

## Article

# Distributed Hydrological Model Based on Machine Learning Algorithm: Assessment of Climate Change Impact on Floods

Zafar Iqbal<sup>1</sup>, Shamsuddin Shahid<sup>1</sup>, Tarmizi Ismail<sup>1,\*</sup>, Zulfaqar Sa'adi<sup>1,2</sup>, Aitazaz Farooque<sup>3</sup>  
and Zaher Mundher Yaseen<sup>4,5,6</sup>

<sup>1</sup> School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Johor, Malaysia; izafar2@graduate.utm.my (Z.I.); sshahid@utm.my (S.S.); zulfaqar@utm.my (Z.S.)

<sup>2</sup> Centre for Environmental Sustainability and Water Security (IPASA), Research Institute for Sustainable Environment (RISE), Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Johor, Malaysia

<sup>3</sup> Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE C1A4P3, Canada; afarooque@upei.ca

<sup>4</sup> Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia; yaseen@ukm.edu.my

<sup>5</sup> USQ's Advanced Data Analytics Research Group, School of Mathematics Physics and Computing, University of Southern Queensland, Toowoomba, QLD 4350, Australia

<sup>6</sup> New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar 64001, Iraq

\* Correspondence: tarmiziismail@utm.my



**Citation:** Iqbal, Z.; Shahid, S.; Ismail, T.; Sa'adi, Z.; Farooque, A.; Yaseen, Z.M. Distributed Hydrological Model Based on Machine Learning Algorithm: Assessment of Climate Change Impact on Floods. *Sustainability* **2022**, *14*, 6620. <https://doi.org/10.3390/su14116620>

Academic Editors: Fadong Li, Gang Chen, Xinlin He and Guang Yang

Received: 9 April 2022

Accepted: 20 May 2022

Published: 28 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/>).

**Abstract:** Rapid population growth, economic development, land-use modifications, and climate change are the major driving forces of growing hydrological disasters like floods and water stress. Reliable flood modelling is challenging due to the spatiotemporal changes in precipitation intensity, duration and frequency, heterogeneity in temperature rise and land-use changes. Reliable high-resolution precipitation data and distributed hydrological model can solve the problem. This study aims to develop a distributed hydrological model using Machine Learning (ML) algorithms to simulate streamflow extremes from satellite-based high-resolution climate data. Four widely used bias correction methods were compared to select the best method for downscaling coupled model intercomparison project (CMIP6) global climate model (GCMs) simulations. A novel ML-based distributed hydrological model was developed for modelling runoff from the corrected satellite rainfall data. Finally, the model was used to project future changes in runoff and streamflow extremes from the downscaled GCM projected climate. The Johor River Basin (JRB) in Malaysia was considered as the case study area. The distributed hydrological model developed using ML showed Nash–Sutcliffe efficiency (NSE) values of 0.96 and 0.78 and Root Mean Square Error (RMSE) of 4.01 and 5.64 during calibration and validation. The simulated flow analysis using the model showed that the river discharge would increase in the near future (2020–2059) and the far future (2060–2099) for different Shared Socioeconomic Pathways (SSPs). The largest change in river discharge would be for SSP-585. The extreme rainfall indices, such as Total Rainfall above 95th Percentile (R95TOT), Total Rainfall above 99th Percentile (R99TOT), One day Max Rainfall ( $R \times 1$  day), Five-day Max Rainfall ( $R \times 5$  day), and Rainfall Intensity (RI), were projected to increase from 5% for SSP-119 to 37% for SSP-585 in the future compared to the base period. The results showed that climate change and socio-economic development would cause an increase in the frequency of streamflow extremes, causing larger flood events.

**Keywords:** satellite rainfall; distributed hydrological model; flood forecast; machine learning; rainfall extremes

## 1. Introduction

The hydrological system involves a complicated interaction between various components [1]. The human interaction with some of these components has made it more intricate over time [2,3]. Hence, the hydrological system is a dynamically complex system that remained difficult to understand and a challenge to model due to its complexity [4]. Hydrological disasters like floods and water stress have become an every-year phenomenon in many other countries across the globe [5]. Floods in a catchment are triggered when precipitation becomes more than the storage and drainage capacity of the catchment [6,7]. Due to rapid population growth, economic development, land-use modifications, and climate change, many catchments across the world have become highly prone to hydrological disasters [8,9]. This is particularly true for Malaysia, where land use and climate changes are often mentioned as the responsible factors for the increased frequency and severity of urban water scarcity and floods [10,11]. This has caused major concern among scientists and policymakers in the context of global environmental changes.

The increase in atmospheric greenhouse gases (GHG) caused a significant rise in global temperature [12]. The changes in precipitation patterns, including intensity, duration, and frequency, have been recorded with the rise in temperature over the last few decades, resulting in frequent hydrological extremes [13]. Water is the most important resource for the survival of living beings [14]. Almost 80% of the world's population lives under different forms of water scarcity [15]. Increasing hydrological disasters may cause a quick depletion of the available water resources [16]. The water management system needs to be advanced with better management policy to attain sustainable development and management of water resources to adapt to climate change [17]. This needs reliable information on climate change projections and implications in catchment hydrological processes.

Rainfall–runoff models simulate the relationships between rainfall and the runoff generated in a catchment [18]. Various methods and techniques have been developed to simplify this complex relationship, ranging from a simple mathematical model to a complex “black box” and physical models [19–21]. According to the methods used to develop the relationship between rainfall and runoff, the models are categorized as empirical, conceptual, and physical [22]. They are also categorized as lumped, semi-distributed, and distributed models based on their ability to consider the spatial variability of catchment properties. Devia et al. conducted a comparative study to compare various rainfall–runoff models [22]. The study revealed that the empirical models require fewer input data but are limited to a certain region or a boundary, whereas the conceptual models are parametric. The parameters are catchment dependent, and thus, their derivations need large hydrological and meteorological data [23]. The physical-based model establishes the rainfall–runoff relationship based on the governing physical laws [24]. These models are most accurate but suffer from scale-related issues and require extensive data [22]. Therefore, they are considered the most complex rainfall–runoff models. The uncertainties associated with extensive data and the parameters used to develop models are specific to the region, making these models more time-consuming and site-specific [25].

In recent years, soft computing or machine learning (ML) methods, such as Artificial Neural Network (ANN), Support Vector Regression (SVR), and Fuzzy Logic and Genetic Algorithm (GA), have been employed to develop rainfall–runoff and other hydrological applications [26–29]. However, these approaches cannot completely manage the dynamics of hydrological processes because of their inherent limitations in the approaches [30]. Potential challenges also arise as these methods require long-term, continuous historical records of hydrological and other variables [31,32]. Furthermore, many of these approaches simplify the multi-factors and often make the nonlinear systems linear, reducing the simulation accuracy [33,34]. The hybridization of ML and conventional physical or conceptual model can improve the capability to model complex interactions. Such an approach also can replicate the functional relationship between input and output by enhancing the original methodologies by data processing, parameter estimation, and routing using machine learning algorithms [35]. The application of such complex problem-solving methodologies

in hydrology and water resources can help to provide a technique for reliable simulation of hydrological disasters, particularly water scarcity and floods, due to the changes in land use driven by physical and socio-economic factors and climate. Incorporating quantitative information on complex interactions of runoff with land use and climate can enhance the model's accuracy in simulating hydrological disasters [36].

The projection of water-related hazards in a catchment is very intricate due to the complex relationship of climate and land use with various ecological and socio-economic factors, including population growth, economic development, urbanization, and policy-related factors, like water management strategies and legislation [37]. Therefore, reciprocating actual hydrological conditions using hydrological models is always challenging [38]. A hydrological model requires a lot of observed data and optimizing different parameters [39]. The data availability or mismatch of any data leads to errors in simulation [40]. Therefore, the major challenge is finding the relationship among the water cycle components that affect a system in various dimensions. Successful simulation of a hydrological cycle using a dynamic approach can address hydrological modelling challenges. The solution to this problem is extremely important for Malaysia, where rapid population and economic growth along with climate and land-use changes have caused a significant change in hydrological disasters. Consequently, a moderate dry spell often forces water rationing, and moderate or extreme rainfall causes floods, especially in rapidly developing urban catchments of Malaysia [13].

The influence of land-use changes, water consumption, temperature rise and groundwater level causes changes in the hydrology of an area [12,41]. Deficiencies are found in studying the impact of climate changes, which are (i) the effect of changes due to a single component, (ii) statistical analysis of time series rather than assessing through a hydrological model, and (iii) not using the updated data for the study. There is a need to analyze the changes in hydrology with the combined effect of all such variables along with the hypothetical climate scenarios based on long-term climate observation of the specific region.

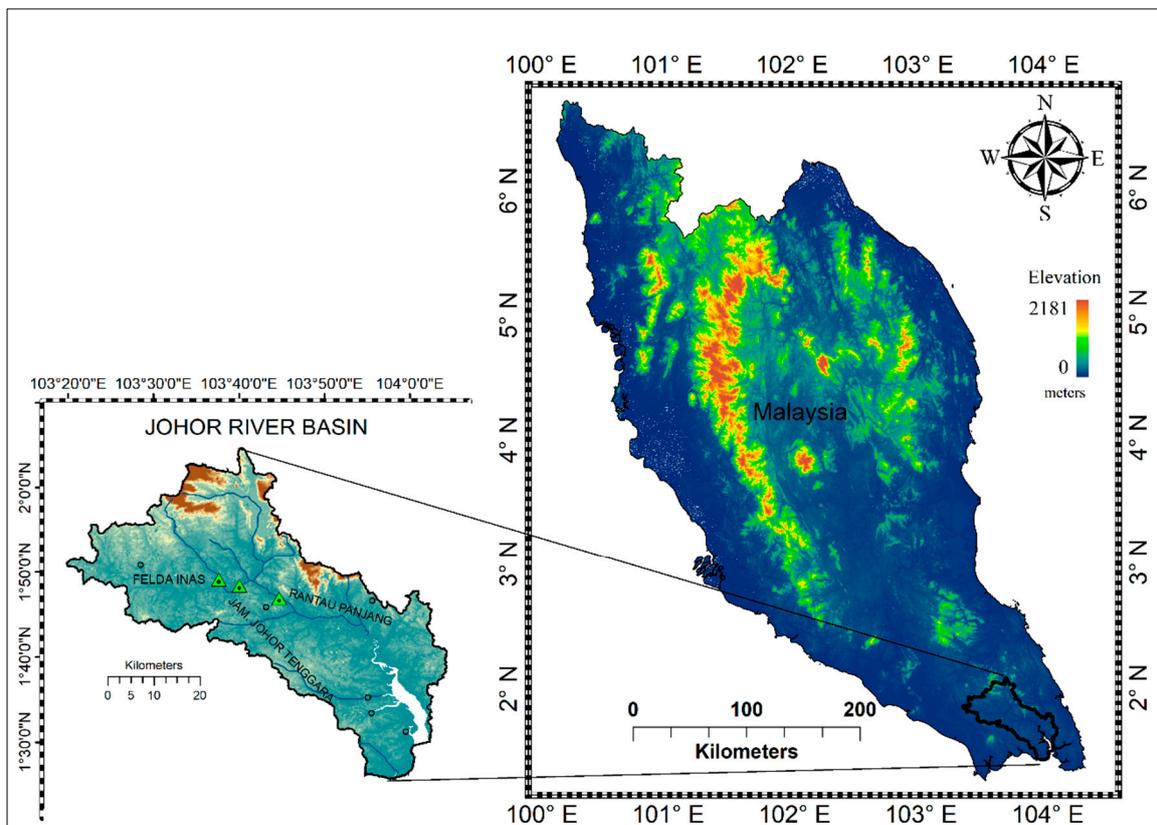
Modelling the dynamics of different factors individually and jointly can help understand the complex nonlinear interrelations and interactions among different elements in the complex physical, environmental, and behavioural systems [42]. The incorporation of quantitative information on complex interactions of various factors can enhance the prediction accuracy of the hydrological model to simulate hydrological disasters. It is expected that the application of complex problem-solving methodologies in hydrology and water resources will provide a reliable simulation of hydrological disasters, particularly water scarcity and floods, due to the changes in land use, climate and other physical and socio-economic aspects factors. Therefore, in this study, we develop an entity-based, distributed hydrological model based on state-of-the-art machine learning algorithms to incorporate various components of the environment to analyze the effect of climate and land-use changes on the flood susceptibility in the Johor River basin, Malaysia. Furthermore, this study includes the projection of rainfall and flood extremes under various SSP scenarios of CMIP6 GCM future projections.

## 2. Study Area and Data Description

### 2.1. Study Area

The study area opted in this research is Johor River Basin (JRB). It is situated in the south-eastern part of the state of Johor. The total catchment area of JRB is approximately 1652 km<sup>2</sup>. JRB is situated in Peninsular Malaysia (Figure 1), also known as West Malaysia. West Malaysia covers 130,598 km<sup>2</sup> and lies between latitudes of 1.20°–6.40° N and longitudes of 99.35°–104.20° E [43]. The JRB has undulating land with elevations ranging up to 366 meters in height. The topography comprises forests and irregular mountains sloping towards the South China Sea. The central and northern regions of JRB are covered with swamps and natural forests; however, in the southern region, rubber and oil palm plantations are the dominant landuse. Approximately 64% of the JRB have a slope angle

ranging from 0 to 50° [44]. The total length of JRB is 122.7 km with major tributaries of Peggeli River, Linggiu River, Sayong River, Jengeli River, and Belitong River [45].



**Figure 1.** Location of the study area.

Malaysia's climate is humid and hot due to its proximity to the equator. The region's rainforest climate is heavily influenced by Asian–Australian atmospheric dynamics and land–sea interaction, varying topography, and monsoon winds [46]. The average daily temperature ranges between 21 and 32 degrees Celsius, with an annual variation of 3 degrees Celsius. The annual average rainfall is approximately 2000–4000 mm, with 150 to 200 rainy days per year [47]. The regional precipitation distribution pattern is determined by the combined response of local topography and wind flow direction.

Peninsular Malaysia experiences two seasons throughout the year: The Southwest Monsoon (SWM) from May to August and the Northeast Monsoon (NEM) from November to February. During NEM, extreme rainfall events are common, but the weather is dry during SWM. Coastal places are affected by the NEM, whilst higher altitude areas are less affected by the monsoon. Peninsular Malaysia has humid weather, with the highest precipitation recorded during the 'inter-monsoon period.'

## 2.2. Data Description

River gauge data of JRB was collected from the Department of Drainage and Irrigation (DID) Malaysia. Daily discharge data of the main tributary was used to calibrate and validate the model. The details of the river gauge are given in Table 1.

