCM4, and CMCC-CM). MMEMs were formulated as MMEM1 of Category-I models, MMEM2 combining Category-I and Category-II models, and MMEM3 as the combination of all three categories. Out of these three different MMEMs, MMEM2 was found to perform better than other MMEMs based on the overall skills, but MMEM1 performed relatively better for the case of extreme wet events. This shows the necessity of forming suitable MMEM for practical purposes of GCM data use, especially for the decadal precipitation.

The outcomes presented in this study are based on the historical data over one catchment (Brisbane River catchment) in Australia where no future projected data were considered. All the considered models in this study contributed to the future projections for decadal timescales, which are called short-term projections. From the performance obtained in this study, it may assume that models' response for the short-term future projection may not vary significantly over the timespan. Note that Australia is very big, and a variety of climates exist in different states and regions. Only one catchment like the Brisbane River catchment will not represent the entire climatology of Australia.

Recently, the Decadal Climate Prediction Project (DCPP) data was released, which contributed to the sixth Coupled Model Intercomparison Project (CMIP6). CMIP6 includes more frequent hindcast start dates considering different climate scenarios than CMIP5. Still, no study has evaluated CMIP5 decadal precipitation over an Australian catchment at a finer resolution. As this is the first attempt to evaluate CMIP5 decadal precipitation for an Australian catchment, a comparative study on similar models that contributed to both CMIP5 and CMIP6 is recommended as a follow-up study. This would provide a more robust understanding of the model performance over the same catchment.

Author Contributions: M.M.H.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing—original draft; A.H.M.F.A.: conceptualization, funding acquisition, project administration, resources, supervision, writing—review and editing; N.G.: data curation, investigation, software, validation, visualization, writing—review and editing; M.P.: data curation, funding acquisition, project administration, resources, supervision, writing—review and editing; M.A.B.: supervision, validation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly funded by CIPRS scholarship of Curtin University of Technology and Data61 PhD student scholarship of CSIRO (Commonwealth Scientific and Industrial Research Organization). These were provided to the first author for his PhD study at Curtin University of Technology, Australia.

Data Availability Statement: This study used open access climate model data available at CMIP5 data portal (<https://esgf-node.llnl.gov/projects/cmip5/>, accessed on 20 June 2018). It is mentioned in Section 3.1. The details of the selected models are presented in Table 1. The processed and analyzed data are stored in the research data repository of Curtin University, Western Australia. Data can be accessed through Curtin University after following its data sharing policy.

Acknowledgments: The authors are thankful to the working groups of the World Climate Research Program who made the CMIP5 decadal experiment data available for the researchers. The authors also would like to thank the Australian Bureau of Meteorology for providing the gridded observed rainfall data and the catchment’s shape file. The authors gratefully acknowledge the financial support from Curtin University, Perth, Australia, and CSIRO Data61, Melbourne, Australia.

Conflicts of Interest: We, all authors, declare no conflicts of interest and no financial issues relating to the submitted manuscript. We warrant that the article is the authors’ original work.

