

Article

Hydrologic Model Evaluation and Assessment of Projected Climate Change Impacts Using Bias-Corrected Stream Flows

Joseph A. Daraio 

Department of Civil Engineering, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada; jadarai@mun.ca; Tel.: +1-709-864-2756

Received: 18 July 2020; Accepted: 14 August 2020; Published: 18 August 2020



Abstract: Hydrologic models driven by downscaled meteorologic data from general circulation models (GCM) should be evaluated using long-term simulations over a historical period. However, simulations driven by GCM data cannot be directly evaluated using observed flows, and the confidence in the results can be relatively low. The objectives of this paper were to bias correct simulated stream flows from calibrated hydrologic models for two basins in New Jersey, USA, and evaluate model performance in comparison to uncorrected simulations. Then, we used stream flow bias correction and flow duration curves (FDCs) to evaluate and assess simulations driven by statistically downscaled GCMs for the historical period and the future time slices 2041–2070 and 2071–2099. Bias correction of stream flow from simulations increased confidence in the performance of two previously calibrated hydrologic models. Results indicated there was no difference in projected FDCs for uncorrected and bias-corrected flows in one basin, while this was not the case in the second basin. This result provided greater confidence in projected stream flow changes in the former basin and implied more uncertainty in projected stream flows in the latter. Applications in water resources can use the methods described to evaluate the performance of GCM-driven simulations and assess the potential impacts of climate change with an appropriate level of confidence in the model results.

Keywords: flow duration curves; bias correction of stream flows; flow regime; precipitation-runoff modeling system (PRMS)

1. Introduction

Estimation and simulation of the flow regime and hydrologic indices require longer time periods, especially for annual time series, which can ensure natural climate variability is included in model calibration and evaluation [1,2]. Calibration of hydrologic models for long-term simulations is done with objective functions on observed daily mean flows, monthly mean flow, and/or peak flows across multiple time resolutions [3,4]. In a non-stationary context, there are limitations to applying calibrated models [5]. Hydrologic simulations for climate change impact analysis are driven by downscaled general circulation models (GCM) [6–8], and such simulations cannot be directly calibrated with observed flows since the inputs differ for historical simulations. Additionally, there are no future observed data; therefore, it is not possible to calibrate or evaluate the model performance of hydrologic simulations for potential future climates. The absence of such data creates a situation similar to the development of models for ungauged basins [9–11], where regionally applicable measures of stream flows are required for model calibration and evaluation. Some questions remain as to the level of confidence for applications of hydrologic models driven by GCMs to local scales [12]. A potential solution to this problem is to use flow duration curves (FDC) and bias correction of stream flows for the evaluation of hydrologic models driven by statistically downscaled meteorological data.

FDCs have been used as objective functions for model calibration [1,13,14] and have a wide range of uses in water resource applications because they are able to simply convey relatively complex information about the hydrology of a system to decision makers [11,15]. Westerberg et al. [1] suggested using FDCs for calibration when the range of observed flows is outside the range of model input data and that important aspects of the flow regime (“hydrological signatures”) and simulated uncertainty have a more direct interpretation than other efficiency measures. A potential additional step to evaluate model performance involves bias correction of simulated stream flows with observed stream flows using FDCs. This is distinct from bias correction techniques that are required for both statistically and dynamically downscaled meteorological data from general circulation models (GCM) used to drive climate change impact assessment in hydrology [16,17], which is not discussed in this paper.

Daraio [18] used the precipitation-runoff modeling system (PRMS), a distributed hydrologic model [19] that has been widely used for climate change impact assessment [20–22], to simulate the hydrologic response of two watersheds on the Coastal Plain in New Jersey (NJ) under selected climate change scenarios. The region is expected to see increases in temperature of around 4 °C and a 6–10% increase in precipitation for high emission scenarios [23]. Model performance was deemed satisfactory, and simulations were able to provide some information on potential seasonal changes in flow and groundwater recharge. However, the model output was biased toward overestimation of low flows and underestimation of high flows. Given that this indicated a systematic bias in simulations, it was conjectured that bias correction of flow output using FDCs could be used to both more fully evaluate model performance and improve the simulation results. Additionally, it was recognized that a more thorough assessment of the models’ abilities to simulate long-term flow regimes was required to increase confidence in the results of climate change projections of stream flows in these basins [24].

Stream flows are well represented by a lognormal distribution, and the FDC is equivalent to the empirical cumulative distribution function for flows, which makes it ideal for parametric bias correction techniques. Bias correction and other correction methods have been applied to hydrologic model output of flows for hydrologic forecasting with success [25–28]. Bias-corrected flows increased the accuracy of ensemble forecasts [29] and improved forecasts using a coupled GCM [30]. Bias correction of flows has the advantage of being applied after parameter estimation and could help address the problem of equifinality [31]. Bias correction of simulated flows from hydrologic models could be used to both more fully evaluate model performance and improve simulation results, in particular for applications to ungauged basins and for future projected flows in a single basin.

The FDC is a representation of the hydrologic regime that is a function of both climatic drivers and watershed properties. Assuming watershed properties are well represented, simulations using climatic drivers from statistically downscaled GCMs should capture the overall long-term hydrologic regime, represented by the FDC. It is hypothesized that bias correction of stream flows and hydrologic analysis using FDCs can be used to evaluate and assess confidence in hydrologic simulations driven by statistically downscaled GCMs. The objectives of this paper were to use FDCs (1) to bias correct simulated stream flows from calibrated PRMS models for two basins in NJ and evaluate model performance in comparison to uncorrected simulations, (2) to use bias correction and FDCs to evaluate the performance of rainfall-runoff simulations driven by statistically downscaled GCMs over the historical period 1956–2005, and (3) to assess potential changes to FDCs using uncorrected and bias-corrected stream flows from climate change projections of the future time slices 2041–2070 and 2071–2099.

2. Methods

2.1. Overview

Previously calibrated PRMS models [18] were used to simulate stream flows in two watersheds in New Jersey, USA, for the historical period (1956–2005) and for the future period 2041–2099. The simulated stream flows were used in the following manner. (1) Simulated stream flows driven

by observed meteorological data from these models were bias corrected using FDCs developed from observed stream flows in each basin. (2) The PRMS simulations driven by statistically downscaled GCM meteorological data over the historical period 1956–2005 for both uncorrected and bias-corrected stream flows were evaluated using FDCs from observed and simulated stream flows. (3) Uncorrected and bias-corrected simulated stream flows driven by statistically downscaled GCM data for climate change projections of the future time slices 2041–2070 and 2071–2099 were assessed using FDCs.

Model performances for simulations driven by observed data for the historical period of water years 1956–2005 were assessed using several goodness-of-fit (GoF) measures and compared with GoF after bias correction of simulated stream flows. Meteorological data (bias corrected with constructed analogs, BCCA, maximum and minimum daily temperature and daily precipitation) from 15 different GCMs from the CMIP5 multi-model ensemble (Table 1) were used to drive calibrated PRMS models for the two watersheds. The GCM-driven stream flow simulations were evaluated and bias corrected for model error, and results from these simulations were used in the analyses described below.

Table 1. List of models from the CMIP5 multi-model ensemble used in this paper.

Modeling Center (or Group)	Institute ID	Model Name
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	CSIRO-BOM	ACCESS1.0
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1
College of Global Change and Earth System Science Beijing Normal University	GCESS	BNU-ESM
Canadian Centre for Climate Modelling and	CCCMA	CanESM2
National Center for Atmospheric Research	NCAR	CCSM4
Community Earth System Model Contributors	NSF-DOE-NCAR	CESM1(BGC)
Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5
Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3.6.0
Institute for Numerical Mathematics	INM	INM-CM4
Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR IPSL-CM5A-MR
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	MIROC-ESM MIROC-ESM-CHEM
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC5
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)	MPI-M	MPI-ESM-MR MPI-ESM-LR
Meteorological Research Institute	MRI	MRI-CGCM3
Norwegian Climate Centre	NCC	NorESM1-M

2.2. Site

The Batsto and the upper Maurice basins are located on the coastal plain of NJ, which ranges in elevation from 119 m to sea level and has a surficial geology that consists of unconsolidated to semi-consolidated material. The Batsto watershed is located entirely within the Pinelands region, which extends from central NJ to just north of Delaware Bay (Figure 1), and the majority of the basin is within the Pinelands National Reserve. The area primarily consists of sandy, acidic soil that sits atop the Kirkwood-Cohansey Aquifer. Coarse sands within the Pinelands are porous in nature, which allows for rapid infiltration. The upper Maurice watershed is adjacent to this region and includes a small part of the western border of the Pinelands, and the soils in the upper Maurice basin are similar to those in Batsto. Soils in the Batsto basin are 85% sand, and soils in the Maurice basin are 78% sand. The Maurice River becomes an estuary downstream of Union Lake, in Millville, NJ, and the extent of the area of the upper Maurice begins upstream of the entrance to the lake. The Batsto watershed is dominated by forest (60%) and wetlands (25%), and the upper Maurice watershed is mixed urban (28%), forest (30%), and agriculture (22%) lands. These two watersheds are near enough to share a very similar climatic regime with important differences in land use and degree of urbanization.

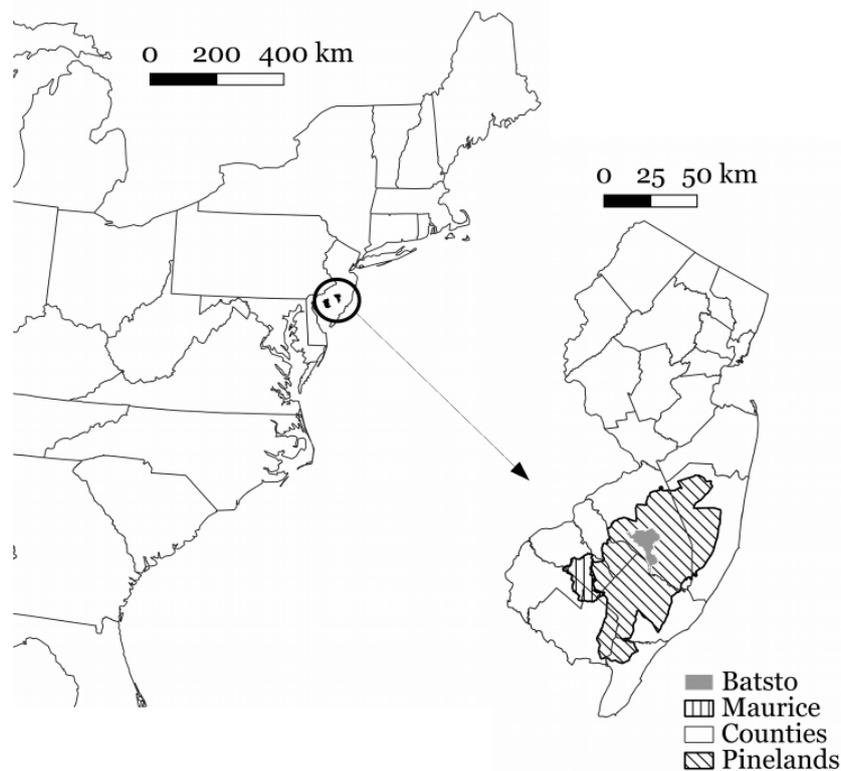


Figure 1. Locations of the Upper Maurice River watershed and the Batsto River watershed. Outlets of each watershed are shown as triangles. The gauge at Norma, NJ, was used for the Maurice River (USGS Site ID 01411500), and the gauge at Batsto, NJ, was used for the Batsto River (Site ID 01409500).