**Table 1.** Description of River Flow data.

Station ID	Station Name	River Basin	Catchment Area (km <sup>2</sup> )	Analysis Period
1737451	SG. JOHOR at RANTAU PANJANG	Sg Johor	1130	2007–2017

ERA-5-Land is a post-processed reanalysis of the European Commission, the European Centre for Medium-Range Weather Forecasts (ECMWF). It has a higher resolution than the previous products, such as the ERA-interim and ERA-5. This product provides complete information on various climates and land variables over a longer period with higher resolution. Physical laws govern the model output generated to produce a consistent set of data by using observed data across the globe as an input. ERA-5 land contains the data of 50 variables which helps to study the energy and water cycles with a one-hour temporal resolution spread globally at a 9km resolution spatial resolution [48]. ERA-5-Land data contains over 50 climate variables. Downscaling the CMIP6 GCMs, long term continuous higher-resolution data is required. Therefore rainfall and temperature data of ERA-5-Land for the period 1981–2014 were used to downscale GCMs, whereas soil moisture was used for the development of the distributed hydrological model. All these data are freely available at <https://cds.climate.copernicus.eu> (accessed on 16 November 2021).

The CMIP6 GCMs were employed in this study for the simulation of the future runoff in the basin. The new multi-model ensemble of CMIP6 was used in this study. The model ensemble allows climate change evaluation and regional projections under various future socio-economic scenarios. The fourth and fifth IPCC reports of the Earth System model (ESM) and Atmosphere–Ocean General Circulation Models (AOGCMs) were coupled as an input for CMIP6, which are known as General Circulation Model (GCM). The GCMs were selected by Iqbal et al. [49] for (Mainland South East Asia) MSEA using a robust selection method that uses the categorical and spatial indices. They selected three GCMs for the region. These GCMs were downscaled in this study to use as an input to the hydrological model to simulate future floods. Further details on the GCMs used can be found in the article [49].

Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) performed better among five other satellite products for precipitation over Peninsular Malaysia. Iqbal et al. [50] bias corrected the IMERG data using rain gauge station data. They developed a two-step bias correction method to improve the performance of IMERG data up to 55% in RMSE. The method was also found better than the conventional bias correction method such as Linear scaling and quantile regression [50]. The details of the study and method can be found in Iqbal et al. [50]. This bias corrected IMERG dataset was used in this study to develop the distributed hydrological model over Johor River Basin.

The details of various other datasets used in this modelling are given in Table 2.

**Table 2.** Details of various parameters used in this study.

	Data Set	Resolution	Source
Land use	MODTBGA (MODIS/Terra Thermal Bands Daily L2G-Lite Global 1km SIN Grid V006)	1 km	<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a> (accessed on 13 June 2021)
Rainfall	MOD16A2 (MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500 m SIN Grid V006)	500 m	<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a> (accessed on 13 June 2021)
Land Surface Temperature	MOD11A1-MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid	1 km	<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a> (accessed on 14 June 2021)
Elevation	ALOS/PALSAR DEM 12.5 m	12.5 m	<a href="https://asf.alaska.edu/">https://asf.alaska.edu/</a> (accessed on 19 July 2021)
Soil Type	SoilGrids250m version 2.0	250 m	<a href="https://soilgrids.org/">https://soilgrids.org/</a> (accessed on 22 July 2021)

### 3. Methodology

#### 3.1. Procedure

The methodology adopted in this study consists of the following steps:

1. The catchment is divided into grids of 10 km each.

2. All the data sets are interpolated to 10 km to achieve a similar resolution.
3. The distributed hydrological model is developed using bias Corrected IMERG data for the catchment.
4. The model is calibrated and validated with the observed river flow data (details given in Table 1).
5. The selected GCMs are downscaled to 10 km resolution for the basin.
6. The downscaled GCMs data is used in the distributed model to simulate the future flow condition under different SSP scenarios. The details of the methods used to complete the analysis are given below.

### 3.2. K-Nearest Neighbour

The KNN is an efficient nonparametric classification algorithm that assigns data to a class based on its nearest neighbours [51]. In the particular classification problem, assuming that  $T = \{x_n \in R^d\}_{n=1}^N$  indicates a training set comprises of N samples within each M class in d-dimension; the sample  $x_n$  is assigned the class mark " $c_n$ ", the distance between the unknown point  $x$  and  $x_i^{NN}$  is estimated using Euclidean distance method as shown in Equation (1).

$$d(x, x_i^{NN}) = \sqrt{(x - x_i^{NN})^T (x - x_i^{NN})} \quad (1)$$

Next, the class name of the query point  $x$  is estimated based on the majority voting of its neighbours, as shown in Equation (2).

$$\hat{c} = \operatorname{argmax}_c \sum_{(x_i^{NN}, c_i^{NN}) \in \bar{T}} \delta(c = c_i^{NN}) \quad (2)$$

where  $c$  is a class label and  $c_i^{NN}$  is the class label of  $i$ -th nearest neighbour.  $\delta(c = c_i^{NN})$ , an indicator function, can have a value of one of the class  $c_i^{NN}$  of the neighbour  $x_i^{NN}$ . This research used KNN to interpolate different data sets to a specific grid.

### 3.3. Downscaling of GCMs

GCM simulations were downscaled to a finer spatial scale for their use by end-users. This study used the MOS approach to downscaling the selected GCMs into fine resolution. In MOS, the statistical calibration between simulated and observed predictors is usually done [52]. The advantage of using this method is that it improves reliability while keeping the original accuracy. The MOS has been found more advantageous in studies related to climate change [53].

The bias correction using MOS can be expressed as in Equations (3) and (4).

$$Bias = \frac{1}{N} \sum_{t=1}^N [F(t) - O(t)] \quad (3)$$

where  $F$  = Forecast;  $O$  = Observations and  $N$  = days in training sample.

$$F' = F(t) - Bias \quad (4)$$

where  $F'$  = Corrected Value.

#### 3.3.1. Gamma Quantile Mapping

Gamma Quantile mapping (GammaQM) was introduced by Piani et al. [54]. This method assumes that a gamma distribution approximates the observed and simulated intensity distribution well. GammaQM builds a model variable  $P_m$  using probability

integral transform to make the new build distribution equal to the distribution of the observed variable  $P_o$ . The mathematical expression of this method is given in Equation (5)

$$P_o = F_o^{-1}(F_m(P_m)) \quad (5)$$

where  $F_m$  = Cumulative function of  $P_m$ ; and  $F_o^{-1}$  = inverse cumulative function of  $P_o$ .

$$pdf(x) = \frac{e^{(-\frac{x}{\theta})} x^{(k-1)}}{\Gamma(k)\theta^k} \quad (6)$$

where in Equation (6),  $k$  signifies the form parameter,  $x$  denotes the Normalized daily precipitation, while  $\theta$  denotes the scaling parameter.

GammaQM could not be applied if the  $k$  value is less than 1 or 0; therefore, the value is presumed greater than 1. GammaQM deliberates mean and extreme values, making it an effective bias-correction method [54–56]. GammaQM is only valid for precipitation data.

### 3.3.2. Power Transformation

Power Transformation (PowerTr) considers the bias in the mean and the differences in the variance for the correction of data [57]. In power transformation, a nonlinear correction in the exponential form such as  $aP^b$  is used for the adjustment of variance. According to this method daily precipitation of  $P$  was transformed into a corrected amount of  $P^*$  using Equation (7).

$$P^* = aP^b \quad (7)$$

A distribution-free approach can calculate the parameter,  $b$ . It is first identified by matching the coefficient of variation (CV) corrected daily GCM precipitation ( $P^b$ ) with the CV of observed daily precipitation for each month of training. The value of  $b$  was determined iteratively. Data grouping was done every five days to reduce sampling variability [58]. The value of  $b$  was used to calculate the transformed precipitation by Equation (8):

$$P^* = P^b \quad (8)$$

A parameter is subjected to the observed and the transformed mean values. It is a dependent parameter to the value of  $b$  parameter and subsequently  $b$  is a dependent to the magnitude of CV Both  $a$  and  $b$  are differed for every block-annual of 5 days.

### 3.3.3. Generalized Quantile Mapping

Generalized quantile mapping (GenQM) is a kind of parametric quantile mapping. Its main differ is the implementation of gamma distribution and generalized Pareto distribution (GPD). Mathmatically, GenQM can be expressed as follows:

$$P_o = F_o^{-1}(F_m(P_m)) \quad (9)$$

The pdf is chnaged with the value of GPD and gamma distribution. The value of the GPD is tailed the extreme distribution [59], as expressed in Equation(10).

$$\Pr(X - u \leq x | X > u) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}}\right), & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{\xi x}{\tilde{\sigma}}\right), & \text{if } \xi = 0 \end{cases} \quad (10)$$

Here, the  $u$  value is the 95th% threshold,  $\tilde{\sigma} = \sigma + \xi(u - \mu)$ ,  $\tilde{\sigma}$  is the scale parameter.  $\xi$  is presented the shape parameter. In this equation, gamma distribution was employed on a smaller threshold. In addition, the GPD was employed on values larger than this threshold, as given in Equation (11):

$$y = \begin{cases} F_{\text{obs, gamma}}^{-1}(F_{\text{CCLM, gamma}}), & \text{if } x < 95 \text{ thpercentile} \\ F_{\text{obs, GPD}}^{-1}(F_{\text{CCLM, GPD}}), & \text{if } x \geq 95 \text{ thpercentile} \end{cases} \quad (11)$$

### 3.3.4. Linear Scaling

The Linear Scaling (LS) approach was introduced by Lenderink et al. [60]. This bias correction method utilizes the monthly correction values “the difference modeled and observed daily dataset”. The monthly scaling factor is then applied to uncorrected daily data. The daily precipitation  $P$  was corrected by the following Equations (12) and (13).

$$P^* = \alpha P \quad (12)$$

While the temperature,  $T$ , is corrected using the following equation,

$$T^* = \alpha T \quad (13)$$

whereas  $\alpha$  is the monthly scaling factor for precipitation is calculated by Equation (14),

$$\alpha = \frac{P_o}{P_s} \quad (14)$$

$P_o$  is the observed monthly mean and  $P_s$  is the simulated monthly mean.

For the bias correction of temperature, the scaling factor is calculated by Equation (15).

$$\alpha = T_o - T_s \quad (15)$$

$T_o$  is the observed temperature mean whereas,  $T_s$  is the monthly mean simulated temperature. The LS method is simple and requires less information, such as only monthly data is required to calculate the scaling factor [61].

## 3.4. Hydrological Model Development

Hydrological interactions, such as transpiration, evaporation, streamflow, rainfall, groundwater flow, and infiltration constitute a hydrological system. The interaction among the hydrological system’s components is complex and variable in space and time. However, four major components mostly govern the hydrological cycle: precipitation, infiltration, runoff, and evapotranspiration. Various methods are adopted to develop the relationships between these major hydrological components and understand the hydrology of any region. The major interacting components of a hydrological cycle are shown in Figure 2. The study is divided into several grid boxes to model the distributed nature of hydrological processes. The divisions of the study area into grid boxes are also shown in Figure 2.

### 3.4.1. Concept of the Distributed Model

A multiple bucket modelling approach was used in this research to account for the spatial variability of the land and climate variables in the catchment. The study area was divided into grids of 10 km each to calculate the cumulative flow in the catchment, considering the variability of soil and climate inputs at a coarser scale. Each bucket was subdivided based on the major hydrological processes. The unsaturated region, evapotranspiration, and surface runoff simulate the flow generation. The model was developed for JRB on a daily time scale using the bias-corrected satellite data. A simple water balance equation was used for each bucket, including rainfall, evaporation, and saturated excess flow. A nonlinear storage discharge relationship was established to generate surface runoff using near-real-time RS data for creating an early flood hazard system to minimize the flood causalities.

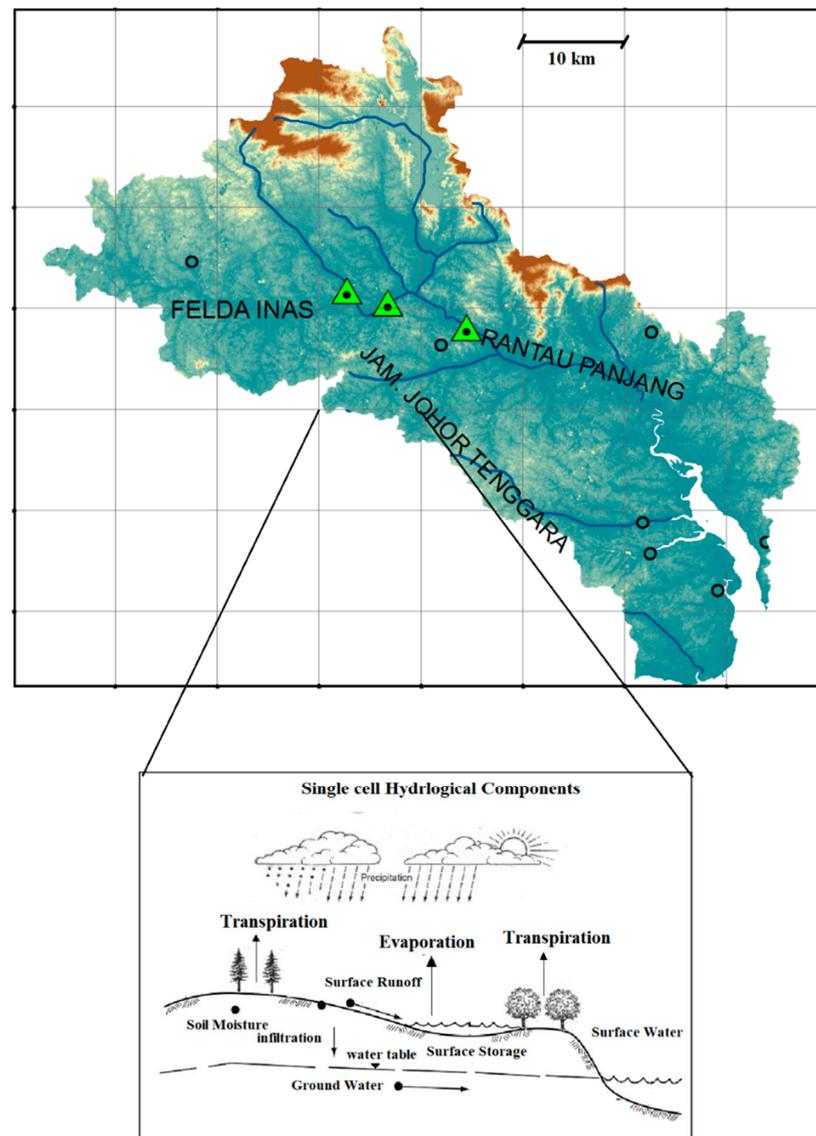


Figure 2. Distribution of the study area into grids.

The main water balance equation used to calculate the discharge from each bucket is given in Equation (16).

$$\frac{d_s(t)}{dt} = R(t) - f_{ss}(t) - f_{se}(t) - et(t) \quad (16)$$

where  $R(t)$  is the amount of rainfall at a certain timestep  $t$ .  $q_{ss}(t)$  and  $q_{se}(t)$  are the runoff at subsurface and a surface scale, whereas  $et$  represents the evapotranspiration for the pertinent bucket.