References

1. Flato, G.; Marotzke, J.; Abiodun, B.; Braconnot, P.; Chou, S.C.; Collins, W.; Cox, P.; Driouech, F.; Emori, S.; Eyring, V.; et al. Evaluation of Climate Models. In *Climate Change 2013—The Physical Science Basis*; Intergovernmental Panel on Climate Change, Ed.; Cambridge University Press: Cambridge, UK, 2013; Volume 9781107057, pp. 741–866.
2. Kumar, S.; Merwade, V.; Kinter, J.L.; Niyogi, D. Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations. *J. Clim.* **2013**, *26*, 4168–4185. [[CrossRef](#)]
3. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [[CrossRef](#)]
4. Choi, J.; Son, S.W.; Ham, Y.G.; Lee, J.Y.; Kim, H.M. Seasonal-to-interannual prediction skills of near-surface air temperature in the CMIP5 decadal hindcast experiments. *J. Clim.* **2016**, *29*, 1511–1527. [[CrossRef](#)]
5. Lovino, M.A.; Müller, O.V.; Berbery, E.H.; Müller, G.V. Evaluation of CMIP5 retrospective simulations of temperature and precipitation in northeastern Argentina. *Int. J. Climatol.* **2018**, *38*, e1158–e1175. [[CrossRef](#)]
6. McKellar, C.; Cordero, E.C.; Bridger, A.F.C.; Thrasher, B. Evaluation of the CMIP5 Decadal Hindcasts in the State of California. Master’s Thesis, San José State University, San Jose, CA, USA, 2013.
7. Gaetani, M.; Mohino, E. Decadal prediction of the sahelian precipitation in CMIP5 simulations. *J. Clim.* **2013**, *26*, 7708–7719. [[CrossRef](#)]
8. Mehrotra, R.; Sharma, A.; Bari, M.; Tuteja, N.; Amirthanathan, G. An assessment of CMIP5 multi-model decadal hindcasts over Australia from a hydrological viewpoint. *J. Hydrol.* **2014**, *519*, 2932–2951. [[CrossRef](#)]
9. Choudhury, D.; Mehrotra, R.; Sharma, A.; Gupta, A.S.; Sivakumar, B. Effectiveness of CMIP5 Decadal Experiments for Interannual Rainfall Prediction Over Australia. *Water Resour. Res.* **2019**, *55*, 7400–7418. [[CrossRef](#)]
10. Bhend, J.; Whetton, P. Evaluation of simulated recent climate change in Australia. *Aust. Meteorol. Oceanogr. J.* **2015**, *65*, 4–18. [[CrossRef](#)]
11. Moise, A.; Wilson, L.; Grose, M.; Whetton, P.; Watterson, I.; Bhend, J.; Bathols, J.; Hanson, L.; Erwin, T.; Bedin, T.; et al. Evaluation of CMIP3 and CMIP5 Models over the Australian Region to Inform Confidence in Projections. *Aust. Meteorol. Oceanogr. J.* **2015**, *65*, 19–53. [[CrossRef](#)]
12. Hossain, M.M.; Garg, N.; Anwar, A.H.M.F.; Prakash, M.; Bari, M. Intercomparison of drift correction alternatives for CMIP5 decadal precipitation. *Int. J. Clim.* **2021**, *42*, 1015–1037. [[CrossRef](#)]
13. Hossain, M.M.; Garg, N.; Anwar, A.H.M.F.; Prakash, M.; Bari, M. Drift in CMIP5 decadal precipitation at catchment level. *Stoch. Environ. Res. Risk Assess.* **2021**, *8*, 5. [[CrossRef](#)]
14. Choudhury, D.; Sharma, A.; Gupta, A.S.; Mehrotra, R.; Sivakumar, B. Sampling biases in CMIP5 decadal forecasts. *J. Geophys. Res. Atmos.* **2016**, *121*, 3435–3445. [[CrossRef](#)]
15. Knutti, R.; Furrer, R.; Tebaldi, C.; Cermak, J.; Meehl, G.A. Challenges in Combining Projections from Multiple Climate Models. *J. Clim.* **2010**, *23*, 2739–2758. [[CrossRef](#)]