2.3. Hydrologic Simulations and Bias Correction

Simulations for the historical period were run from water years 1956–2005 using observed meteorologic data and meteorologic data derived from 15 GCMs (Table 1). Simulations were run using GCM derived data for the time slices 2041–2070 and 2071–2099.

2.3.1. Simulations Using Observed Data

The PRMS models were previously calibrated over the years 1989–1995 and validated over 1996–2003 by Daraio [18]. In this study, simulations were evaluated against observed flows for the historical period of water years 1956–2005 using several GoF measures (Table 2). Uncorrected simulated flows were bias corrected using quantile mapping based on the full record of flows. Bias correction was done using the flow duration curves from the entire period (1956–2005) instead of using part of the period for calibration and part for evaluation. This choice was made because a significant part of any bias is random, and a long time period is recommended for calibration of the bias correction [17].

Lognormal distributions were fit to FDCs, and parameter estimates were obtained, μ and σ , for all simulations using maximum likelihood estimation (MLE) with the R package “fitdistrplus” [32]. FDCs for observed and simulated flows were developed, and PRMS model bias was estimated and corrected in each basin using quantile mapping in R [33] with the R package “qmap” [34,35]. The general approach (see Gudmundsson et al. [34] for details) transforms the simulated data, P_m , using the following.

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

where P_o is the observed data, F_o^{-1} is the inverse CDF for P_o , and F_m is the CDF for P_m . The CDF was the quantile function estimated from the data that used a quantile step of 0.01. Bias correction was applied for daily mean flows over the full period 1956–2005 based on the composite FDC. As with uncorrected simulations, performance for bias-corrected simulations was evaluated using GoF measures.

Confidence intervals for FDCs were estimated at the 10% and 90% exceedance levels. The 10% and 90% FDCs and a median FDC were estimated by sorting each annual flow in decreasing order, high flows to low flows, then taking the 10%, 50%, and 90% quantiles for each associated probability of exceedance $p_{e,i}$, where:

$$p_{e,i} = \frac{i}{366} \text{ for } i = 1, 2, \dots, 365 \quad (2)$$

from the annual maximum flow to the lowest annual flow. The 10% FDC represents the FDC for flows that were in the 90th percentile (greater than 90% of all values) of flows for each $p_{e,i}$. Likewise, the 90% FDC represents the FDC for flows that were in the 10th percentile of flows for each $p_{e,i}$.

2.3.2. Goodness-of-Fit

Goodness-of-fit for simulations driven by observed historical climate data were evaluated using the R package “hydroGoF” [36] (Table 2). Note that the seasonal definition for this analysis does not follow the traditional definitions that were used in the GoF evaluations. The season definition was based on the exceedance hydrograph for observed flows (Figure 2 and Daraio [18]): winter = JFM and part of April, spring = AMJJ, summer = JASO, and autumn = OND. Seasonal definitions for the GoF measures were winter = DJF, spring = MAM, summer = JJA, and autumn = SON. Therefore, the GoF measures for seasonal flows provided a robust measure of model performance at these time scales.

Table 2. Goodness-of-fit measures.

ME	Mean Error	$\frac{1}{N} \sum (Q_s - Q_o)$
R ²	Coefficient of variation	$\left[\frac{1}{N} \frac{(Q_o - \bar{Q}_o)(Q_s - \bar{Q}_s)}{\sigma_o \sigma_s} \right]^2$
RMSE	Root Mean Squared Error	$\sqrt{\frac{1}{N} \sum (Q_s - Q_o)^2}$
PBIAS	Percent Bias	$\frac{\sum (Q_s - Q_o)}{\sum Q_o} \times 100$
NSE	Nash–Sutcliffe Efficiency	$1 - \frac{\sum (Q_o - Q_s)^2}{\sum (Q_o - \mu_o)^2}, -\infty \leq NSE \leq 1$
d	Index of agreement	$1 - \frac{\sum (Q_o - Q_s)^2}{\sum [(Q_s - \mu_o) + (Q_o - \mu_o)]^2}, 0 \leq d \leq 1$
KGE	Kling–Gupta Efficiency	$1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$ $\beta = \frac{\mu_s}{\mu_o}; \gamma = \frac{\sigma_s / \mu_s}{\sigma_o / \mu_o}, -\infty \leq KGE \leq 1$
VE	Volumetric efficiency	$1 - \frac{\sum (Q_o - Q_s) }{\sum Q_o}, 0 \leq VE \leq 1$

Q_o = observed flow (m³/s); Q_s = simulated flow (m³/s); N = total number of observations/simulations; μ_o = mean observed flow; μ_s = mean simulated flow; σ_o = standard deviation of observed flow; σ_s = standard deviation of simulated flow; summations are done over daily mean flow for the indicated time period.

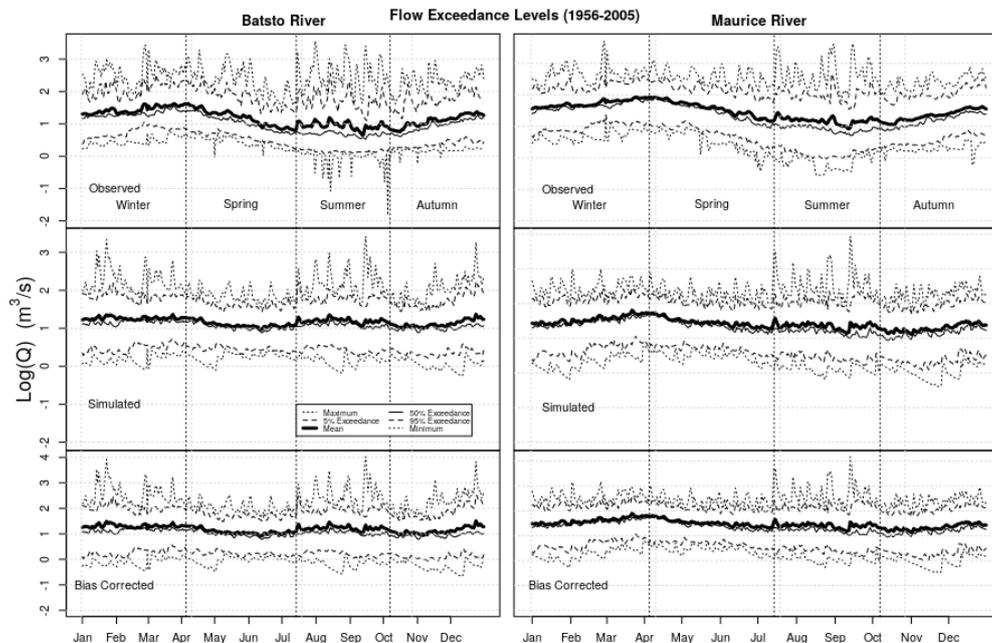


Figure 2. Flow exceedance for water years 1956–2005 for the Batsto (left) and Maurice Rivers. Plots show from top to bottom: observed, simulated uncorrected (UC), and bias-corrected (BC) for each river. The lower curve is the minimum flow in the record; the next curve is the 95% exceedance level (flow exceeded 95% of the time); then the median flow, then mean flow (thick line), then 5% exceedance, then the maximum flow in the record for each day of the year.

2.4. Simulations Using GCM Data

Statistically downscaled bias-corrected (1/8° × 1/8°) constructed analogs (BCCA) V2 daily climate projections from the CMIP5 multimodel ensemble [37–39] were used to derive meteorological data (i.e., precipitation, maximum, and minimum air temperature) to drive PRMS models. The CMIP5 general circulation models used in this analysis are listed in Table 1. Shapefiles of both basins with delineated HRUs were uploaded to the USGS Geo Data Portal, and the area-weighted grid (12 km) statistics algorithm was used to obtain daily mean precipitation, maximum air temperature, and minimum air temperature for each HRU. These data were evaluated and validated prior to being

made publicly available on the data portal by Bracken [40]. Data for the historical period of 1955–2005 and future projections for the period 2020–2100 were downloaded for each CMIP5 model. A total of 132 climate projections were used for simulations in each basin to drive PRMS simulations and obtain stream flow under two different representative concentration pathways (RCP) for emissions leading to 4.5 and 8.5 additional Wm^{-2} by the end of the 21st Century. The selection of CMIP5 climate projections was based on data availability from the USGS Geo Data Portal. PRMS simulations for the historical period (water years 1955–2000) provided baseline data from which climate change measures were calculated. The baseline values were calculated from either one historical run, for GCMs with only one such run, or the ensemble mean from all historical runs combined for GCMs with multiple historical runs. Simulations driven by statistically downscaled meteorological data were corrected for PRMS model bias, and results from uncorrected simulations and model-corrected were compared for the historical period (1956–2005) and for the future periods 2041–2070 and 2071–2099. Bias correction of stream flows was also done directly on uncorrected GCM-driven stream flow simulations with observed data

Changes in FDCs were analyzed qualitatively by plotting the projected median FDC, 10% exceedance FDC, and 90% exceedance FDC curves and baseline FDC curves at the same exceedances. Quantitative estimates of the proportional change PC (percent change expressed as a decimal) were obtained using GCM-driven simulations for the historical period as a baseline. Changes in stream flows from the projected median FDC, 10% exceedance FDC, and 90% exceedance FDC curves were estimated using:

$$PC_{p_{e,i}} = \frac{(Q_{f_{p_{e,i}}} - Q_{h_{p_{e,i}}})}{Q_{h_{p_{e,i}}}} \text{ for } i = 1, 2, \dots, 365 \quad (3)$$

where $Q_{f_{p_{e,i}}}$ is the GCM-driven projected stream flow for exceedance probability $p_{e,i}$ in the 10%, median, or 90% FDC, $Q_{h_{p_{e,i}}}$ is the GCM-driven simulated historical flow in the 10%, median, or 90% FDC for exceedance probability $p_{e,i}$, and $PC_{p_{e,i}}$ is the proportional change in stream flow in the 10%, median, or 90% FDC for exceedance probability $p_{e,i}$. Positive values of PC indicate an increase in flow in the future, and negative values represent a decrease of flow in the future. Future uncorrected simulations were compared with baseline uncorrected simulations, and future model-corrected simulations were compared with model-corrected baseline simulations in all cases.