#### 3.4.2. Excess Saturation Runoff Rate

The overland flow or saturated excess flow of a bucket was calculated by Equation (17), given the following conditions were fulfilled.

$$f_{se} = \frac{V - V_b}{\Delta t} \quad \text{if } V > V_b \quad f_{se} = 0 \quad \text{if } V < V_b \quad (17)$$

where  $V$  is the volume of soil water storage and  $V_b$  is the soil moisture storage capacity.  $V_b$  depends on the average soil depth ( $L$ ) and the average soil porosity  $\Phi$ .

### 3.4.3. Subsurface Runoff

The subsurface runoff is the function of soil storage and catchment response time. The function to calculate the subsurface runoff depends on the values of soil water storage capacity and the soil water storage at the pertinent grid, as given in Equation (18) below.

$$f_{ss} = \frac{V - V_f}{t_c} \quad \text{if } V > V_f \quad f_{ss} = 0 \quad \text{if } V < V_f \quad (18)$$

where  $V$  represents the soil water storage and  $V_f$  is the threshold storage assumed to be equal to  $V_f = f_c L$ , the product of soil field capacity and the average soil depth. Darcy's Law is used to calculate the catchment response time considering the hydraulic gradient equal to the hillslope of the ground calculated using the DEM. The equation for the calculation of catchment response time is given in Equation (19),

$$t_c = \frac{L \Phi}{2 K_s \tan \beta} \quad (19)$$

where  $L$  is the hillslope length,  $\tan \beta$  is the average ground surface slope, and  $K_s$  is the average saturated hydraulic conductivity.

### 3.4.4. Evapotranspiration

The evapotranspiration in the water balance model is calculated using an empirical relationship that uses minimum parameters. The FAO Blaney–Criddle was used in this study to find the evapotranspiration using the precipitation and temperature at a specific grid [62], using Equation (20) below.

$$et_{o(i)} = p_{(i)} \left( 0.46 T_{mean(i)} + 8.13 \right) \quad (20)$$

where,  $p_{(i)}$  is the average precipitation and  $T_{mean(i)}$  is the mean temperature of the grid ( $i$ ).

### 3.4.5. Flow Routing

This distributed hydrological model's flow routing relies on the "Eight Direction Pour point model". It calculates the direction of flow of a single grid based on the difference in elevation of the surrounding eight grids. Furthermore, the flow direction of each cell is determined using the "Direction of steepest descent" method. The steps to calculate the flow routing are as follows:

1. The average elevation of each grid is calculated for all the cells.
2. The flow direction of each cell is calculated using the Eight Direction Pour point model.
3. The flow accumulation in each cell is calculated using the bucket model developed in Section 3.3.1.
4. Flow accumulation is calculated by adding the cumulated flow of the grids flowing into the particular grid
5. The flow route is calculated by connecting the low water accumulated cells with high water accumulated cells.

A machine learning model was used to generate runoff from rainfall in each grid cell. R-packages "r.watershed" tool and "rdwplus" were used to estimate the routing of the generated runoff from the cell to the catchment outlet.

### 3.4.6. Projections of Climate Change Impacts on Hydrological Extremes

The framework to analyze the impact of climate changes on hydrological extremes is shown in Figure 3. The extremes in the flow at Ratu Panjang station were simulated for historical and future scenarios. To simulate the flow using the spatially distributed hydrological model developed in this study, the downscaled GCMs data was used as an input. The model generated the flow for historical and future scenarios for various SSPs.

The output was used to calculate various flow quantiles, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. The changes in each quantile were calculated by taking the difference between simulated future flow to the historical flow by each GCM.

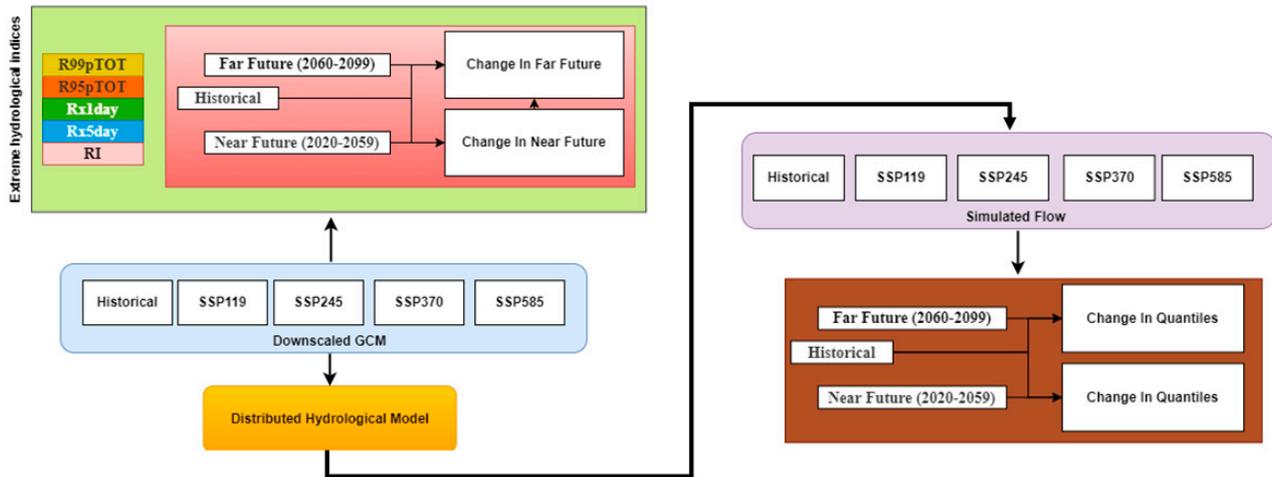


Figure 3. Framework to analyze the impact of climate change on hydrological extremes.

The downscaled GCMs data was used to analyze various rainfall extremes, as given in Table 3. These indices were calculated for historical and each SSP. The difference in these indices was calculated with the historical period as the reference. The details of each index are shown in Table 3.

Table 3. WMO hydrological extreme indices used in this study.

Indices	Symbol	Description	Formula
Total rainfall above 95th Percentile	R95pTOT	Annual total rainfall when rainfall > 95p	$R95p = \sum_{w=1}^W RR_{wj}^*$ where $RR_{wj} > RR_{wn95}$
Total Rainfall above 99th Percentile	R99pTOT	Annual total rainfall when rainfall > 99p	$R99p = \sum_{w=1}^W RR_{wj}^*$ where $RR_{wj} > RR_{wn99}$
One day Max Rainfall	$R \times 1 \text{ day}$	Annual maximum 1-day rainfall	$Rx1day_j = \max(RR_{ij}^{**})$
Five-day Max Rainfall	$R \times 5 \text{ day}$	Annual maximum 5-day rainfall	$Rx5day_j = \max(RR_{ij}^{**})$
Rainfall Intensity	RI	Average rainfall on the rainy days	$RI = \frac{\sum_{w=1}^w RR_{wj}}{W^{***}}$

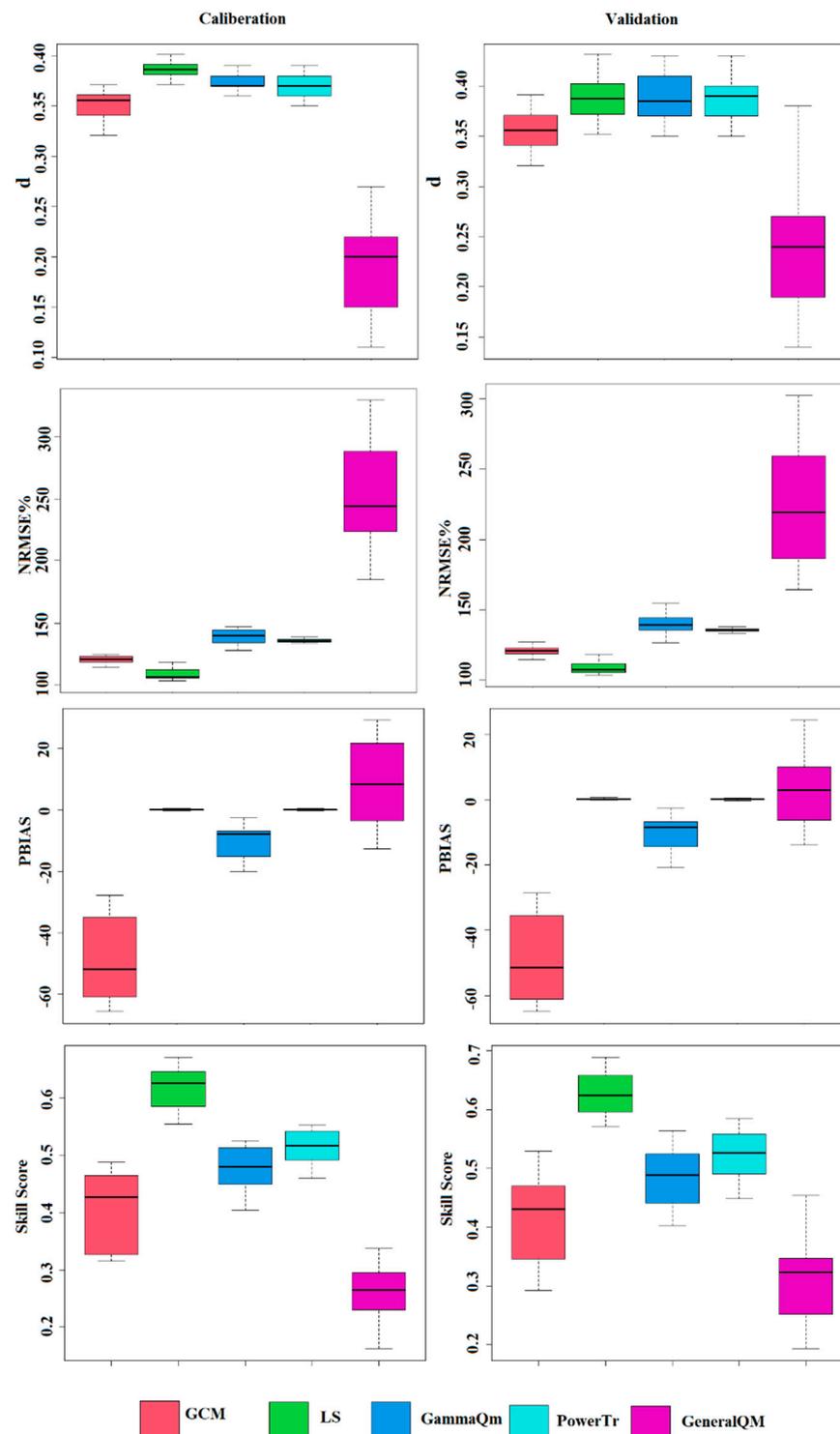
\* Daily Rainfall amount on wet days (Rainfall > 0). \*\* Daily rainfall amount on the day, *i*, in period *j*. \*\*\* Number of wet days (Rainfall > 0).

### 4. Application Results

#### 4.1. Downscaling of GCMs

##### 4.1.1. Downscaling of Precipitation

ERA-5-Land data was used as the reference data to downscale the precipitation and temperature of EC-Earth, EC-Earth-Veg, and MRI-ESM2. LS, GammaQM, PowerTr, and GenQM were used to downscale the historical GCMs. Index of agreement (*d*), Normalized Root Mean Square Error (NRMSE), Percentage Bias (Pbias), and Skill Score (SS) of the downscaled precipitation are shown in Figure 4. The downscaled result of one out of three GCMs are shown below. The results of the remaining GCMs are provided as supplementary material for further reference (Figures S1, S2, S5, S6, S9 and S10). The results showed that LS performs better than the other bias correction methods. Compared to the other methods, LS has improved the *d* values by up to 20% for each GCM.



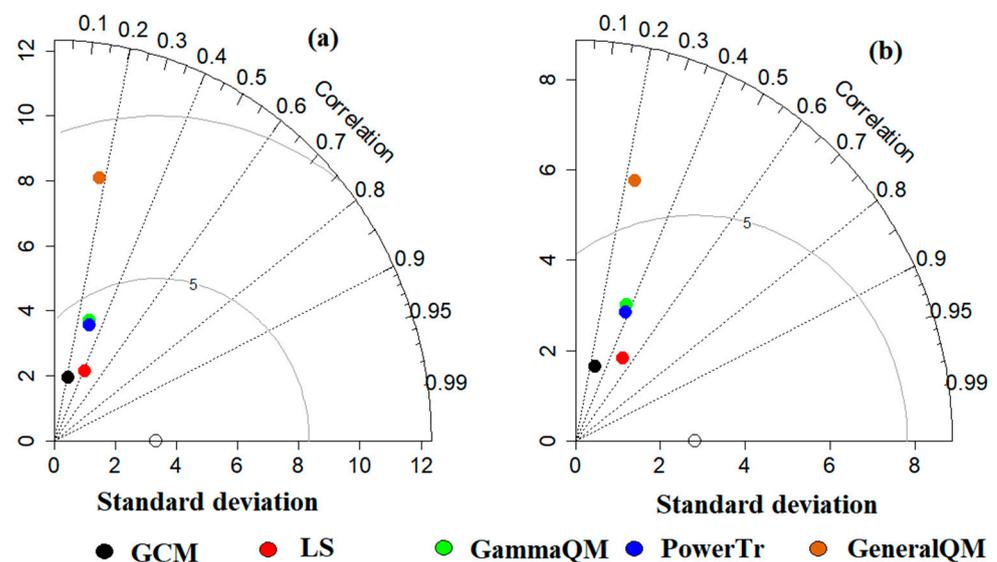
**Figure 4.** The statistical performance of the downscaled and raw GCMs.

The error in downscaling was compared using NRMSE %. Plots between ERA-5 and GCM with different bias correction methods are shown in Figure 4. The results show that the NRMSE of raw GCM ranges between 120 and 130%, whereas the LS reduced the NRMSE by 100–110%. GammaQM, PowerTr, and GenQM showed poor performance in reducing the NRMSE.

The PBIAS in the bias-corrected outputs is shown in Figure 4. The results showed that the raw GCM biases range from  $-40$  to  $-45\%$  compared to ERA-5 data. The most suitable model which reduced the bias very close to zero was the LS and PowerTr. The SS

of the raw and the bias-corrected GCMs were also compared to show the model's accuracy. The best SS was found for the LS. The raw GCM of EC-Earth showed a mean SS of 0.42, whereas it improved to 0.63 using LS, 0.48 using GammaQM and 0.53 using PowerTr, and it reduced to 0.25 for GenQM, as shown in Figure 4. Similar improvements were observed for EC-Earth-Veg and MRI-ESM2 using the LS method.

Taylor diagram was used to compare the degree of correspondence between the bias-corrected data, as shown in Figure 5. The figure shows the bias-corrected outputs for the EC-Earth during calibration and validation periods. The results showed that, in terms of three statistical matrices (Standard Deviation, Correlation and RMSE), the LS method performed better to reduce the bias in both the calibration and validation periods. The LS showed a correlation coefficient higher than 0.4 during calibration and validation, while all other methods showed less than 0.4. The root-mean-square error of LS corrected data was less than the other models, while the standard deviation was nearer to the observed one as it is radially nearer to the observation (hollow circle on the x-axis). The Taylor diagrams of the remaining GCMs are provided as supplementary materials (Figures S3, S4, S7, S8, S11 and S12).