16. McSweeney, C.F.; Jones, R.G.; Lee, R.W.; Rowell, D.P. Selecting CMIP5 GCMs for downscaling over multiple regions. *Clim. Dyn.* **2015**, *44*, 3237–3260. [[CrossRef](#)]
17. Islam, S.A.; Bari, M.A.; Anwar, A.H.M.F. Hydrologic impact of climate change on Murray-Hotham catchment of Western Australia: A projection of rainfall-runoff for future water resources planning. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3591–3614. [[CrossRef](#)]
18. Sheffield, J.; Camargo, S.J.; Fu, R.; Hu, Q.; Jiang, X.; Johnson, N.; Karnauskas, K.B.; Kim, S.T.; Kinter, J.; Kumar, S.; et al. North American Climate in CMIP5 Experiments. Part II: Evaluation of Historical Simulations of Intraseasonal to Decadal Variability. *J. Clim.* **2013**, *26*, 9247–9290. [[CrossRef](#)]
19. Meehl, G.A.; Teng, H. CMIP5 multi-model hindcasts for the mid-1970s shift and early 2000s hiatus and predictions for 2016–2035. *Geophys. Res. Lett.* **2014**, *41*, 1711–1716. [[CrossRef](#)]
20. Hossain, M.M.; Garg, N.; Anwar, A.H.M.F.; Prakash, M.; Bari, M. A comparative study on 10 and 30-year simulation of CMIP5 decadal hindcast precipitation at catchment level. In Proceedings of the MODSIM2021, 24th International Congress on Modelling and Simulation, Sydney, NSW, Australia, 5–10 December 2021; pp. 609–615. [[CrossRef](#)]
21. Frost, A.J.; Ramchurn, A.; Smith, A. *The Bureau's Operational AWRA Landscape (AWRA-L) Model*; Bureau of Meteorology Technical Report; Australian Government, Bureau of Meteorology: Melbourne, Australia, 2016.
22. Salathé, E.P. Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *Int. J. Climatol.* **2003**, *23*, 887–901. [[CrossRef](#)]
23. Grotch, S.L.; MacCracken, M.C. The Use of General Circulation Models to Predict Regional Climatic Change. *J. Clim.* **1991**, *4*, 286–303. [[CrossRef](#)]
24. Fowler, H.J.; Blenkinsop, S.; Tebaldi, C. Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* **2007**, *27*, 1547–1578. [[CrossRef](#)]
25. Hossain, M.M.; Garg, N.; Anwar, A.H.M.F.; Prakash, M. Comparing Spatial Interpolation Methods for CMIP5 Monthly Precipitation at Catchment Scale. *Indian Water Resour. Soc.* **2021**, *41*, 28–34.
26. Skelly, W.C.; Henderson-Sellers, A. Grid box or grid point: What type of data do GCMs deliver to climate impacts researchers? *Int. J. Climatol.* **1996**, *16*, 1079–1086. [[CrossRef](#)]
27. Jones, P.W. First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon. Weather Rev.* **1999**, *127*, 2204–2210. [[CrossRef](#)]
28. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2011; Volume 100, 676p.
29. Wilmot, C.J. Some Comments on the Evaluation of Model Performance. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 1309–1313. [[CrossRef](#)]
30. Roberts, N.M.; Lean, H.W. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* **2008**, *136*, 78–97. [[CrossRef](#)]
31. Jain, S.; Salunke, P.; Mishra, S.K.; Sahany, S. Performance of CMIP5 models in the simulation of Indian summer monsoon. *Theor. Appl. Climatol.* **2019**, *137*, 1429–1447. [[CrossRef](#)]
32. Sakamoto, T.T.; Komuro, Y.; Nishimura, T.; Ishii, M.; Tatebe, H.; Shiogama, H.; Hasegawa, A.; Toyoda, T.; Mori, M.; Suzuki, T.; et al. MIROC4h-A new high-resolution atmosphere-ocean coupled general circulation model. *J. Meteorol. Soc. Jpn.* **2012**, *90*, 325–359. [[CrossRef](#)]
33. Purwaningsih, A.; Hidayat, R. Performance of Decadal Prediction in Coupled Model Intercomparison Project Phase 5 (CMIP5) on Projecting Climate in Tropical Area. *Procedia Environ. Sci.* **2016**, *33*, 128–139. [[CrossRef](#)]
34. Homsí, R.; Shiru, M.S.; Shahid, S.; Ismail, T.; Bin Harun, S.; Al-Ansari, N.; Chau, K.-W.; Yaseen, Z.M. Precipitation projection using a CMIP5 GCM ensemble model: A regional investigation of Syria. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 90–106. [[CrossRef](#)]
35. Kamworapan, S.; Surussavadee, C. Evaluation of CMIP5 global climate models for simulating climatological temperature and precipitation for southeast Asia. *Adv. Meteorol.* **2019**, *2019*, 1067365. [[CrossRef](#)]
36. Kumar, D.; Kodra, E.; Ganguly, A.R. Regional and seasonal intercomparison of CMIP3 and CMIP5 climate model ensembles for temperature and precipitation. *Clim. Dyn.* **2014**, *43*, 2491–2518. [[CrossRef](#)]
37. Ta, Z.; Yu, Y.; Sun, L.; Chen, X.; Mu, G.; Yu, R. Assessment of Precipitation Simulations in Central Asia by CMIP5 Climate Models. *Water* **2018**, *10*, 1516. [[CrossRef](#)]
38. Chen, J.; Brissette, F.P.; Lucas-Picher, P.; Caya, D. Impacts of weighting climate models for hydro-meteorological climate change studies. *J. Hydrol.* **2017**, *549*, 534–546. [[CrossRef](#)]
39. Meehl, G.A.; Teng, H.; Maher, N.; England, M.H. Effects of the Mount Pinatubo eruption on decadal climate prediction skill of Pacific sea surface temperatures. *Geophys. Res. Lett.* **2015**, *42*, 10840–10846. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.