Annual and full record FDCs for observed, simulated, and GCM-driven derived stream flows were analyzed using the R package “hydroTSM” [36]. GCM-driven simulations were analyzed for each individual GCM, but results were pooled for simulations with multiple runs (CanESM, 5 runs; CCSM4, 2; CSIRO-Mk3.6.0, 10; IPSL-CMA5-LR, 4; MIROC5, 3; MPI-ESM-LR, 3; MPI-ESM-MR, 3). Flow duration curves from the time slices 2041–2070 and 2071–2099 were developed, and stream flow projections were corrected for model error (model corrected) using quantile mapping with the R package “qmap” [35]. Simulations were model corrected using pooled parameter estimates for GCMs with multiple runs.

3. Results

3.1. Simulations Driven by Observed Meteorological Data

3.1.1. Goodness-of-Fit

Uncorrected simulations of stream flow for water years 1956–2005 performed relatively well in both basins based on several GoF measures used to evaluate the models (Figures 3 and 4). Overall, based on GoF measures, the model performed better in the Batsto basin than in the Maurice basin for daily flows by most measures with less of a difference for monthly flows. This was also the case for GoF measures for uncorrected annual and seasonal flows (Table 3). In the Maurice basin, the monthly R^2 and Nash–Sutcliffe Efficiency (NSE) measures for uncorrected flows improved from

≈0.4 to 0.7 compared with daily flows, and most measures indicated a better fit of model simulations for monthly mean flows over daily flows in both basins.

Uncorrected simulations showed a wide range in GoF measures for seasonally averaged flows. For both basins, absolute PBIAS was greater for seasonally averaged flows than for daily, monthly, and annual mean flows. In Batsto, seasonal bias ranged from −17% (winter) to 30% (summer), while the daily, monthly, and annual PBIASs were −1.4%, −1.4%, and −2.8%, respectively (Table 3). The same trend was apparent in Maurice.

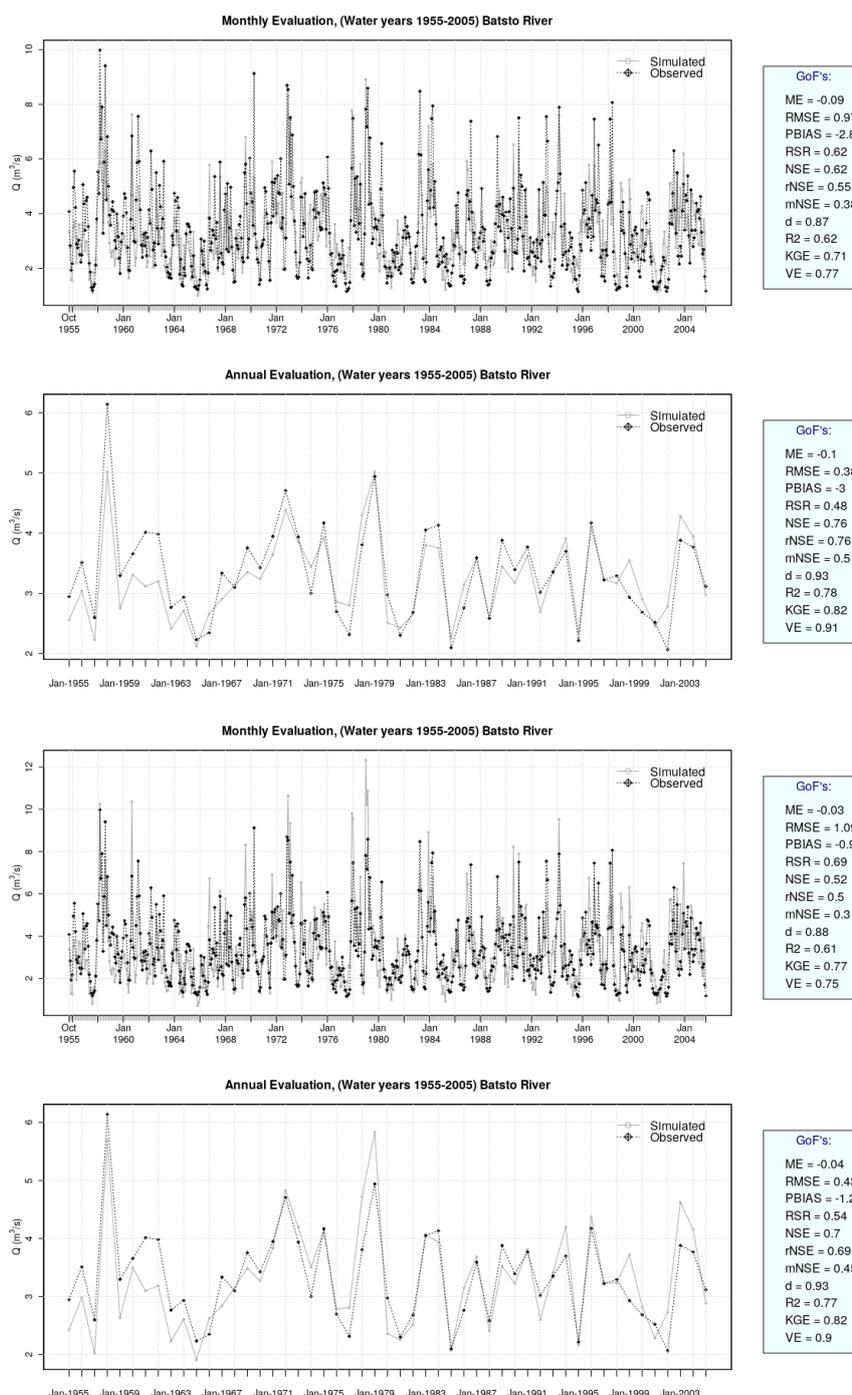


Figure 3. Goodness-of-fit for monthly mean flow (m^3s^{-1} ; top) and annual mean flow (second row) for uncorrected (UC) and monthly mean flow (third row) and annual mean flow (bottom) for bias-corrected (BC) PRMS simulations in the Batsto River.

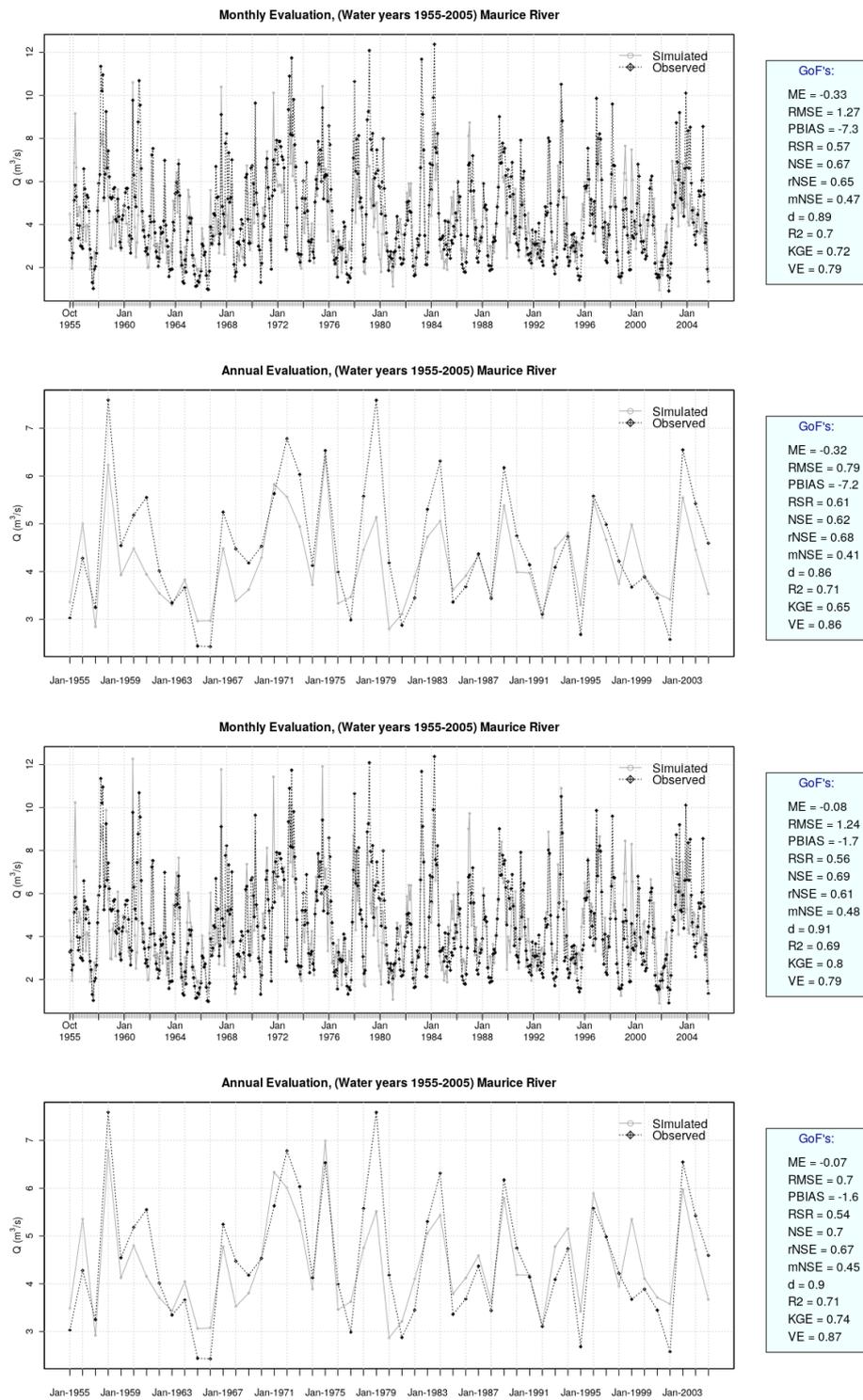


Figure 4. Goodness-of-fit for monthly mean flow (m^3s^{-1} ; top) and annual mean flow (second row) for uncorrected (UC) and monthly mean flow (third row) and annual mean flow (bottom) for bias-corrected (BC) PRMS simulations in the Maurice River.