**Figure 5.** Taylor diagram shows bias correction methods' performance to correct EC-Earth during (a) calibration; (b) validation periods.

#### 4.1.2. Downscaling of Maximum Temperature

The comparison of  $d$  in Figure 6 showed that the PowerTr downscaling method improved the  $d$  value from 0.48 to 0.56, a comparison of NRMSE % values showed that the GammaQM and GenQM methods failed to downscale the GCM because the NRMSE % values increased for these two models. However, LS and PowerTr showed a slight improvement in the NRMSE % by 5–10%.

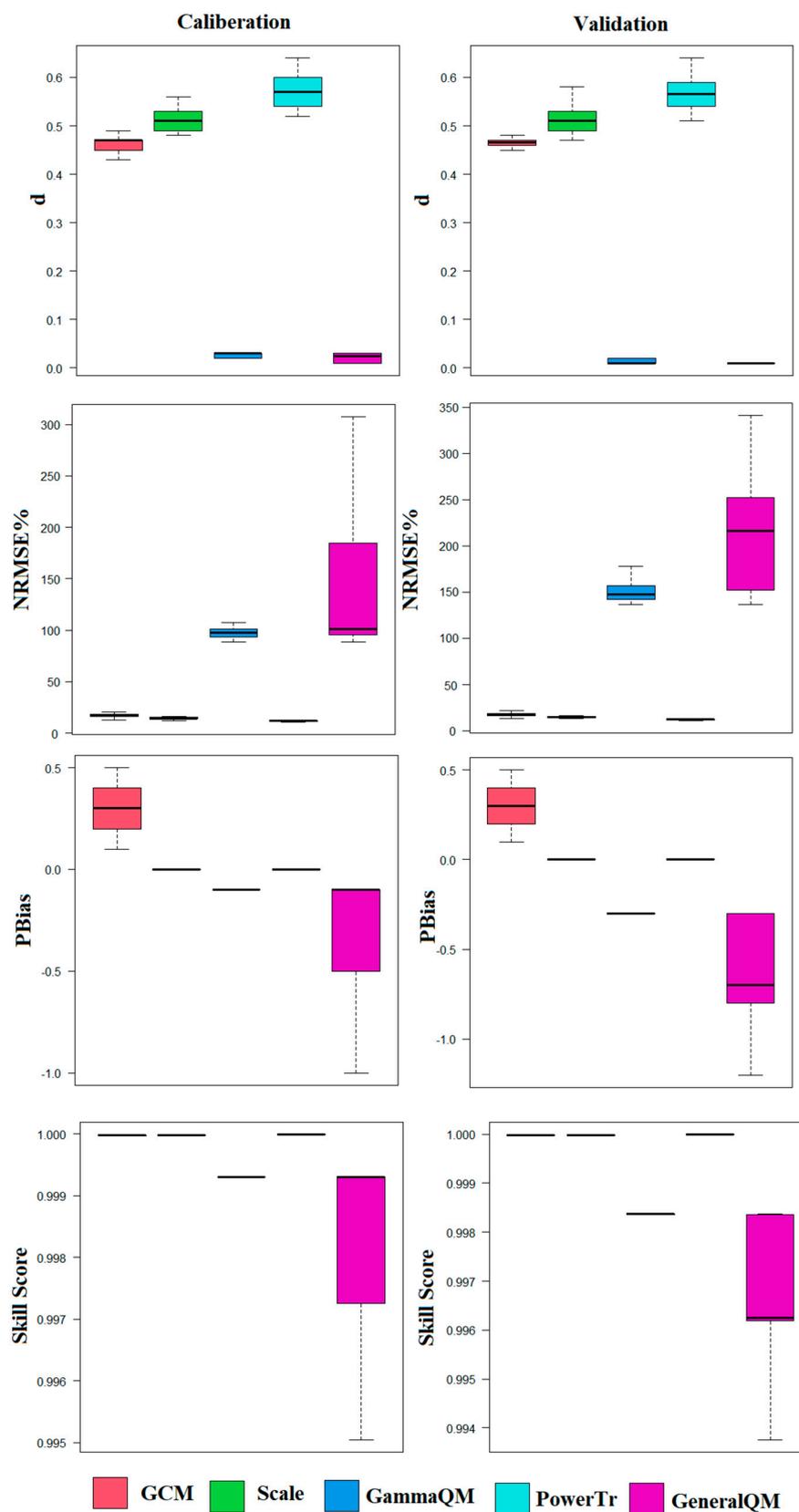
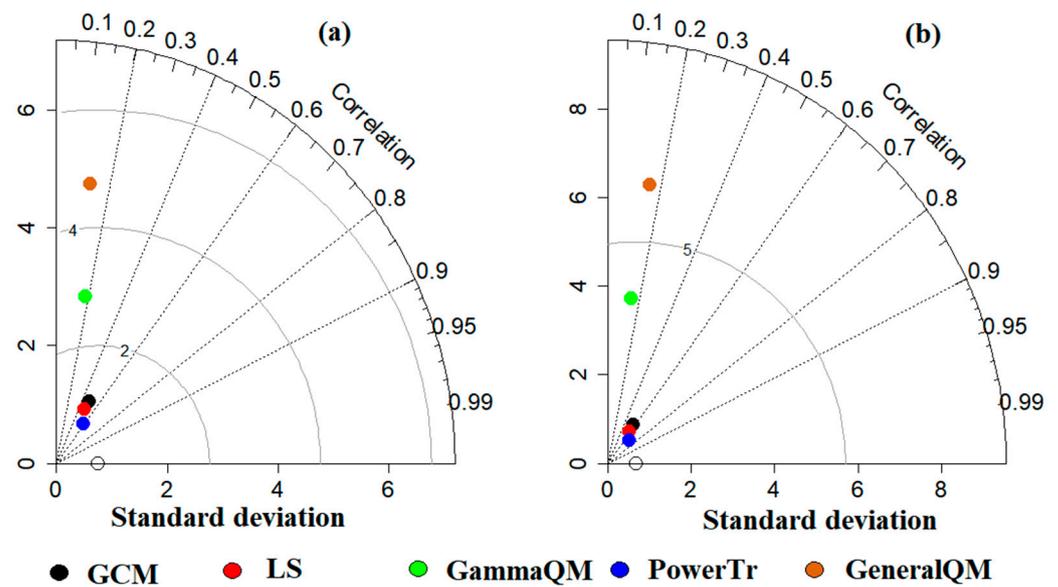


Figure 6. The statistical performance of the downscaled and raw GCMs for maximum temperature.

The output of the downscaling models is compared in Figure 6. The average percentage biases in downscaled data compared to the ERA-5 ranged from 0.1 to 0.4. The results showed that the PowerTr model reduced the biases in the three GCM by an average

of 20–30%. Similar results were demonstrated by the LS method, whereas the GenQM and GammaQM showed unsatisfactory performance in the bias correction. The SS of the models was also compared to the ERA-5 data, as shown in Figure 6. The SS of raw GCM was 0.995, which was further improved up to 0.999 by the PowerTr method in most cases. However, the SS was reduced in the case of GammaQM and GenQM. Therefore, in terms of improving these indices, the PowerTr downscaling method proved to be a better model than the others.

Figure 7 shows the Taylor diagram for EC-Earth GCM. The calibration and validation period results showed that the best model to downscale the EC-Earth is the PowerTr as it reduced the RMSE and increased the correlation. The results of EC-Earth-Veg and MRI-ESM2 are given in appendices.



**Figure 7.** Taylor diagram shows bias correction methods' performance to correct EC-Earth during (a) calibration; (b) validation period.

#### 4.1.3. Downscaling of Minimum Temperature

The results for downscaling minimum temperature are presented in this section. Figure 8 shows the efficacy of the downscaling models in terms of the  $d$ . The EC-Earth raw values showed a  $d$  value of 0.5, whereas the downscaling models showed improvement, specially PowerTr increased the  $d$  values up to 0.57. Figure 8 shows the NRMSE % values of GCM and other downscaled output compared to the ERA-5 historical data. The NRMSE % values of the GCM were observed between 0 and 50% for all the three GCMs. However, the improvement in MRI-ESM2 using PowerTr and LS was up to 10% for NRMSE %.

The biases in the downscaling GCMs are compared in Figure 8. The average biases in these models compared to the ERA-5 were in the range of  $-0.2$  to  $0.7\%$ . The results showed that the PowerTr model reduced the biases in the three GCM by an average of 20–30%. Similar results were shown by the LS method, whereas the GammaQM and GenQM showed very high errors in the bias correction. The SS of the models were also compared to the ERA-5 data. The SS of raw GCM was 0.997, which was improved up to 0.999 by the PowerTr method in most cases. However, the SS was noticed to decrease for GammaQM and GenQM.

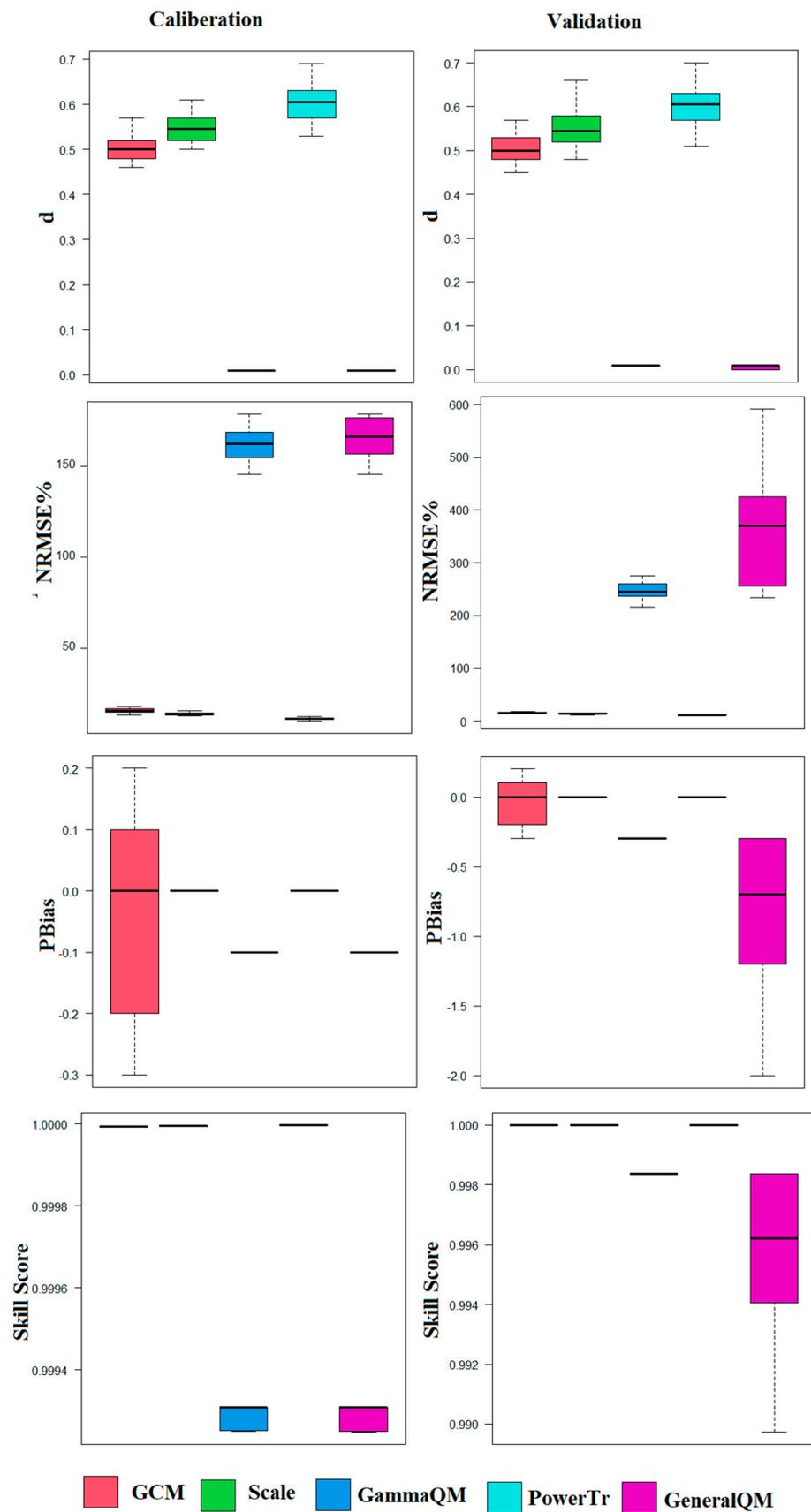
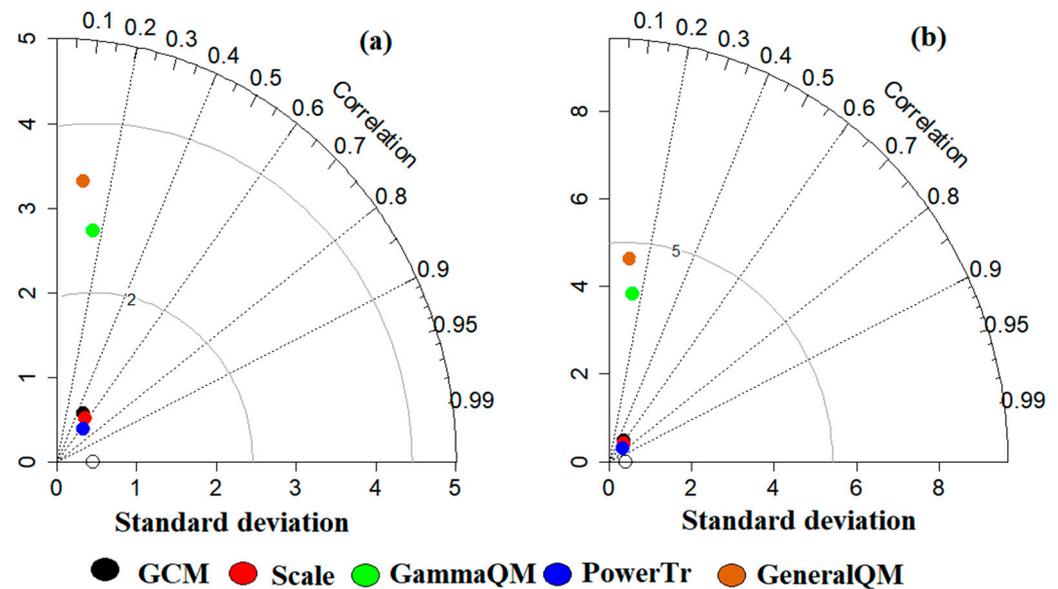


Figure 8. The statistical performance of the downscaled and raw GCMs for minimum temperature.