Table 3. Goodness-of-fit (GoF) measures for uncorrected (UC) and bias-corrected (BC) simulated flows. GoF was measured based on seasonally averaged flows. DJF = December, January, and February, etc. Note that seasonal bias correction was not done based on the seasonal definitions used in the evaluations.

GoF	UC				BC			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Batsto River								
ME (m^3s^{-1})	−0.71	−0.32	0.74	0.04	−0.54	−0.36	0.85	0.06
R ²	0.57	0.61	0.65	0.63	0.56	0.63	0.66	0.61
RMSE (m^3s^{-1})	1.69	1.41	1.61	1.11	1.86	1.32	1.74	1.34
PBIAS%	−16.70	−9.70	28.90	1.40	−12.60	−10.70	32.90	2.10
NSE	0.48	0.54	0.55	0.63	0.37	0.60	0.47	0.45
d	0.82	0.80	0.84	0.88	0.85	0.86	0.87	0.87
KGE	0.62	0.51	0.56	0.74	0.69	0.65	0.61	0.72
VE	0.74	0.76	0.57	0.78	0.70	0.75	0.57	0.74
Maurice River								
ME (m^3s^{-1})	−0.96	−0.57	0.60	−0.25	−0.63	−0.31	0.82	−0.07
R ²	0.39	0.51	0.46	0.47	0.38	0.50	0.45	0.46
RMSE (m^3s^{-1})	2.39	1.90	2.25	1.68	2.41	1.92	2.52	1.78
PBIAS%	−16.70	−11.60	18.50	−6.40	−11.00	−6.30	25.20	−1.70
NSE	0.21	0.45	0.36	0.40	0.20	0.44	0.20	0.33
d	0.75	0.81	0.80	0.82	0.77	0.83	0.79	0.82
KGE	0.56	0.63	0.62	0.67	0.59	0.69	0.58	0.68
VE	0.71	0.74	0.63	0.71	0.71	0.74	0.59	0.69

Bias correction of flows using the FDC for the full record in each basin improved the fit of the overall FDC as expected. However, GoF measures did not improve consistently for bias-corrected simulations (Figure 4 and Table 3). For bias-corrected simulations of Batsto River flows, there were improvements in fit according to the measures of mean error (−0.09 to −0.03) and PBIAS (−2.8% to \approx −0.9%) for daily and monthly mean flows and KGE from 0.62 to 0.75 and 0.71 to 0.77 for daily and monthly mean flows, respectively. For bias-corrected simulations of Maurice River flows, there were improvements in fit according to the measures of mean error (\approx 0.32 to \approx −0.08) and PBIAS (\approx 7% to −1.7%) for daily and monthly mean flows and Kling–Gupta Efficiency (KGE) from 0.63 to 0.87 and 0.72 to 0.80 for daily and monthly mean flows, respectively.

Bias correction of flows did not improve the overall performance of the model on simulations of annual mean flows in either basin, with the exception of PBIAS, which was expected from the bias correction. Goodness-of-fit for seasonally averaged bias-corrected flows was variable in a similar manner to the uncorrected simulated flows. Model PBIAS increased in some seasons and decreased in others for bias-corrected simulations compared to uncorrected simulations. Overall, for both basins, PRMS simulations over the time period (1956–2005) simulated here compared with the shorter evaluation period (1997–2003) from [18] performed as well or better for some measures and worse for others.

3.1.2. Flow Duration Curves

Each basin showed a different pattern of variation in annual FDCs (Figure 5). Observed annual FDCs varied to about the same magnitude across all flow exceedances in the Maurice basin. Observed annual FDCs showed less variability at low flows (high exceedance) in the Batsto basin and greater variability for high flows. Simulated FDCs did not fully capture the overall FDC or the variation of annual FDCs in either basin. Simulations in both basins tended to underestimate high flows and overestimate low flows (Figure 6). Variation at low exceedance levels was underestimated in the Maurice River, and variation at high exceedance levels was overestimated in the Batsto River.

In the Maurice basin, there was a greater tendency for uncorrected simulations to underestimate high flows than to overestimate low flows. Simulations for the Batsto River, on the other hand, underestimated high and overestimated low flows to a similar degree (Figure 7). The FDC in the Maurice basin was smoother than that for the Batsto for flows with $p_e > 0.95$ (low flows). Simulated flows were able to capture this sharper drop in low flows for the Batsto; however, this part of the FDC was not fit well by a lognormal distribution based on the entire range of flows.

The variations in observed annual FDCs (Figure 6) were relatively well represented by uncorrected simulated flow duration curves in both basins. Differences in uncorrected simulated flow bias for each basin were more apparent when comparing the annual FDCs and the 10% and 90% FDCs. For instance, the consistent underestimation of high flows and underestimation of low flows was clearer in Batsto. The middle of the FDC ($0.25 \leq p_e \leq 0.75$) was well represented by uncorrected simulations, as well as the variance in annual FDCs. In Maurice, for $p_e \leq 50\%$, the tendency of uncorrected simulations to underestimate flows can be seen clearly. Uncorrected simulations in Maurice also showed slightly lower variation than observed flows in the annual FDCs.

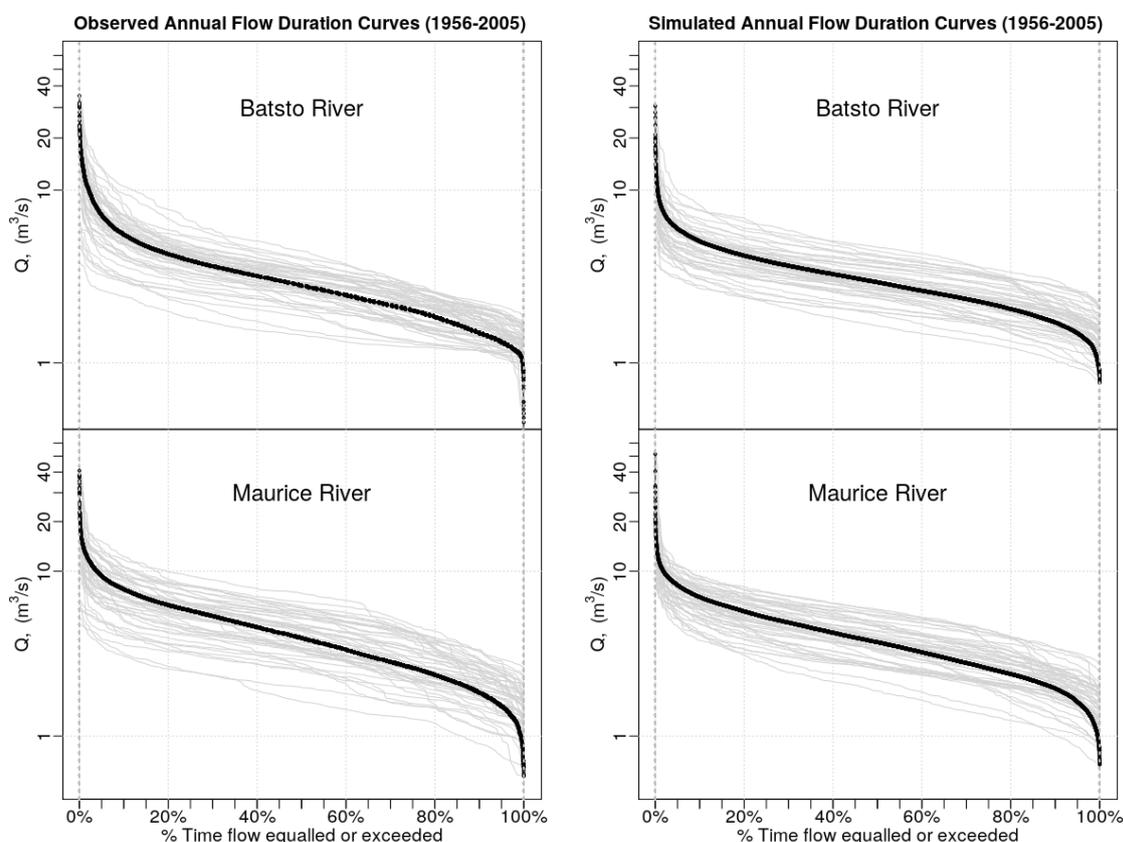


Figure 5. Observed (left) and PRMS simulated (right) annual flow duration curves for the Batsto and Maurice Rivers for water years 1956–2005. Thick black line represents the FDC for the full period of record.

Bias correction of stream flow for the calibrated PRMS model was able to correct the systematic model error of the overestimation of low flows and underestimation of high flows in both basins (Figure 6). Bias-corrected simulations for both basins tended to increase the spread of the 10% and 90% FDCs over the full range of p_e with the exception of at low p_e in Maurice, which indicated greater uncertainty with bias correction. Overall, FDCs representing bias-corrected simulations indicated that the model performed better in the Maurice River than in the Batsto River for both high and low flows and with respect to the variation in flows.

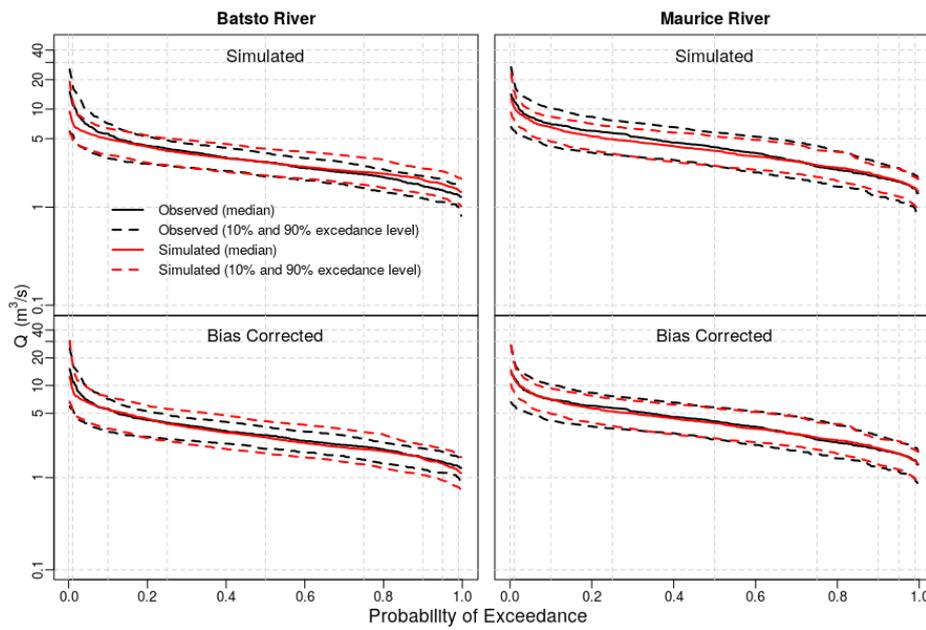


Figure 6. Median, 10%, and 90% exceedance levels for annual flow duration curves for simulations in each basin. The upper two plots show the calibrated PRMS simulations for uncorrected stream flows compared to observed flows. The second row of plots shows results from bias-corrected simulations compared to observed flows.