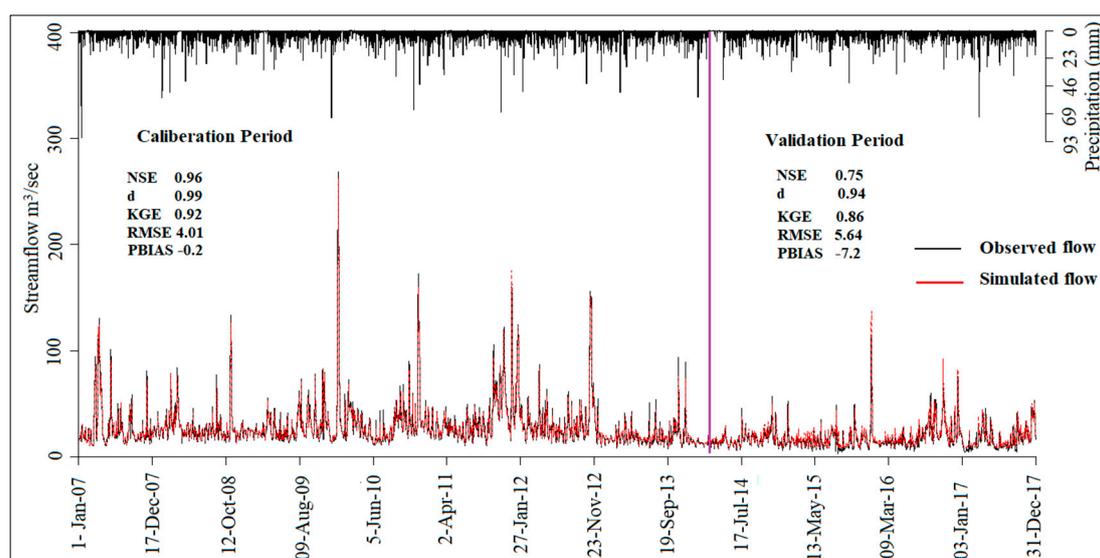
Figure 9 shows the Taylor diagram for EC-Earth minimum temperature. The calibration and validation period show that the best model to downscale the EC-Earth is the PowerTr as it has reduced the RMSE and increased the correlation.



**Figure 9.** Taylor diagram shows bias correction methods' performance to downscale EC-Earth minimum temperature during (a) calibration; (b) validation periods.

#### 4.2. Calibration and Validation of Hydrological Model

The integrated hydrological model was developed for each grid using the BIMERG data. Model calibration and validation were performed using river flow data at Ratu Panjang from 2007 to 2017. Seven years' data starting from 2007 was used for calibration, whereas the remaining four years of river flow data were used for validation. The model was developed using RF by tuning the parameters using repeated cross-validation for random sampling. The 'repeatedcv' package in R was used to split the data into ten parts. Nine parts were used to train and the remaining part for validation in each of the ten iterations used. The performance during each iteration was measured using evaluation metrics. The average performance of ten folds with ten repetitions was calculated to summarize the performance. The calibration and validation results are shown in Figure 10. The model showed good performance for the validation period, giving the NSE,  $d$ , KGE, RMSE, and Pbias of 0.96, 0.99, 0.92, 4.01, and  $-0.2$ , respectively. The NSE value of 0.96 was much better than the reported NSE values for similar other models, such as SWAT and HSPF, APEX, and SAC-SMA [63]. Conventionally, the NSE value greater than 0.65 is considered good for model evaluating criteria [64]. The model showed satisfactory values for the coefficient of determination ( $d$ ) during calibration (0.99) and validation period (0.94). The results also showed very less bias in calibration and validation periods. Pbias showed overestimation or underestimation of the measured flow. The acceptable range of Pbias is less than 10%, whereas in this case, the Pbias was  $-0.2\%$  for calibration and  $-7.2\%$  for the validation period. The RMSE values range from 4.01 to 5.64 for calibration and validation. The KGE values were also in the acceptable range of 0.92 for calibration and 0.86 for validation.



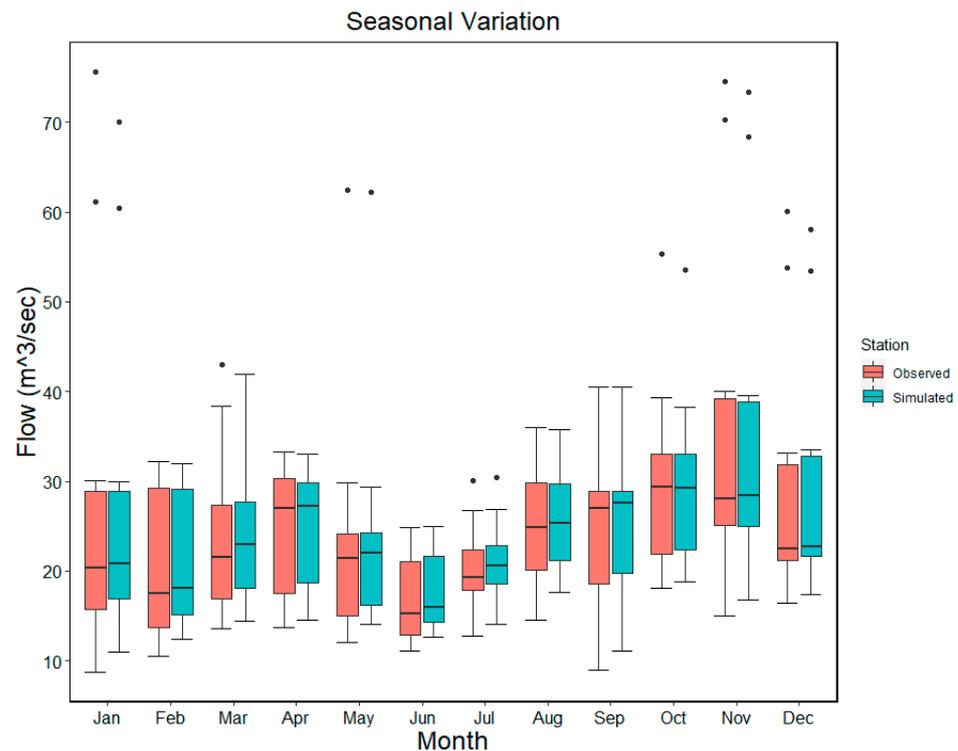
**Figure 10.** Observed and modelled river flow during calibration and validation period of 2007–2017.

The model was evaluated using other statistical indices, such as MAR,  $d$ ,  $md$ , KGE, RMSE, and Pbias. The results of these statistical indices are given in Table 4. The correlation terms, such as  $d$ ,  $md$ , and  $R^2$  were in the acceptable range, showing a good performance in simulating the runoff during the calibration and validation period. The error terms showed very negligible values, indicating the model's good performance in simulating observed flow.

**Table 4.** Performances of the model during the calibration and validation periods.

	MAE	RMSE	NRMSE%	Pbias	NSE	$d$	$md$	$R^2$	KGE
Caliberation	2.24	4.01	20.2	−0.2	0.96	0.99	0.91	0.97	0.92
Validation	3.8	5.64	50.2	−0.7	0.75	0.94	0.76	0.78	0.86

A boxplot of simulated and observed flow for all the months during the analysis period (2007–2017) was plotted to analyze the seasonal streamflow variations. Figure 11 shows that the model simulates the seasonal variation well. The mean and quantile ranges of the mean monthly rainfall of each month depicted a very good range of values. It can be observed in Figure 11 that the extreme values of the streamflow during January, May, July, October, November, and December were also well simulated by the model. The simulated mean values of each month were approximately equal to the observed flow. The result indicates that the model can simulate the seasonal variation and the seasonal extremes. Therefore, it can be used for climate change impact on river flow in the basin.



**Figure 11.** Seasonal variations in mean monthly stream flow.

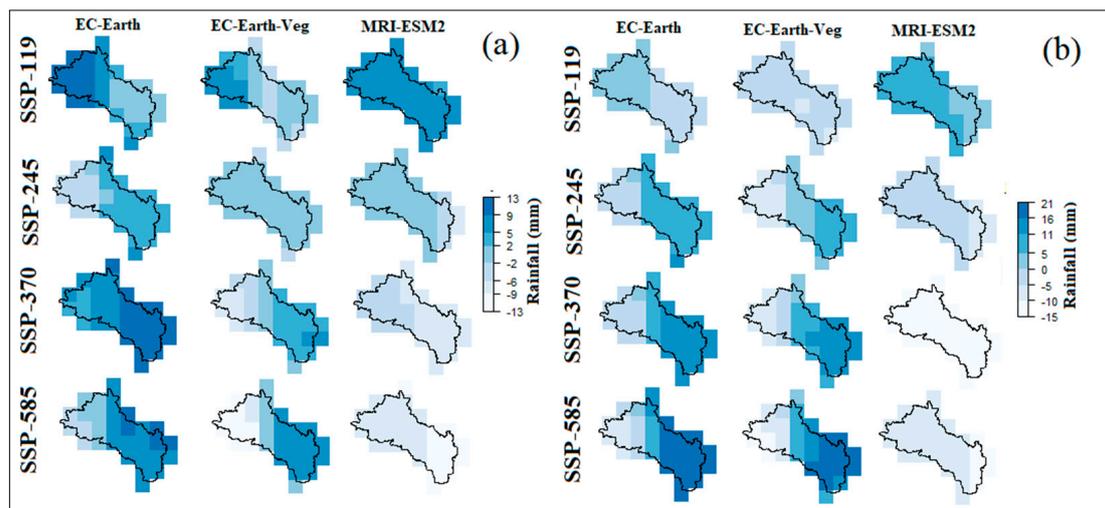
#### 4.3. Hydrological Changes under Future Scenarios

##### 4.3.1. Projected Rainfall Extremes

This section evaluated five extreme rainfall indices defined by WMO for four SSPs and three GCMs. The indices include R95pTOT, R99pTOT, Rx1day, Rx5day, and RI. The historical and future precipitation simulations of the most suitable GCMs. (As discussed in our previous paper [49], EC-Earth, EC-Earth-Veg, and MRI-ESM2 were used to assess the changes in precipitation extremes. The simulated extremes were analyzed for each GCM individually to cover the maximum uncertainty range in the near (2020–2059) and far (2060–2099) future. The changes in these indices compared to the historical period are discussed in the following sections).

##### Total Rainfall above 95th Percentile (R95pTOT)

Figure 12a shows the changes in R95TOT for different SSPs using three GCMs compared to the historical period. The results for EC-Earth showed the highest increase in R95pTOT for SSP-119. The increase was 13 in the northern part of JRB. For SSP-370, a moderate increase over the whole basin ranging from 2 to 10 mm was observed. However, SSP-585 projected a decrease in the northern part but an increase in total annual rainfall in the southern region up to 10 mm.



**Figure 12.** Change in annual total rainfall above 95 percentile (a) in the near future; (b) far future.

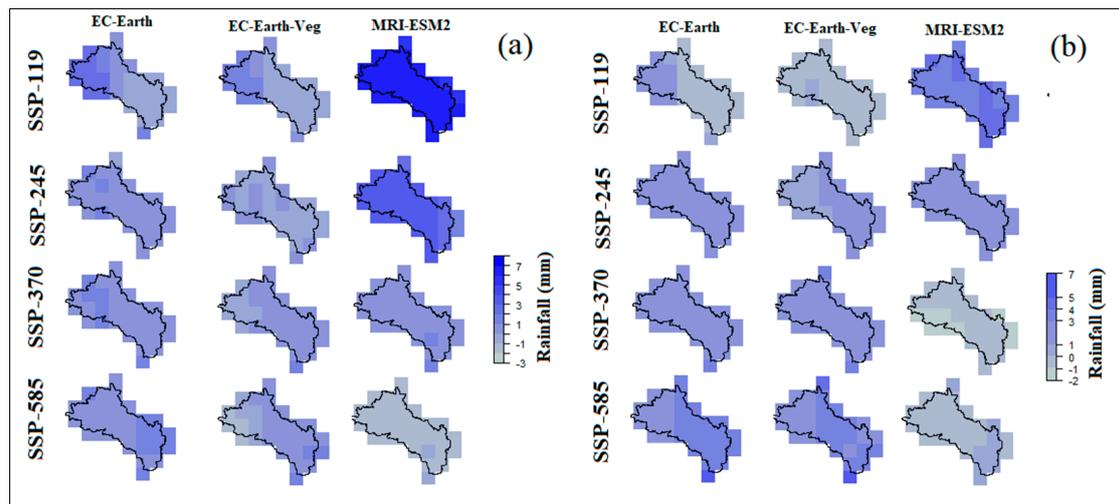
The EC-Earth-Veg GCMs showed a slightly less increase than the EC-Earth for all the scenarios. Under SSP-119, the R95pTOT value range from 5 mm in the northeastern region of JRB to  $-2$  mm in the southern region. Similarly, SSP-245 showed a small change of 1–2 mm over the entire basin, whereas SSP-370 showed an increase of 5 mm in the south and a decrease of  $-6$  mm in the northern parts. For the fossil fuel development scenarios (SSP-585), the R95pTOT showed an increase of 8 mm in the southern region and a decrease of  $-9$  mm in the northern part.

MRI-ESM2 showed a moderate increase in the near future for SSP-119, whereas the percentage change in the R95pTOT showed a reduction for SSP-370 and SSP-585. The rainfall showed an increase up to 5–9 mm in the sustainability scenario (SSP-119), whereas for the middle of the road scenario, the increase in R95pTOT was minimal (2 mm). Furthermore, the reduction of  $-2$  to  $-13$  mm in the R95pTOT was noticed during the regional rivalry (SSP-370) and fossil fuel development (SSP-585) scenarios.

The change in R95pTOT of the far future (2060–2099) compared to the base period is shown in Figure 12b. The maps show a gradual increase in R95pTOT from 5 mm for SSP-119 to 21 mm for SSP-585 for EC-Earth GCM. EC-Earth-Veg showed a similar pattern of increase in R95pTOT for SSP-119 and SSP-585. A decrease in R95pTOT was observed for SSP-199, whereas an increase by 5 and 21 mm was observed for SSP-245 and SSP-585, respectively. MRI-ESM2 showed a reverse pattern for R95pTOT under SSP-119, whereas a decrease in R95pTOT was projected for SSP-585. SSP-370 showed no changes in R95pTOT in the far future.

#### Total Rainfall above 99th Percentile (R99pTOT)

The percentage changes in R99pTOT are shown in Figure 13a. For SSP-119, the EC-Earth showed an increase of 3 mm in the north and 1 mm in the southern part. EC-Earth-Veg showed a similar increase, whereas MRI-ESM2 showed a higher increase up to 9 mm in the entire JRB for rainfall which exceeded the 99 percentiles of daily rain during the near future. For SSP-245, the EC-Earth showed an increase of 3 mm. Similarly, EC-Earth-Veg showed an increase of 3–5 mm, whereas the MRI-ESM2 showed an increase of 5–7 mm in the basin.



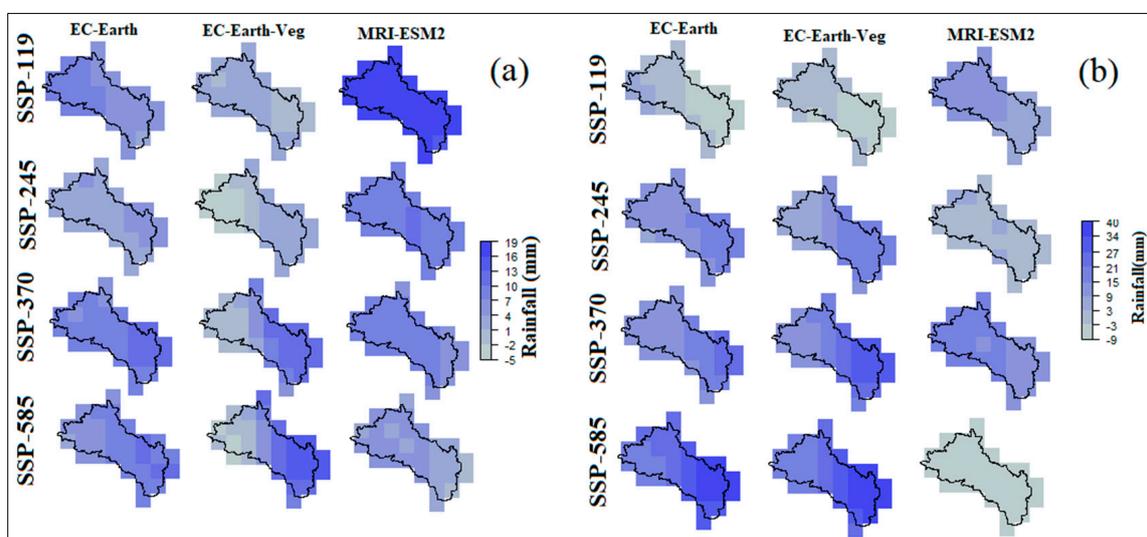
**Figure 13.** Change in annual total rainfall above 99 percentile (a) in the near future; (b) far future.

The EC-Earth and EC-Earth Veg showed a similar change for SSP-370 like the previous scenarios, but MRI-ESM2 showed a slight increase in the R99pTOT (up to 1–2 mm). EC-Earth and EC-Earth-Veg showed a positive change at the majority of the grid points for SSP-585, whereas MRI-ESM2 showed a slight decrease of -1 mm in most of the grids in JRB during 2020–2059.

The changes in R99pTOT for the far future period are shown in Figure 13b. For SSP-119, the highest increase of 7 mm was observed for MRI-ESM2, whereas EC-Earth-Veg and EC-Earth showed an increase of 3–5 mm. An average increase of 4–5 mm was observed for all the GCMs for SSP-245 and SSP-375, except MRI-ESM2, which showed a negative change of -1 to -2 mm at some grid points. Under fossil fuel development scenarios (SSP-585), the EC-Earth and EC-Earth-Veg showed an increase in R99pTOT up to 5 mm, whereas MRI-ESM2 showed a decrease by -1 to -2 mm.

#### Changes in One Day Max Rainfall ( $R \times 1$ day)

Figure 14a shows the changes in maximum one-day rainfall for 2020–2059. All GCMs showed an increase in  $R \times 1$  day at the majority of the grids for all the scenarios. For SSP-119, the least change of 1–4 mm was observed for EC-Earth-Veg, whereas EC-Earth showed an increase of up to 10 mm and MRI-ESM2 by 19 mm in the near future.



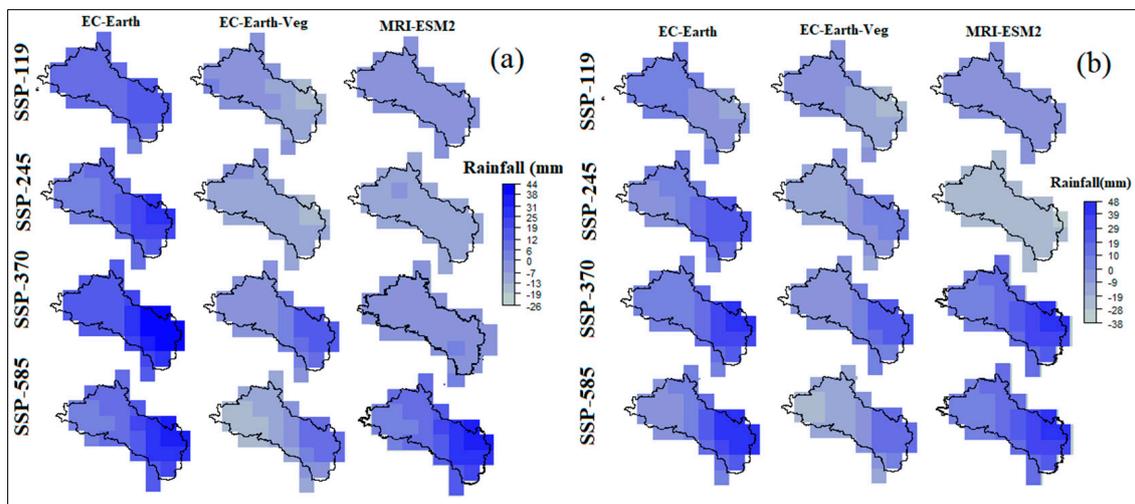
**Figure 14.** Change in one-day max rainfall (a) in the near future; (b) far future.

All GCMs projected a lower increase for SSP-245 than SSP-119. The maximum change of 10 mm followed by an average increase of 4–7 mm and 1–4 mm was observed for MRI-ESM2, EC-Earth, and EC-Earth-Veg, respectively. For SSP-370, the  $R \times 1$  day showed an increase by 4–7 mm for EC-Earth, whereas a slight decrease in the northern part by  $-2$  mm for EC-Earth-Veg and an overall increase of 7–10 mm for MRI-ESM2. For SSP-585, the  $R_{x1day}$  showed an increase of 4–7 mm for EC-Earth. EC-Earth-Veg showed a slight decrease in the northern region, whereas an increase of 4–7 mm in the southern part. MRI-ESM2 showed a slight decrease at the majority of the grid points with a value ranging from  $-2$  to  $-5$  mm.

The changes in the far future are shown in Figure 14b. For SSP-119, a decrease in  $R_{x1day}$  was observed for EC-Earth and EC-Earth-Veg by  $-3$  to  $-9$  mm. MRI-ESM2 showed an increase in the northern region up to 15 mm and an increase of 9 mm in the southern part. The change in  $R_{x1day}$  was in the range of 9 to 21 mm for EC-Earth and EC-Earth-Veg, whereas MRI-ESM2 showed an increase of 3 to 9 mm for SSP-245. SSP-370 showed a further increase of  $R_{x1day}$ , EC-Earth (15–21 mm), EC-Earth-Veg (15–27 mm), and MRI-ESM2 (15–21 mm).  $R_{x1day}$  was projected to increase by 34 to 40 mm for the fossil fuel development scenario. EC-Earth-Veg showed a similar increase of 21–34 mm, whereas MRI-ESM2 decreased to  $-3$  mm in the entire basin.

#### Changes in 5-Day Max Rainfall ( $R \times 5$ day)

Figure 15a shows the changes in  $R \times 5$  day for 2020–2059. EC-Earth showed the greatest changes of up to 44 mm for all SSPs, whereas EC-Earth-Veg showed a reduction. For SSP-119, the change ranges from 12 to 25 mm for EC-Earth,  $-7$  to 6 mm for EC-Earth-Veg, and 0 to 6 mm for MRI-ESM2.



**Figure 15.** Change in five-day max rainfall (a) in the near future; (b) far future.

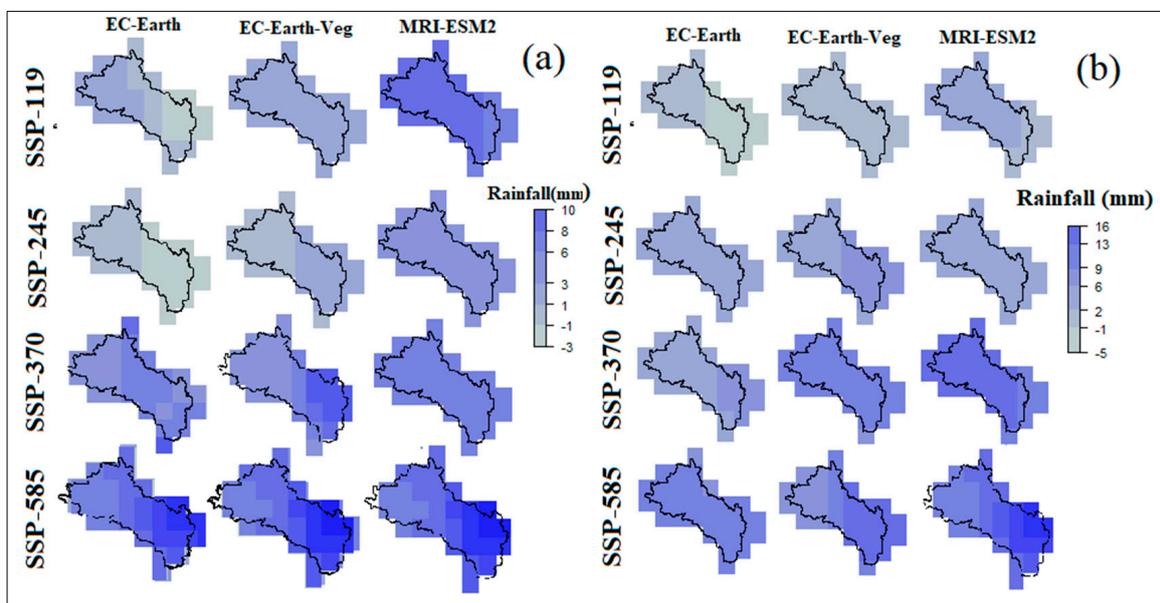
EC-Earth showed a  $-5$  to 12 mm change at different grids for SSP-245. EC-Earth showed a decrease by 0 to  $-7$  mm. MRI-ESM2 also showed a decrease like EC-Earth-Veg. For SSP-370, the southern region showed an increase of up to 44 mm for EC-Earth GCM, whereas MRI-ESM2 and EC-Earth-Veg showed an increase of 2 to 12 mm. MRI-ESM2 showed an increase up to 38 mm in the southern part of JRB for SSP-585. In contrast, EC-Earth-Veg showed a slight increase in the  $R \times 5$  day in the southern part and a decrease up to  $-3$  mm in the northern region. All grids under EC-Earth showed an increase in the  $R \times 5$  day index ranging from 12–25 mm in the near future.

The changes in  $R \times 5$  day for the far future are shown in Figure 15b. The plot shows an overall increase of up to a maximum of 10 mm under SSP-119 for all GCMs. There was a further increase for EC-Earth up to 19 mm, whereas a decrease up to  $-19$  mm in the northern region for EC-Earth-Veg and MRI-ESM2 for SSP-245. An overall increase in the  $R$

$\times 5$  day was observed for SSP-370. MRI-ESM2 showed the maximum increase of 39 mm, followed by EC-Earth (30 mm) and EC-Earth-Veg (19 mm). For SSP-585, EC-Earth  $R \times 5$  day showed an increase of 40 mm in the southern region and 10–19 mm in the northern region of JRB. EC-Earth-Veg showed a slight decrease in the  $R \times 5$  day in the northern part in the range of  $-9$  to  $-19$  mm. However, positive changes up to 10 mm were observed in the southern grids. MRI-ESM2 showed the highest increase of 48 mm in the JRB.

#### Changes in Rainfall Intensity (RI)

The changes in RI under various climate change scenarios for the near and far future are shown in Figure 16a, respectively. The Figure 16a shows the decrease in RI for SSP-119. EC-Earth projected a decrease by  $-1$  to  $-3$  mm, EC-Earth-Veg showed a small increase of up to 1 mm, and MRI-ESM2 showed an increase of 6 to 8 mm. The RI for SSP-245 showed no major variations for GCMs. It showed an increase for SSP-370 in the range of 3–8 mm for EC-Earth, 3–10 mm for EC-Earth-Veg and 6–8 mm for MRI-ESM2. The highest increase in rainfall intensity was for SSP-585. It was projected to increase up to 6–10 mm for EC-Earth, 3–10 mm for EC-Earth-Veg, and 8–10 mm for MRI-ESM2 over most parts of JRB.



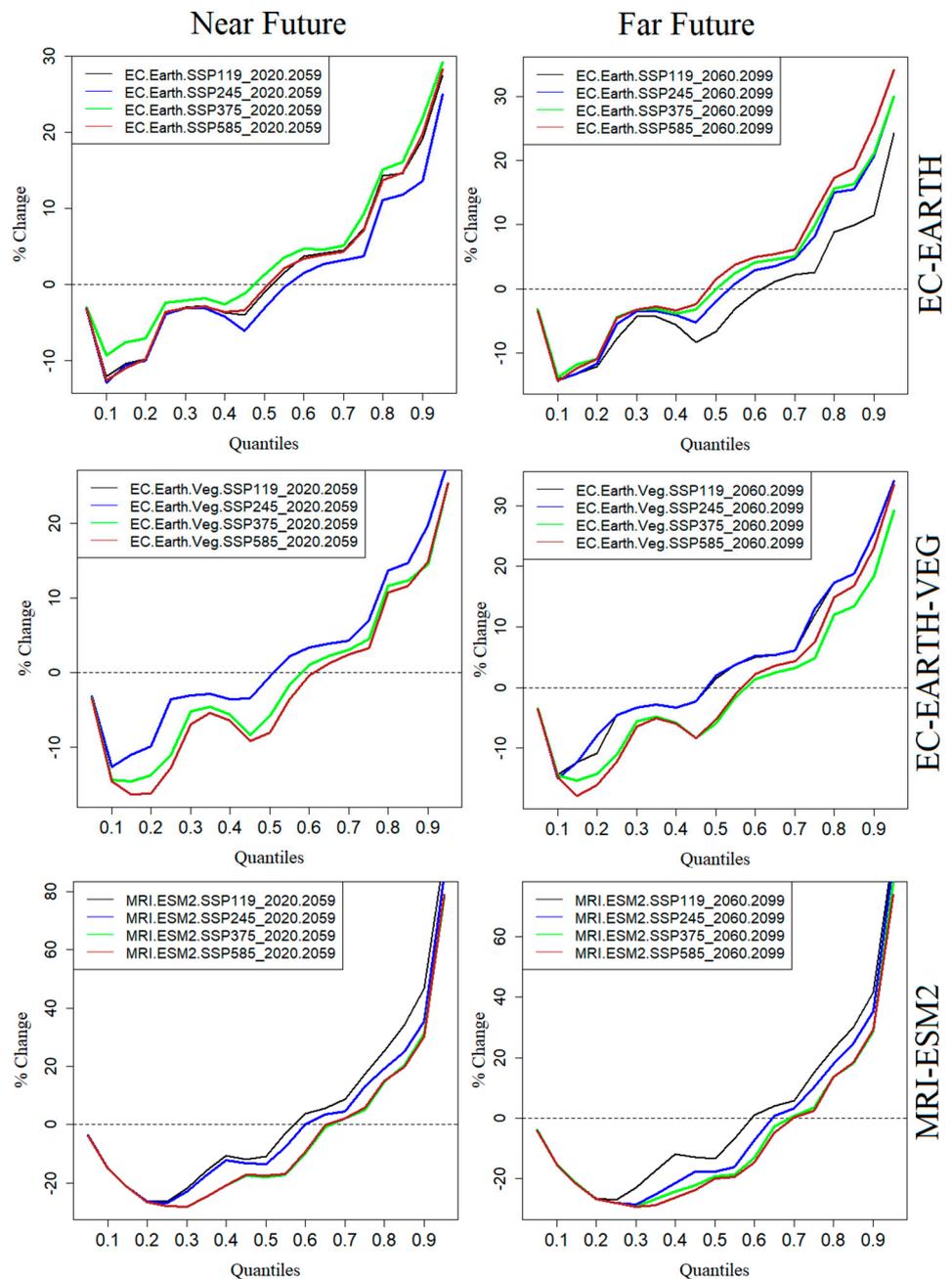
**Figure 16.** Change in rainfall intensity (a) in the near future; (b) far future.