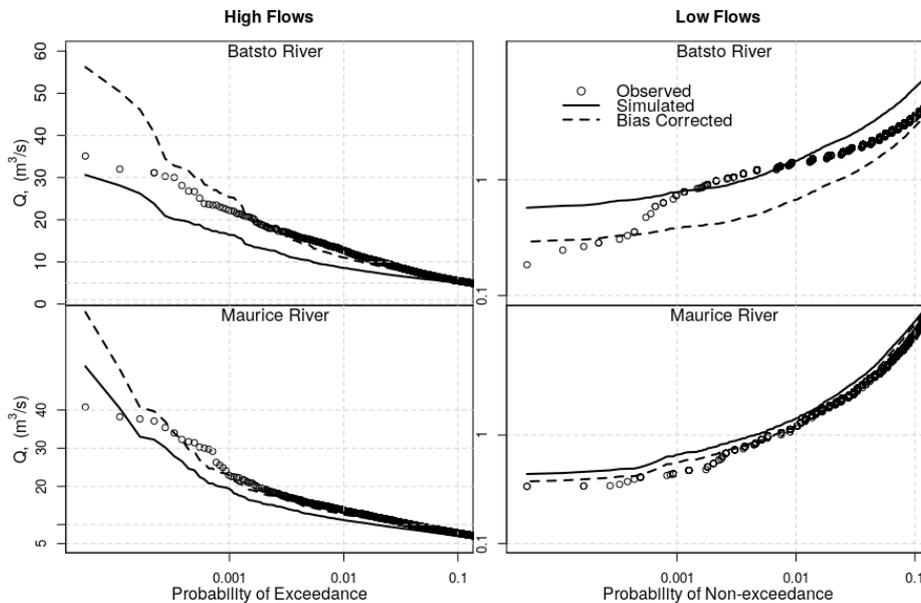


Figure 7. Tail of the full record flow duration curves for high and low flows (right and left plot in each row, respectively) in each river as indicated.

3.2. Simulations Using GCM Data

Hydrologic simulations driven by GCM data could only be evaluated using FDCs. Estimated parameters for the fitted lognormal distributions for uncorrected and model-corrected simulations driven by GCM data showed greater variation than parameters bias corrected directly with observed data (Figure 8). GCM-driven simulated stream flows showed relatively low overall variation for estimated parameters over the historical period and increased variation in parameters for climate change projections, which indicated over-correction of stream flows. It is likely that bias

correction of GCM-driven stream flow simulations with observed flow data, in this case, reduced the simulation of the natural variation of the hydrologic model. Therefore, the analysis of climate change impacts on flow duration curves was done using only model-corrected GCM simulated stream flows.

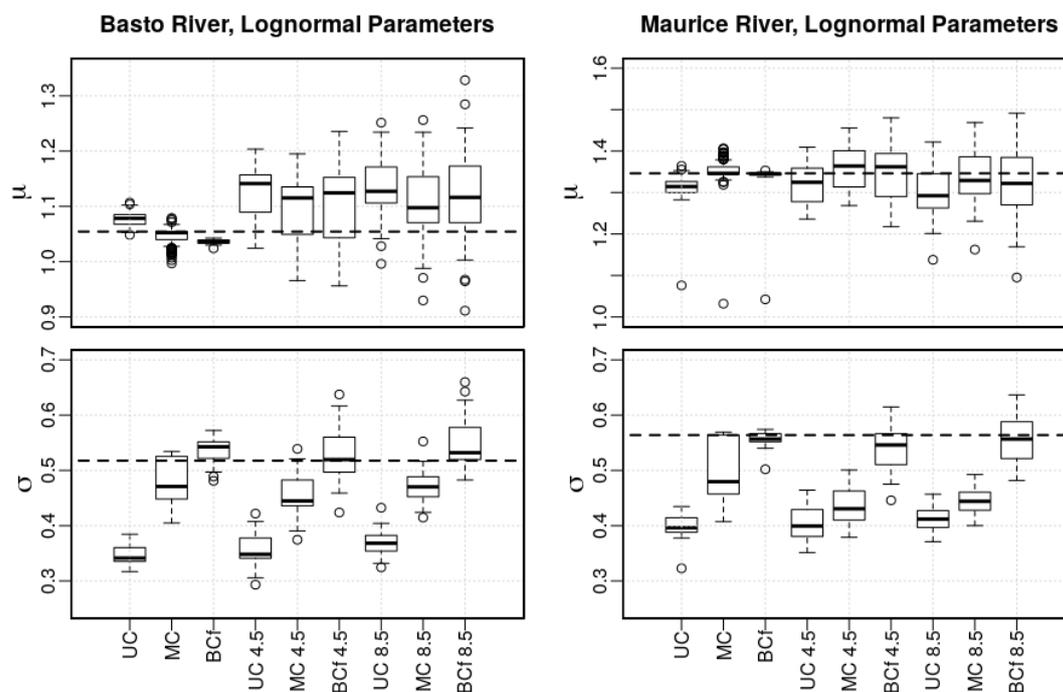


Figure 8. Lognormal distribution parameters, mean μ (top) and standard deviation σ (bottom), for uncorrected (UC), model-corrected (MC), and bias-corrected (BC) GCM simulations. UC, MC, and BC are GCM simulations for the historical period (1956–2005). Parameters are shown for GCM simulations for RCP 4.5 and RCP 8.5 for UC, MC, and BC. Parameter values for observed fitted lognormal distributions are shown as the dotted line in each plot.

3.2.1. Historical Simulations

Overall, the annual FDCs from GCM-driven stream flow simulations for the historical period for uncorrected and model-corrected stream flows fell within the range of observed annual FDCs (Figure 9). Annual FDCs from GCM-driven simulated stream flows for the historical period showed a wide range of variability dependent on the GCM (supplemental figures), and some models showed greater variation than others. FDCs based on the ensemble of uncorrected GCM simulations for the historical period tended to overestimate low flows and underestimate high flows (Figure 9) to an even greater degree than uncorrected simulations driven by observed meteorological data. Application of the model correction to GCM-driven simulations greatly reduced this bias, but did not fully eliminate it. Model correction of GCM-driven simulations increased the variance of annual FDCs in both basins.

The 90% FDC for high flows from uncorrected GCM-driven simulations of the Batsto River were aligned with the median FDC of observed high flows, $p_e \lesssim 0.05$, whereas the opposite was the case for low flows with $p_e \gtrsim 0.95$. Implementing the model correction for GCM simulated stream flows in Batsto was able to correct for bias in most of the high flows and greatly improved the estimates of low flow frequencies. A similar trend was apparent in the Maurice River, though to a lesser extent at both high and low flows. The model correction of GCM-driven simulations in Maurice improved estimates of the frequency distribution, but not as much as in Batsto.

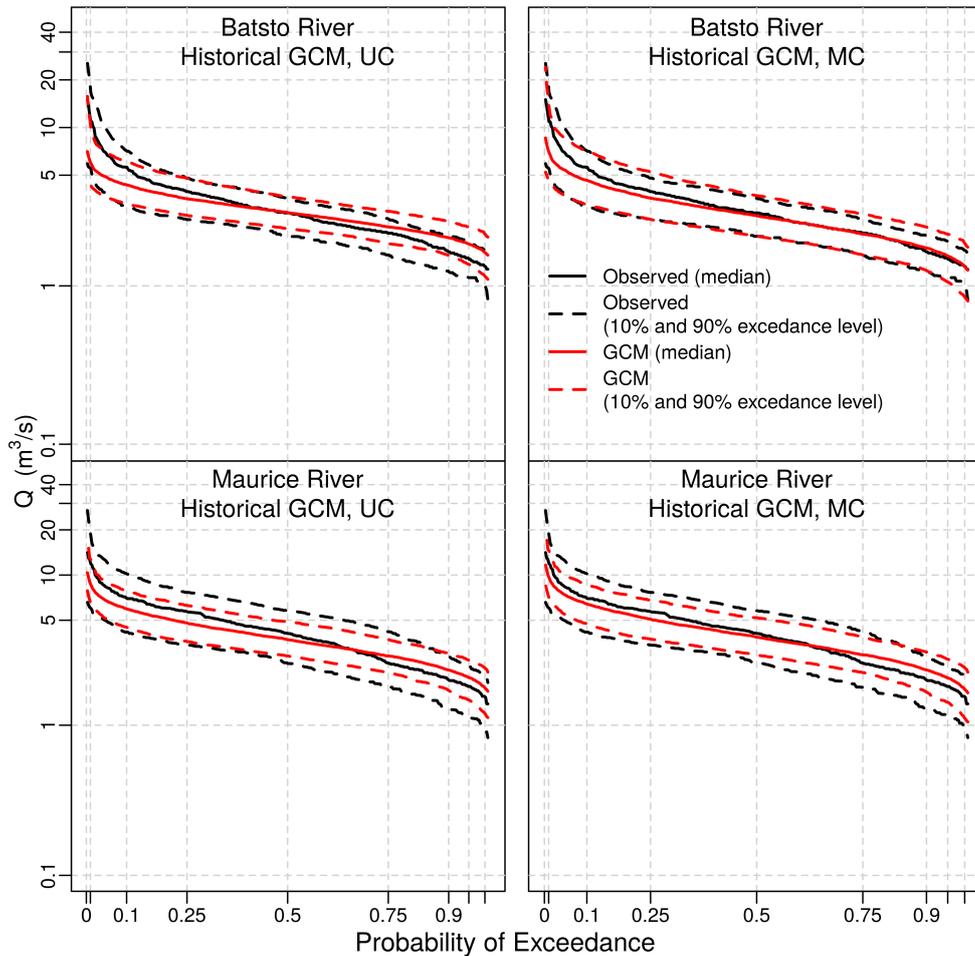


Figure 9. Annual flow duration curves for stream flow simulated for the Batsto River using downscaled climate projections for each GCM simulation (uncorrected (UC), and model corrected (MC)) used in the analysis. The observed FDC is the thick black line, and the thick blue line is the FDC for the simulated data using the full period.

3.2.2. Projected Simulations

The projected changes in parameter values for the full FDCs indicated a much greater uncertainty in estimates of FDC than can be seen in the qualitative assessment of annual FDCs. Overall estimates of σ did not change much for projected distributions; however, there was much greater variation in estimates of the mean, μ , for projected distributions of stream flows, or FDCs (Figure 8).

The annual FDCs for projected stream flows under RCPs 4.5 and 8.5 in both time slices (2041–2070 and 2071–2099) indicated some change in overall frequency of flows for both basins (Figures 10 and 11). As with historical simulations, FDCs from GCM simulated stream flows for the historical period showed a wide range of variability dependent on the GCM (supplemental figures). The FDCs showed an increase in variation (wider spread in 10% and 90% FDCs) with climate change across the full range of p_e in both basins. The projected impacts of climate change on the FDCs were well quantified by measures of the proportional change in flows by quantile, PC_{p_e} (Figures 12 and 13).

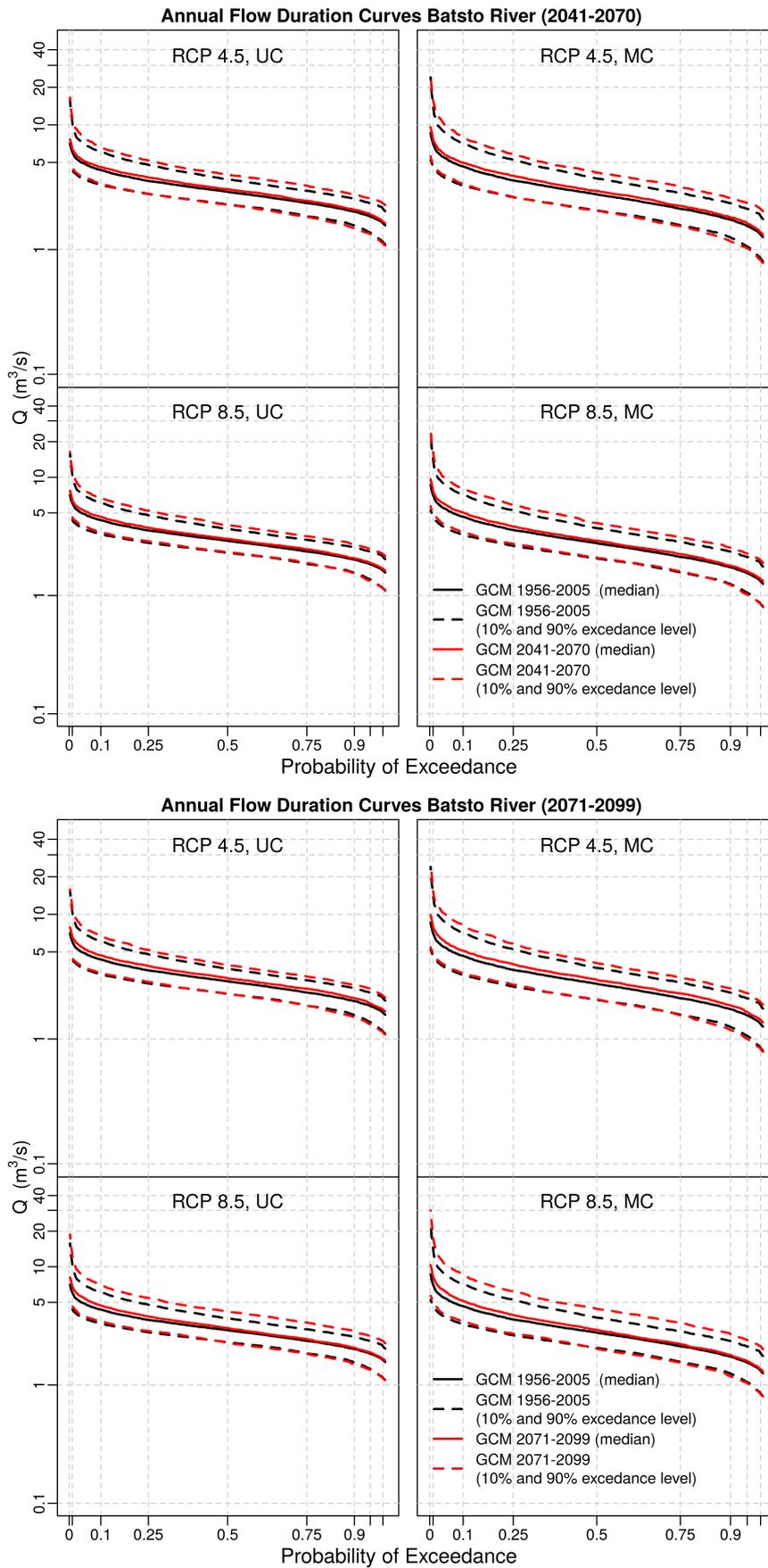


Figure 10. Projected flow duration curves for the Batsto River under emission scenarios RCP 4.5 and 8.5 for the two time slices as indicated.

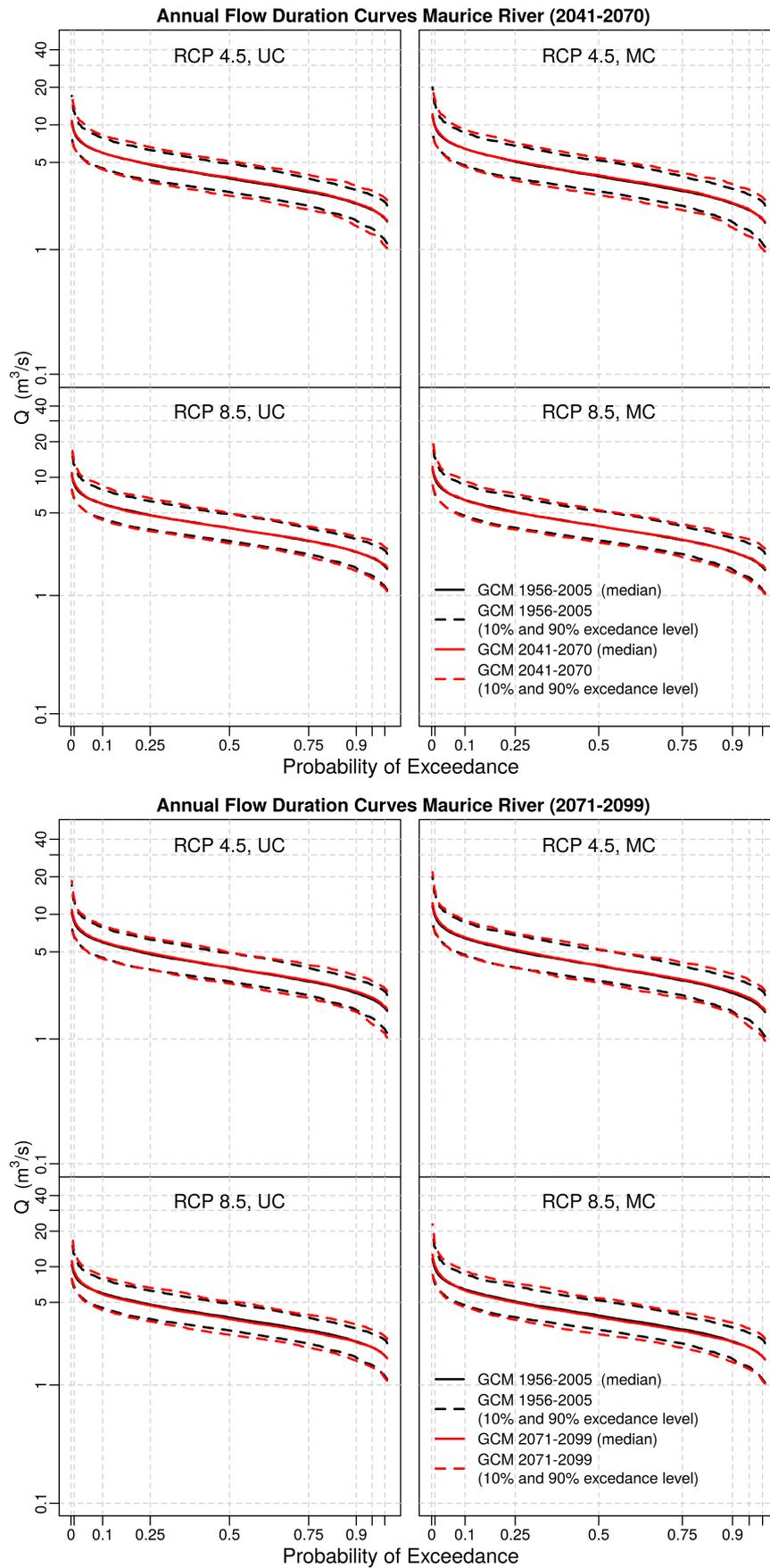


Figure 11. Projected flow duration curves for the Maurice River under emission scenarios RCP 4.5 and 8.5 for the two time slices as indicated.