Change in RI for the far future is shown in Figure 16b. The results show a decrease in RI for SSP-119 for all GCMs. However, the RI showed an increase all over JRB for SSP-370 and 585.

#### 4.3.2. Changes in River Flow

The distributed hydrological model developed using RF was used to simulate the historical and future river flows using the downscaled data of CMIP6 GCMs. The historical and future precipitation and temperature data of the most suitable GCMs (As discussed in our previous paper [49], EC-Earth, EC-Earth-Veg, and MRI-ESM2 were used as input for the model. The simulated flow for these projected data sets was analyzed for each GCM individually to cover the maximum uncertainty range in the near (2020–2059) and far (2060–2099) future. A comparison of quantiles for various SSPs for EC-Earth is shown in Figure 17. Changes for low quantile flow were observed to decrease in the near future. The maximum reduction ( $-14\%$ ) was for SSP-245, whereas a  $-8\%$  reduction was for 0.1 quantiles compared to the historical flow period. For mid and higher quantiles, the change in river flow was projected to increase up to 28%. The projected highest increase

was in the near future period for the SSP-375. However, the higher quantiles, such as 0.90, the minimum change were recorded for SSP-275 for the near future.



**Figure 17.** Changes in simulated flow for GCM EC-Earth.

The changes in quantiles for EC-Earth show that the river flow reduces in all scenarios for lower quantiles, while the maximum increase of 32% was observed for higher quantiles for SSP-585, followed by 22% for SSP-119.

The model simulation using EC-Earth-Veg showed a reduction in river flow for lower quantiles in the near future, as shown in Figure 17. However, an increase in river flow was noted for higher quantiles, indicating an increase in extreme flows in future periods. The maximum change of 25% was observed in higher streamflow quantiles in the near future for SSP-245, whereas the lowest changes in the higher extremes were for SSP-585. In the far future period, a similar flow pattern for lower quantiles and an increase in higher quantiles were recorded with a maximum increase of 32% for SSP-119 and 27% for SSP-245.

Figure 17 shows the changes in river flow for MRI-ESM2. The figure shows the greatest change in the higher quantiles for SSP-119 (80%) and the lowest change in lower quantiles ( $<0.3$ ) for SSP-585. The percentage reduction for SSP-585 was  $-23\%$  in the near future. The far future showed an increase in river flow extremes, whereas lower flows showed a reduction. The maximum change of  $68\%$  was observed for SSP-119 for higher quantiles in the far future. The lowest reduction in the flow was  $-24\%$  for SSP-585 in the near future.

## 5. Discussion

### 5.1. Reliability of the Newly Developed Model

Estimating river flow is an intricate process, especially in data scarce catchments. A huge set of data with a long temporal resolution is required for parameter estimation and optimization. The recent use of ML in hydrological modelling is gaining more attention in the scientific community. Integrated hydrological model development with the help of ML proved to be efficient as compared to the conventional modelling methods. However, there remains a gap for improvement by optimizing its internal parameters. Therefore, this study developed a distributed hydrological model using RF for parameter estimation. The calibration and validation results are provided in Section 4.2. The statistical indices used to show the efficiency of the model output shows the model's good capability to simulate the river flow in JRB. The model showed good performance for the calibration period, giving the NSE,  $d$ , KGE, RMSE, and Pbias of 0.96, 0.99, 0.92, 4.01, and  $-0.2$ , respectively. The NSE value of 0.96 is much better than the reported NSE values for the calibration period for similar other models, such as SWAT and HSPF, APEX and SAC-SMA [63]. Conventionally, the NSE value greater than 0.65 is considered good for model evaluating criteria [64]. The model showed satisfactory values for  $d$  during the calibration (0.99) and validation period (0.94). The simulation of the model train with machine learning algorithms showed very less bias in calibration and validation periods. Pbias revealed the overestimation or underestimation of the simulated flow compared to the measured flow. From the literature, the acceptable range of Pbias for model simulation is less than 10%, whereas in this case, the Pbias is  $-0.2\%$  for calibration and  $-7.2\%$  for the validation period. The error terms such as RMSE values range from 4.01 to 5.64 for calibration and validation. The KGE values are also in the acceptable range of 0.92 for calibration and 0.86 for validation. Tan et al. [45] validated the re-known SWAT model in JRB. The statistical results of the SWAT model for calibration and validation are NSE 0.66 and 0.62, respectively, whereas the model developed in this study showed NSE values of 0.96–0.75. The model calibration and validation result shows that this model can be used to simulate river flows in JRB for other datasets.

### 5.2. Changes in Precipitation Flood Frequency under Future Scenario

Variation in intensity and frequency of various climate and weather extremes has been found in the literature. The climate extremes are mostly found increasing in many parts of the world [65,66]. Over MSEA, substantial changes in daily rainfall intensity and the number of heavy rainfall days (R20 mm) have been observed for various future scenarios of CIMP6 [67]. Therefore, in this study, CMIP6 model outputs of three selected models were used for the historic forcing and future scenarios, such as SSP-119, SSP-245, SSP-370 and SSP-585, to study the changes in precipitation extremes in JRB. These data of CIMP6 GCMs were also used in the hydrological model to determine the change in flow extremes in JRB under various SSP scenarios. The results are shown in Section 4.3.2. Under different SSP scenarios, the flow at Ratu Panjang station was observed to be increasing for higher quantiles. However, the changes in lower quantiles compared to the historical flow were decreasing. Similarly, the precipitation extremes such as five-day max rainfall and rainfall intensity were observed to be increasing in the latter part of the century. The changes in these rainfall extremes and the substantial increase in some other extremes such as  $R \times 1$  day, R95pTOT, and R99pTOT justify the increasing river flow for higher quantiles than the historical flow at the river station.

Similar changes in climate extremes have also been reported by Kharin et al. [68]. They used the transient non-stationary GEV to study the global scale frequency change in climate extremes and risk ratio. The risk ratio determined in the study showed an increase from 0.65 to 1.22, while the global temperature increased from the preindustrial level under scenarios of CMIP5. Our study also strengthens the hypothesis that “*The contrast in relative frequency changes between more extreme and weaker events is projected to become larger as climate warms*”. Li et al. [69] analyzed 20 GCM from CMIP6 to study the change in temperature and precipitation extremes over the globe. The study found that most of the model increases the intensity and frequency of precipitation extremes, especially over tropical regions. The maximum one- and five-day rainfall events  $R \times 1$  day and  $R \times 5$  day increased up to 7.2% compared to the historical extremes. In the majority region of the world, the temperature and precipitation extremes were following the “intense gets intenser” tendency. However, comparable results of flow quantiles were observed in our study. The lower flow quantiles were found to decrease in most scenarios, whereas the higher quantiles were increasing for all models and scenarios. Therefore, it can be remarked that, in JRB, the precipitation and river flow extremes at Ratu Panjang gauge station will be increasing in the future.

### 5.3. Significance of the Study

Hydrological disasters like floods and water stress have become a common phenomenon in many countries globally. Consequently, a moderate dry spell often forces water rationing and moderate or extreme rainfall causing floods, especially in rapidly developing urban catchments [70]. The changing pattern of hydrological disasters due to environmental changes is a major concern for scientists and policymakers all over the globe. Numerous hydrological models have been developed to estimate runoff from rainfall to predict hydrological disasters [19]. Distributed hydrological models have been found to be most reliable for runoff prediction. However, they need extensive data and parameters in space and time for reliable runoff estimation. The outputs of such models are also prone to uncertainties due to the simple approximation of many hydrological processes. This study attempt to introduce a machine learning algorithm to improve the performance of the distributed hydrological model. The results showed that the hybridization of ML and conventional physical or conceptual model improved the capability to model complex interactions and runoff prediction. The model can provide a more accurate estimation of streamflow extremes. Therefore, the model can be used for reliable simulation of hydrological disasters, particularly water scarcity and floods, due to the changes in land use driven by physical and socio-economic factors and climate. This is particularly important for developing countries where rapid landuse changes have significantly affected local hydrology. This study also showed the suitability of the model in reliable projections of hydrological changes due to climate change. Therefore, the model’s output can be used for climate change adaptation and mitigation planning.

## 6. Conclusions

The distributed hydrological model for JRB was developed by using ML algorithms. RF was used to estimate the parameters to calculate the simulated flows. The model was developed using the bias-corrected IMERG data with an approximate resolution of 10 km. The soil properties and the topographical characteristics were included in calculating the model output. The model showed a varying flow simulation at Ratu Panjang compared to the observed flow. The efficiency of the model was assessed by calculating statistical indices. These indices values, such as RMSE, NSE, and  $R^2$ , proved that the distributed hydrological model can simulate the flow of any catchment. The calibration and validation results and the processing time prove that the ML-based models are good in flood simulation in any catchment with insufficient historical data. The model developed in this study can efficiently simulate the hydrological behaviour like the physical models, and also it can be applied to generate the long-term simulation. The model provided a near real-time

flood simulation using the bias-corrected IMERG data and used it to indicate the flood susceptibility of any region.

The study found that the river flow under the change climate scenarios increases with the higher carbon concentration pathways. The results also revealed that the rainfall extremes are also getting worse in intensity and frequency. The reduction of flow up to  $-14\%$  at lower quantiles and an increase of  $28\%$  at mid and higher quantiles were recorded in this analysis. Similarly, the sustainability pathway (SSP1) showed a reduction in projected river flow extremes, whereas the middle of the road (SSP2) showed a balance increase in the higher flow quantiles. Contrary to these, the regional rivalry (SSP3) and fossil fuel development (SSP5) showed a higher increase in streamflow extremes by up to  $68\%$  at the end of the century. The framework developed in this study can simulate the historical and future surface runoff very effectively with very few parameters. The model's efficacy is improved due to the use of RF in parameters estimation and GCM data, enabling it to simulate the effect of climate change on the river discharge in the region. The simulation takes less time, showing that the model can also be considered for NRT flood simulation in any region.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su14116620/s1>, Supplementary materials are available with this paper as appendices.

**Author Contributions:** Conceptualization, Z.I. and S.S.; Methodology, Z.I.; Software, Z.I. and S.S.; Validation, S.S.; Formal analysis, Z.I. and S.S.; Resources, T.I.; data curation, Z.S.; writing—original draft preparation, Z.I., Z.M.Y. and S.S.; writing—review and editing, A.F., T.I. and Z.M.Y.; funding acquisition, T.I. and A.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** Ministry of Higher Education Malaysia (FRGS) (No. R.J130000.78515F092).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The observed streamflow data is not available to be shared with a third party as per instruction from the Department of Irrigation and Drainage Malaysia. However, the GCM and Satellite data sets are freely available on the website/references given in the article.

**Acknowledgments:** The authors would like to acknowledge Higher Education Commission Pakistan (HEC) and Ministry of Higher Education Malaysia (FRGS) (No. R.J130000.78515F092) for providing financial support to conduct this research. We also acknowledge the Department of Irrigation and Drainage Malaysia for providing the rainfall data of entire Peninsular Malaysia.

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

## References

1. Yaseen, Z.M.; Sulaiman, S.O.; Deo, R.C.; Chau, K.-W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *J. Hydrol.* **2019**, *569*, 387–408. [[CrossRef](#)]
2. Vogel, R.M.; Lall, U.; Cai, X.; Rajagopalan, B.; Weiskel, P.K.; Hooper, R.P.; Matalas, N.C. Hydrology: The interdisciplinary science of water. *Water Resour. Res.* **2015**, *51*, 4409–4430. [[CrossRef](#)]
3. Halder, B.; Haghbin, M.; Farooque, A.A. An Assessment of Urban Expansion Impacts on Land Transformation of Rajpur-Sonarpur Municipality. *Knowl.-Based Eng. Sci.* **2021**, *2*, 34–53. [[CrossRef](#)]
4. Ahmad, S.; Simonovic, S.P. Spatial System Dynamics: New Approach for Simulation of Water Resources Systems. *J. Comput. Civ. Eng.* **2004**, *18*, 331–340. [[CrossRef](#)]
5. Sa'adi, Z.; Shiru, M.S.; Shahid, S.; Ismail, T. Selection of general circulation models for the projections of spatio-temporal changes in temperature of Borneo Island based on CMIP5. *Theor. Appl. Climatol.* **2020**, *139*, 351–371. [[CrossRef](#)]
6. Ehteram, M.; Othman, F.B.; Yaseen, Z.M.; Afan, H.A.; Allawi, M.F.; Malek, M.B.A.; Ahmed, A.N.; Shahid, S.; Singh, V.P.; El-Shafie, A. Improving the Muskingum flood routing method using a hybrid of particle swarm optimization and bat algorithm. *Water* **2018**, *10*, 807. [[CrossRef](#)]
7. Sharafati, A.; Khazaei, M.R.; Nashwan, M.S.; Al-Ansari, N.; Yaseen, Z.M.; Shahid, S. Assessing the uncertainty associated with flood features due to variability of rainfall and hydrological parameters. *Adv. Civ. Eng.* **2020**, *2020*, 7948902. [[CrossRef](#)]