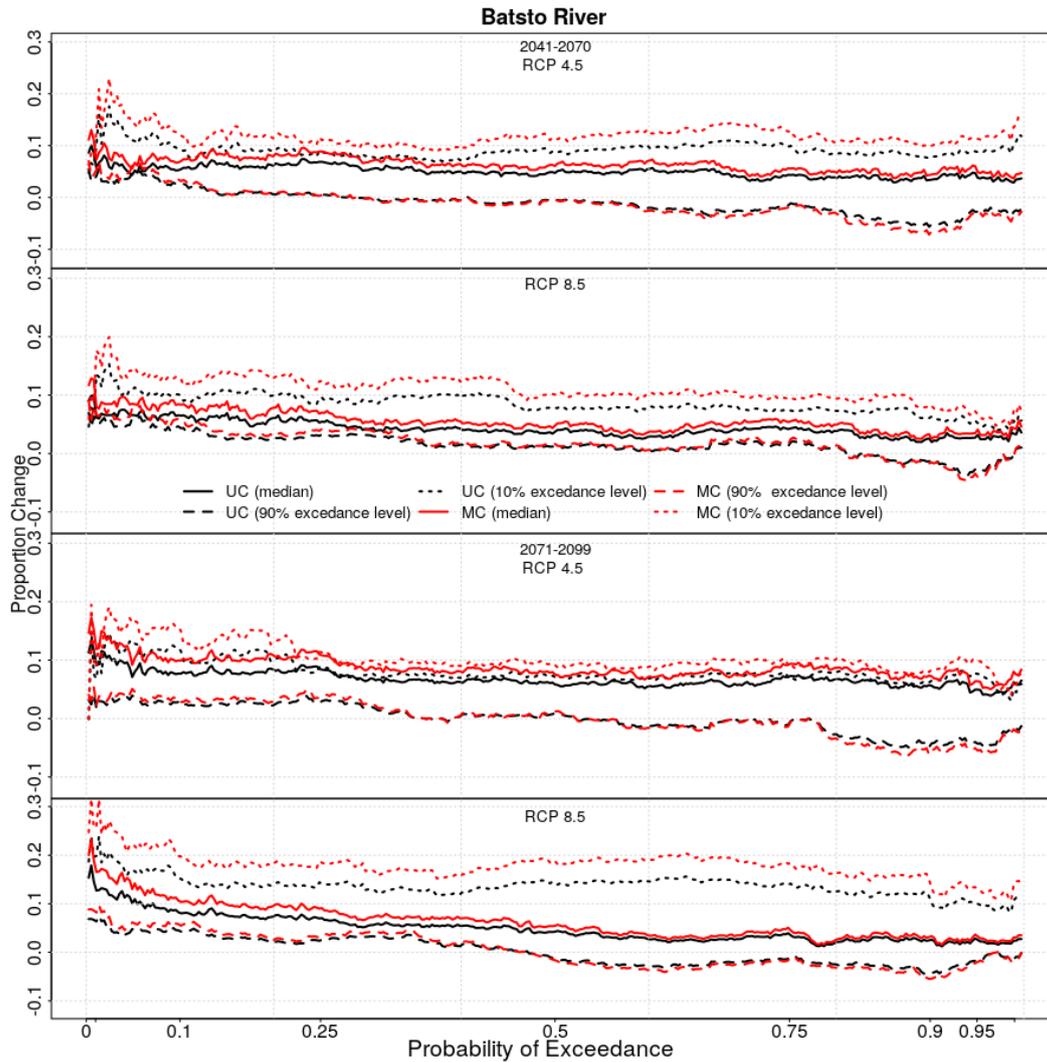


Figure 12. Projected change in stream flows as indicated by the proportional change in stream flow at each exceedance level of the FDC for the Batsto River. Projected change in flow is shown for each RCP for uncorrected (UC) and model-corrected (MC) GCM simulations.

For the Batsto River, the median FDC and 10% FDC were projected to increase across all probabilities under both RCPs in both time slices (2041–2070 and 2071–2099). Model-corrected simulations indicated a slightly greater increase in flow levels than uncorrected projected stream flows. The 10% FDC for both model-corrected and uncorrected projected flows increased for higher flows (low p_e) and decreased for low flows (high p_e). There was a greater increase in the projected median FDC for high flows than for low flows as indicated by a slight negative slope of the median FDC difference curve. Under RCP 4.5 for 2071–2099, both median and 90% FDCs changed about the same amount, and under RCP 8.5, the median FDC seemed to return closer to the baseline FDC despite a relatively large increase in the 90% FDC. The magnitude of low flows showed an increase under RCP 4.5 of around 5% (Figure 12), but under RCP 8.5, projections indicated little or no increased flow at these high p_e levels. At $p_e < 0.1$, stream flows were projected to increase by a greater amount than for low flows ($p_e > 0.9$) under both RCPs, and a greater increase of high flows ($p_e \geq 0.2$) was projected under RCP 8.5.

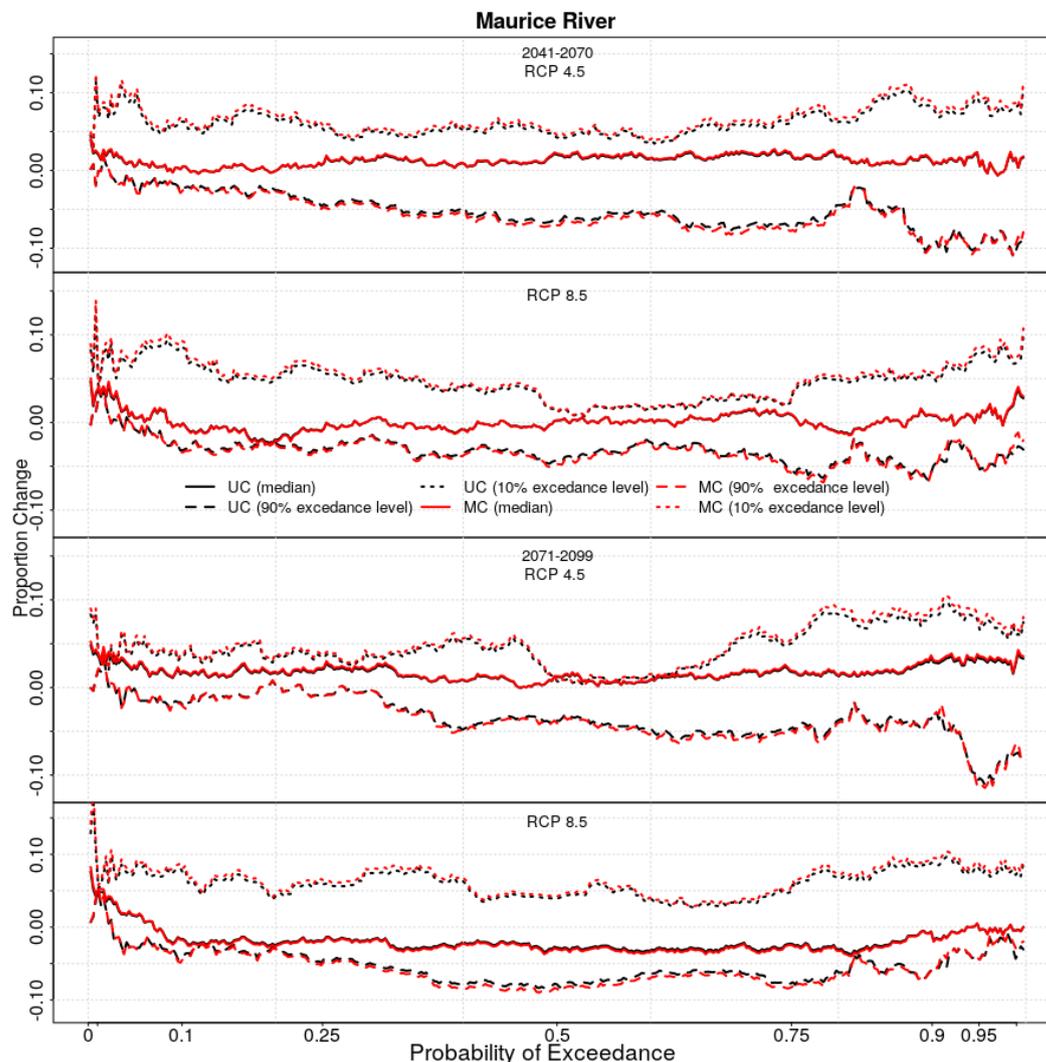


Figure 13. Projected change in stream flows as indicated by the proportional change in stream flow at each exceedance level of the FDC for the Maurice River. Projected change in flow is shown for each RCP for uncorrected (UC) and model-corrected (MC) GCM simulations.

For the Maurice River, there was little or no difference between uncorrected and model-corrected simulations of projected stream flows (Figure 13). Under RCP 4.5 for both time slices, the change in median FDCs indicated that flows were not projected to increase by more than around 2.5% at most p_e , with slightly greater increases in flow indicated at $p_e \lesssim 0.02$. Under RCP 8.5, projections for 2041–2070 indicated that median flows will not change except at the extremes, $p_e > 0.9$ and < 0.1 . Projections also indicated an increase of 5–10% in the 10% FDC for $p_e > 0.75$ and $p_e < 0.1$ suggesting that higher low flows and greater high flows may occur in the future. Projections in FDCs for 2071–2099 under RCP 8.5 showed a decrease in the median FDC in Maurice for all $p_e \gtrsim 0.07$, but with an increase in flows with $p_e \lesssim 0.05$ and extreme high flows could increase by $> 10\%$. Furthermore, projections suggested that flows ranging from $0.1 \leq p_e \leq 0.9$ could decrease slightly.

4. Discussion

An important consideration when evaluating hydrologic model performance is whether or not the model can be considered an appropriate representation, or hypothesis, of the relevant processes in a simulated basin [41,42], which can be difficult due to uncertainty in the modeling chain. For climate change impact assessments, top-down approaches propagate uncertainty from emission scenarios, global climate models, regional climate models, through downscaling and bias

correction (of meteorological data) methods [17] to data that are used to drive hydrologic models. Additionally, there is error in observed data used to estimate model parameters (e.g., land use and soils), meteorological data used as input (temperature and precipitation), and flows used for calibration/evaluation. An important question remains, “Are we getting the right answers for the right reasons?” [1]. This limits the application of hydrologic models to assess the potential impacts of climate change for local and regional decision making due to the uncertainty of model outputs.

Results were reported by Daraio [18] with relatively low overall confidence in large part because of the limited evaluation period and existing bias in stream flow simulations. The extended evaluation of model simulations done here provided an assessment of the performance over a long-term historical time period that extended beyond the calibration period, which should be done before its use to project potential future stream flows [24]. The analysis using bias correction of stream flows from PRMS simulations allowed for more confidence in projected climate change impacts in both basins, and in the Maurice River in particular.

4.1. Performance of the Hydrologic Model

Goodness-of-fit (GoF) measures are the traditional way to assess the performance of hydrologic models. Moriasi et al. [43] suggested that for watershed models, GoF measures for daily, monthly, or annual simulations should have values of $R^2 > 0.60$, $NSE > 0.50$, and $PBIAS \leq \pm 15\%$. Evaluation of model performances using these metrics suggested that the model in the Batsto River performed adequately on daily mean, monthly mean, annual mean, and (most) seasonal mean flows, but only monthly mean flows were adequately simulated in the Maurice River. Ritter and Muñoz-Carpena [44] suggested that an acceptable model would have an $NSE \geq 0.65$, in which case only annual flow simulations in the Batsto River and annual and monthly mean flows in the Maurice River would be considered adequate. The use of the coefficient of variation (R^2) and other correlation measures for model calibration and evaluation has been questioned [45], but it was included here primarily for model comparison. Additionally, several variants of the NSE were used to supplement this efficiency measure. The Kling–Gupta efficiency improves on the NSE because it includes more information on model performance including correlations, bias, and variability [36,46]. The index of agreement d provides a measure of differences in the means for observed and simulated flows and does not include a correlation measure, though the measure is sensitive to the influence of extreme values due to the use of squared differences [45]. The Volumetric Efficiency (VE) represents the fraction of water delivered at the proper time and in itself provides some measure of the skill of the model to capture an important aspect of the flow regime. Including these GoF measures for models in both basins indicated that PRMS simulations over the time period 1956–2005, compared with the shorter evaluation period (1997–2003) from Daraio [18], performed as well or better for some measures and worse on others. Overall, this increased the confidence in model performance for both basins.