8. Mango, L.M.; Melesse, A.M.; McClain, M.E.; Gann, D.; Setegn, S.G. Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: Results of a modeling study to support better resource management. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2245–2258. [[CrossRef](#)]
9. Halder, B.; Ameen, A.M.S.; Bandyopadhyay, J.; Khedher, K.M.; Yaseen, Z.M. The impact of climate change on land degradation along with shoreline migration in Ghoramara Island, India. *Phys. Chem. Earth Parts A/B/C* **2022**, 103135. [[CrossRef](#)]
10. Saudi, A.S.M.; Juahir, H.; Azid, A.; Azaman, F. Flood risk index assessment in Johor River Basin. *Malays. J. Anal. Sci.* **2015**, *19*, 991–1000.
11. Muzamil, S.A.H.B.S.; Zainun, N.Y.; Ajman, N.N.; Sulaiman, N.; Khahro, S.H.; Rohani, M.M.; Mohd, S.M.B.; Ahmad, H. Proposed Framework for the Flood Disaster Management Cycle in Malaysia. *Sustainability* **2022**, *14*, 4088. [[CrossRef](#)]
12. Shahid, S.; Alamgir, M.; Wang, X.; Eslamian, S. Climate Change Impacts on and Adaptation to Groundwater. *Handb. Drought Water Scarcity Environ. Impacts Anal. Drought Water Scarcity* **2017**, *2*, 107–124.
13. Ziarh, G.F.; Asaduzzaman, M.; Dewan, A.; Nashwan, M.S.; Shahid, S. Integration of catastrophe and entropy theories for flood risk mapping in peninsular Malaysia. *J. Flood Risk Manag.* **2021**, *14*, e12686. [[CrossRef](#)]
14. Connor, R. *The United Nations World Water Development Report 2015: Water for a Sustainable World*; UNESCO Publishing: Bonn, Germany, 2015; Volume 1, ISBN 9231000713.
15. Chemicals, U. *Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases*; United Nations Environment Programme: Geneva, Switzerland, 2003; Volume 194.
16. Iqbal, Z.; Shahid, S.; Ahmed, K.; Ismail, T.; Nawaz, N. Spatial distribution of the trends in precipitation and precipitation extremes in the sub-Himalayan region of Pakistan. *Theor. Appl. Climatol.* **2019**, *137*, 2755–2769. [[CrossRef](#)]
17. Ahmad, S.; Simonovic, S.P. System Dynamics Modeling of Reservoir Operations for Flood Management. *J. Comput. Civ. Eng.* **2000**, *14*, 190–198. [[CrossRef](#)]
18. Sitterson, J.; Sinnathamby, S.; Parmar, R.; Koblich, J.; Wolfe, K.; Knightes, C.D. Demonstration of an online web services tool incorporating automatic retrieval and comparison of precipitation data. *Environ. Model. Softw.* **2020**, *123*, 104570. [[CrossRef](#)]
19. Young, P.C. Advances in real-time flood forecasting. *Philos. Trans. R. Soc. London. Ser. A Math. Phys. Eng. Sci.* **2002**, *360*, 1433–1450. [[CrossRef](#)]
20. Fahimi, F.; Yaseen, Z.M.; El-shafie, A. Application of soft computing based hybrid models in hydrological variables modeling: A comprehensive review. *Theor. Appl. Climatol.* **2017**, *128*, 875–903. [[CrossRef](#)]
21. Naganna, S.R.; Beyaztas, B.H.; Bokde, N.; Armanuos, A.M. On the evaluation of the gradient tree boosting model for groundwater level forecasting. *Knowl.-Based Eng. Sci.* **2020**, *1*, 48–57. [[CrossRef](#)]
22. Devia, G.K.; Ganasri, B.P.; Dwarakish, G.S. A Review on Hydrological Models. *Aquat. Procedia* **2015**, *4*, 1001–1007. [[CrossRef](#)]
23. Perrin, C.; Michel, C.; Andréassian, V. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol.* **2001**, *242*, 275–301. [[CrossRef](#)]
24. Agrawal, N.; Desmukh, T.S. Rainfall Runoff Modeling using MIKE 11 Nam—A Review. *Int. J. Innov. Sci. Eng. Technol.* **2016**, *3*, 659–667.
25. Yaseen, Z.M.; Ebtehaj, I.; Kim, S.; Sanikhani, H.; Asadi, H.; Ghareb, M.I.; Bonakdari, H.; Wan Mohtar, W.H.M.; Al-Ansari, N.; Shahid, S. Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis. *Water* **2019**, *11*, 502. [[CrossRef](#)]
26. Khosravi, K.; Golkarian, A.; Booi, M.J.; Barzegar, R.; Sun, W.; Yaseen, Z.M.; Mosavi, A. Improving daily stochastic streamflow prediction: Comparison of novel hybrid data-mining algorithms. *Hydrol. Sci. J.* **2021**, *66*, 1457–1474. [[CrossRef](#)]
27. Johari, A.; Javadi, A.A.; Habibagahi, G. Modelling the mechanical behaviour of unsaturated soils using a genetic algorithm-based neural network. *Comput. Geotech.* **2011**, *38*, 2–13. [[CrossRef](#)]
28. Omeje, O.E.; Maccido, H.S.; Badamasi, Y.A.; Abba, S.I. Performance of Hybrid Neuro-Fuzzy Model for Solar Radiation Simulation at Abuja, Nigeria: A Correlation Based Input Selection Technique. *Knowl.-Based Eng. Sci.* **2021**, *2*, 54–66.
29. Khan, N.; Shahid, S.; Juneng, L.; Ahmed, K.; Ismail, T.; Nawaz, N. Prediction of heat waves in Pakistan using quantile regression forests. *Atmos. Res.* **2019**, *221*, 1–11. [[CrossRef](#)]
30. Wang, X.; Zhang, J.; He, R.; Amgad, E.; Sondoss, E.; Shang, M. A strategy to deal with water crisis under climate change for mainstream in the middle reaches of Yellow River. *Mitig. Adapt. Strateg. Glob. Chang.* **2010**, *16*, 555–566. [[CrossRef](#)]
31. Qin, H.-P.; Su, Q.; Khu, S.-T. An integrated model for water management in a rapidly urbanizing catchment. *Environ. Model. Softw.* **2011**, *26*, 1502–1514. [[CrossRef](#)]
32. Tidwell, V.C.; Passell, H.D.; Conrad, S.H.; Thomas, R.P. System dynamics modeling for community-based water planning: Application to the Middle Rio Grande. *Aquat. Sci.* **2004**, *66*, 357–372. [[CrossRef](#)]
33. Roper, R.F.; Rumí, R.; Aguilera, P.A. Modelling uncertainty in social–natural interactions. *Environ. Model. Softw.* **2016**, *75*, 362–372. [[CrossRef](#)]
34. Yaseen, Z.M.; Shahid, S. Drought Index Prediction Using Data Intelligent Analytic Models: A Review. In *Intelligent Data Analytics for Decision-Support Systems in Hazard Mitigation*; Springer: Cham, Switzerland, 2020; pp. 1–27.
35. Chandwani, V.; Vyas, S.K.; Agrawal, V.; Sharma, G. Soft computing approach for rainfall-runoff modelling: A review. *Aquat. Procedia* **2015**, *4*, 1054–1061. [[CrossRef](#)]
36. Koch, J.; Demirel, M.C.; Stisen, S. The SPATial Efficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models. *Geosci. Model Dev.* **2018**, *11*, 1873–1886. [[CrossRef](#)]

37. Guo, H.C.; Liu, L.; Huang, G.H.; Fuller, G.A.; Zou, R.; Yin, Y.Y. A system dynamics approach for regional environmental planning and management: A study for the Lake Erhai Basin. *J. Environ. Manag.* **2001**, *61*, 93–111. [\[CrossRef\]](#)
38. Sood, A. Integrated Watershed Management as an Effective Tool for Sustainable Development: Using Distributed Hydrological Models in Policy Making. Ph.D. Thesis, University of Delaware, Newark, NJ, USA, 2009.
39. Minville, M.; Cartier, D.; Guay, C.; Leclaire, L.-A.; Audet, C.; Le Digabel, S.; Merleau, J. Improving process representation in conceptual hydrological model calibration using climate simulations. *Water Resour. Res.* **2014**, *50*, 5044–5073. [\[CrossRef\]](#)
40. Bárdossy, A.; Singh, S.K. Robust estimation of hydrological model parameters. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 1273–1283. [\[CrossRef\]](#)
41. Zhang, L.; Nan, Z.; Xu, Y.; Li, S. Hydrological Impacts of Land Use Change and Climate Variability in the Headwater Region of the Heihe River Basin, Northwest China. *PLoS ONE* **2016**, *11*, e0158394. [\[CrossRef\]](#)
42. Zomorodian, M.; Lai, S.H.; Homayounfar, M.; Ibrahim, S.; Fatemi, S.E.; El-Shafie, A. The state-of-the-art system dynamics application in integrated water resources modeling. *J. Environ. Manag.* **2018**, *227*, 294–304. [\[CrossRef\]](#)
43. Ratnayeke, S.; van Manen, F.T.; Clements, G.R.; Kulaimi, N.A.M.; Sharp, S.P. Carnivore hotspots in Peninsular Malaysia and their landscape attributes. *PLoS ONE* **2018**, *13*, e0194217. [\[CrossRef\]](#)
44. Kia, M.B.; Pirasteh, S.; Pradhan, B.; Mahmud, A.R.; Sulaiman, W.N.A.; Moradi, A. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ. Earth Sci.* **2012**, *67*, 251–264. [\[CrossRef\]](#)
45. Tan, M.L.; Ficklin, D.L.; Ibrahim, A.L.; Yusop, Z. Impacts and uncertainties of climate change on streamflow of the Johor River Basin, Malaysia using a CMIP5 General Circulation Model ensemble. *J. Water Clim. Chang.* **2014**, *5*, 676–695. [\[CrossRef\]](#)
46. Webster, P.J.; Magaña, V.O.; Palmer, T.N.; Shukla, J.; Tomas, R.A.; Yanai, M.; Yasunari, T. Monsoons: Processes, predictability, and the prospects for prediction. *J. Geophys. Res. Ocean.* **1998**, *103*, 14451–14510. [\[CrossRef\]](#)
47. Noor, M.; Ismail, T.B.; Shahid, S.; Ahmed, K.; Chung, E.-S.; Nawaz, N. Selection of CMIP5 multi-model ensemble for the projection of spatial and temporal variability of rainfall in peninsular Malaysia. *Theor. Appl. Climatol.* **2019**, *138*, 999–1012. [\[CrossRef\]](#)
48. Zhang, W.; Villarini, G.; Scoccimarro, E.; Napolitano, F. Examining the precipitation associated with medicanes in the high-resolution ERA-5 reanalysis data. *Int. J. Climatol.* **2020**, *41*, E126–E132. [\[CrossRef\]](#)
49. Iqbal, Z.; Shahid, S.; Ahmed, K.; Ismail, T.; Ziarh, G.F.; Chung, E.-S.; Wang, X. Evaluation of CMIP6 GCM rainfall in mainland Southeast Asia. *Atmos. Res.* **2021**, *254*, 105525. [\[CrossRef\]](#)
50. Iqbal, Z.; Shahid, S.; Ahmed, K.; Wang, X.; Ismail, T.; Gabriel, H.F. Bias correction method of high-resolution satellite-based precipitation product for Peninsular Malaysia. *Theor. Appl. Climatol.* **2022**, *148*, 1429–1446. [\[CrossRef\]](#)
51. Huang, M.; Lin, R.; Huang, S.; Xing, T. A novel approach for precipitation forecast via improved K-nearest neighbor algorithm. *Adv. Eng. Inform.* **2017**, *33*, 89–95. [\[CrossRef\]](#)
52. Maraun, D.; Wetterhall, F.; Ireson, A.M.; Chandler, R.E.; Kendon, E.J.; Widmann, M.; Brienen, S.; Rust, H.W.; Sauter, T.; Themel, M.; et al. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* **2010**, *48*, 2009RG000314. [\[CrossRef\]](#)
53. Eden, J.M.; Widmann, M. Downscaling of GCM-Simulated Precipitation Using Model Output Statistics. *J. Clim.* **2014**, *27*, 312–324. [\[CrossRef\]](#)
54. Piani, C.; Haerter, J.O.; Coppola, E. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* **2010**, *99*, 187–192. [\[CrossRef\]](#)
55. Wilcke, R.A.I.; Mendlik, T.; Gobiet, A. Multi-variable error correction of regional climate models. *Clim. Change* **2013**, *120*, 871–887. [\[CrossRef\]](#)
56. Amengual, A.; Homar, V.; Romero, R.; Alonso, S.; Ramis, C. A statistical adjustment of regional climate model outputs to local scales: Application to Platja de Palma, Spain. *J. Clim.* **2012**, *25*, 939–957. [\[CrossRef\]](#)
57. Leander, R.; Buishand, T.A. Resampling of regional climate model output for the simulation of extreme river flows. *J. Hydrol.* **2007**, *332*, 487–496. [\[CrossRef\]](#)
58. Terink, W.; Hurkmans, R.; Torfs, P.; Uijlenhoet, R. Bias correction of temperature and precipitation data for regional climate model application to the Rhine basin. *Hydrol. Earth Syst. Sci. Discuss.* **2009**, *6*, 5377–5413.
59. Coles, S.; Bawa, J.; Trenner, L.; Dorazio, P. *An Introduction to Statistical Modeling of Extreme Values*; Springer: Cham, Switzerland, 2001; Volume 208.
60. Lenderink, G.; Buishand, A.; Van Deursen, W. Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1145–1159. [\[CrossRef\]](#)
61. Lafon, T.; Dadson, S.; Buys, G.; Prudhomme, C. Bias correction of daily precipitation simulated by a regional climate model: A comparison of methods. *Int. J. Climatol.* **2013**, *33*, 1367–1381. [\[CrossRef\]](#)
62. Allen, R.G.; Pruitt, W.O. FAO-24 Reference Evapotranspiration Factors. *J. Irrig. Drain. Eng.* **1991**, *117*, 758–773. [\[CrossRef\]](#)
63. Moriasi, D.N.; Arnold, J.G.; Liew, M.W.V.; Bingner, R.L.; Harmel, R.D.; Veith, T.L.; Van Liew, M.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [\[CrossRef\]](#)
64. Jeong, J.; Kannan, N.; Arnold, J.; Glick, R.; Gosselink, L.; Srinivasan, R. Development and Integration of Sub-hourly Rainfall–Runoff Modeling Capability Within a Watershed Model. *Water Resour. Manag.* **2010**, *24*, 4505–4527. [\[CrossRef\]](#)
65. Ahmed, K.; Shahid, S.; Chung, E.S.; Ismail, T.; Wang, X.J. Spatial distribution of secular trends in annual and seasonal precipitation over Pakistan. *Clim. Res.* **2017**, *74*, 95–107. [\[CrossRef\]](#)

66. Nashwan, M.S.; Shahid, S. Spatial distribution of unidirectional trends in climate and weather extremes in Nile river basin. *Theor. Appl. Climatol.* **2019**, *137*, 1181–1199. [[CrossRef](#)]
67. Ge, F.; Zhu, S.; Luo, H.; Zhi, X.; Wang, H. Future changes in precipitation extremes over Southeast Asia: Insights from CMIP6 multi-model ensemble. *Environ. Res. Lett.* **2021**, *16*, 24013. [[CrossRef](#)]
68. Kharin, V.V.; Flato, G.M.; Zhang, X.; Gillett, N.P.; Zwiers, F.; Anderson, K.J. Risks from Climate Extremes Change Differently from 1.5 °C to 2.0 °C Depending on Rarity. *Earth's Future* **2018**, *6*, 704–715. [[CrossRef](#)]
69. Li, C.; Zwiers, F.; Zhang, X.; Li, G.; Sun, Y.; Wehner, M. Changes in Annual Extremes of Daily Temperature and Precipitation in CMIP6 Models. *J. Clim.* **2021**, *34*, 3441–3460. [[CrossRef](#)]
70. Pereira, L.S.; Cordery, I.; Iacovides, I. *Coping with Water Scarcity: Addressing the Challenges*; Springer Science & Business Media: New York, NY, USA, 2009; ISBN 978-1-4020-9578-8.