All measures of GoF for daily mean and monthly mean flows in Batsto were slightly better or about the same over 1956–2005 compared with the shorter evaluation period. This was not the case for daily mean flows in the Maurice River where GoFs were better for the shorter evaluation period, but monthly mean flow GoFs for 1956–2005 were about the same for most measures. On a seasonal basis, there were no consistent patterns in the differences in GoF measures between the longer and shorter evaluation periods in each basin. Some measures improved over the longer period, and others indicated poorer performance. Both models performed better for monthly mean flows, which may be more relevant for use in long-term climate change projections. Bias correction of simulated flows using composite FDCs from previously calibrated PRMS models was able to improve model performance on some measures of GoF, and overall, they were as good as uncorrected simulations. The most important improvement was the reduction in the degree of overestimation (underestimation) of low (high) flows in both basins for bias-corrected simulated stream flows.

While it is difficult to objectively interpret the GoF measures, in the context of each model being a hypothesis for the relevant processes in the basins [42], there is not enough evidence to reject them. Overall, the bias-corrected calibrated models of the Batsto and Maurice Rivers are good models, and GoF measures indicated better model performance in the Batsto. Assessment with FDCs, on the other hand, indicated better simulations of flow regime in the Maurice. Models in both basins do have important limitations, and the variation in GoF scores provided an indication of the strengths and weaknesses of the models. In particular, bias correction of stream flows was able to identify areas where a model may not simulate physical processes well.

The primary weakness of the simulations lies in the performance of the model at both high and low p_e levels, or low and high flows, respectively. The observed FDC for the Batsto River shows reduced variation at low flows and a sharper curve in the FDC at the highest p_e level. This is most likely due to the strong groundwater component of the Batsto, especially at low flows [47]. The lognormal distribution fit to the data was not able to capture low flows at high p_e , or the tail of the distribution, so the fitted FDC is more smoothly curved than the FDC from observed flows in Batsto. This is an indication that there may be a structural error in the model that does not correctly simulate the physical processes that dominate at low flows.

The PRMS simulations driven by observed meteorological data did not capture high flows, or low p_e levels, as well either. Bias correction of model simulations improved performance at high flows for simulations, but again, this may indicate a problem with the PRMS model. The error in the PRMS model at high flows seems to be greater in the Maurice River than Batsto, and it may be due to some misrepresentation of runoff processes due to the greater area of mixed urban development in the Maurice watershed. Additionally, the relative performance of the models, both uncorrected and bias corrected, at the seasonal scales indicates the need to test bias correction of stream flows using seasonally based FDCs. This work is in progress by the author's research group.

4.2. Downscaled GCM-Driven Simulations

4.2.1. Bias Correction of Stream Flows

Statistically downscaled GCM or GCM-driven stream flow simulations showed bias at both extremes where the FDC was not as well simulated as the middle of the FDCs ($p_e \approx 0.5$), which can be primarily attributed to the performance of the PRMS models. For uncorrected simulations, use of GCM data increased the PRMS model bias at both ends of the FDC, and correcting for model bias improved the representation of GCM-driven simulations over the historical period.

Direct bias correction of GCM-driven stream flow simulations with observed flows—which includes both PRMS model bias and any bias that can be attributed to the meteorological input data—was able to improve the overall fit of the FDC over model-corrected GCM-driven simulations in both basins. However, the direct bias correction of GCM stream flows seems to reduce the representation of the natural variation of simulated flows over the historic period while increasing variation over climate change projected stream flows. This would occur if the timing of, for instance, high flows in simulations did not coincide with the timing of observed high flows. High flows tend to occur in summer in both basins. If high flows were simulated to occur in winter, then the mismatch in timing would lead to correcting for flows generated by different physical processes in the different seasons, e.g., rain on snow in winter versus a large tropical event in summer. Bias correction of stream flows without model calibration has been shown to reduce seasonal forecast error for forecasts that focused on the volume, but not timing, of flows [48]. Improper timing of bias corrections may improve the overall fit of the FDC, but would lead to poor climate change projections. Therefore, no GCM-driven simulations were bias corrected directly with observed flows, and bias correcting stream flows would need to include timing to ensure better results.

4.2.2. Climate Change Projections

Overall, the correction of model bias was effective for improving the fit of the FDCs, but projected changes in FDCs based on uncorrected and model-corrected GCM-driven simulations were not consistent in the two basins. In Maurice, projected changes in FDCs were almost identical for uncorrected and model-corrected GCM simulated stream flows. In Batsto, model-corrected GCM-driven simulations showed a greater change in flows at the median FDC level compared to uncorrected simulated stream flow and an even greater increase in flows at the 10% FDC level. It seems likely that the lack of difference in projected change in the FDCs between uncorrected and model-corrected stream flow simulations in Maurice indicates higher confidence in climate projections in this basin. The FDC was better simulated in Maurice than in Batsto; therefore, it is likely that correction was done with proper timing, as discussed above, which led to more consistent correction of flows and overall better model fit. Consistent with this interpretation, the analysis for simulations in Batsto provided a better understanding of the limitations of the model and allow for a more clear interpretation of results. Most importantly, it seems to be an indication of greater uncertainty for projections in the Batsto River compared with projections in Maurice.

Projected trends in both basins included an increase in flow variability (spread of the 10% and 90% FDCs), an increase in median flows through 2070 under both RCPs, and reduced changes in the median FDCs by 2099 under RCP 8.5. The former indicates an increase in the variation of flows across all probabilities of exceedance, in particular at high flows. Both rivers are projected to have greater high and low flows, and the projections indicate more variation at high flows in both basins. Model-corrected simulations had greater variation for both projected FDCs and for the baseline (historical) GCM simulations. However, some bias correction methods are known to lead to variance inflation [17], which could increase the uncertainty of simulated flows. Projected high flows were larger with a greater likelihood that annual maximum flows will be more extreme. It is important to note that it is well known that one of the weaknesses of GCMs is an inability to capture extremes [49,50]. While both the PRMS model bias and model-corrected GCM-driven bias in simulated stream flows indicated an underestimation of high flows and overestimation of low flows, the consistency of the error provides a level of confidence in these projections despite the bias. That is, based on these projections, it seems likely that extreme floods will be higher, by the end of this century compared to the end of the last, by 5 to 30% in Batsto and up to 15% in Maurice. However, more detailed analysis of projected changes in extremes, both high and low, is necessary to reduce the uncertainty in the interpretation of these projections.

There were important differences in projections in each basin that are clear based on the changes in the FDCs. For instance, in the Batsto River, climate change projections indicated an increase in mean flows under RCP 4.5—with slightly greater changes indicated by model-corrected compared to uncorrected projected stream flows (see above). Projected stream flows were not greater under the RCP 8.5 emission scenario than for RCP 4.5 in Batsto for 2041–2070. By contrast, variation in flows increased in the Maurice River under both RCPs; however, the overall mean flow was not projected to increase by much, and the median FDC indicated that flows across most probabilities of exceedance will be lower by the end of the century. In Batsto, the range of flow variation was projected to increase while the low flow regime returns to baseline by the end of the century for RCP 8.5.

Impacts of Watershed Characteristics

The Batsto basin lies within the protected Pinelands National Reserve in New Jersey, and there is a relatively small amount of agriculture in the basin ($\approx 10\%$ of the basin). Land use change and population growth will be minimal in this basin because of its location in the Reserve. The potential impacts of the projected changes in stream flow will primarily impact the ecology of the river. The results here indicate some change in the flow regime, but do not provide enough information to fully assess the potential ecological impacts. Such an assessment requires a more detailed analysis of the natural flow regime relevant to ecological characteristics of the river. Poff et al. [51] suggested

that the natural flow regime can be defined using flow magnitudes, the frequency of occurrence, the duration of specified flow conditions, the timing and/or predictability of flows, and the rate of change or flashiness of flow. These indicators of the flow regime can be quantified using a number of measures including the autocorrelation structure of the time series, rising and descending limb density, peak distribution, low flows, and the FDC [52]. On their own, the FDCs analyzed in this paper represent an incomplete picture of the natural flow regime. The author's research group is continuing the current analysis to include ecologically relevant hydrologic indices that can describe most of the natural variation in flows, which is one of the key goals of hydro-ecological classification of rivers [53,54].

The upper Maurice basin is much more developed, and urbanization is expected to continue in the future as the population continues to increase. The climate change projections do not include potential changes to land use and the expected growth of urban area in the basin; therefore, the results must be interpreted in light of this limitation. For instance, the potential increase in high flows due to climate change is likely to be exacerbated by increased urbanization. Additionally, projected decreases in low flows have potential implications for water resources in Maurice since they signify a reduction in base-flow and thus groundwater recharge, and the majority of drinking water is obtained from groundwater sources in the watershed [18]. However, knowing the potential changes due to climate change can help water resource managers and engineers adapt water infrastructure systems and control that which is amenable to such control. The use of FDCs to show these potential changes could be a strong means to convey the impacts of climate change on water resources and guide such decisions.

5. Conclusions

Evaluation of GCM-driven stream flow simulations requires techniques that provide a measure of a regional or long-term flow regime. This entails the use of methods that are commonly used for hydrologic simulations in ungauged basins. This is particularly important for the use of models for climate change projections since future stream flows even in a gauged basin cannot be observed. A model that performs well over longer periods that include natural climate variation presumably can be used with confidence for climate change projections. Overall, PRMS simulations of both the Batsto and upper Maurice watersheds were able to simulate the FDC and the annual variation of FDCs within each basin. The application of a bias correction for model error on simulated stream flows improved the fit of the FDC over the historical period. While correcting for PRMS model bias seemed to increase the uncertainty FDC estimates representing the flow regime, bias correction can help identify potential model uncertainty and distinguish sources of uncertainty. Bias correction of stream flows indicated the presence of some potential structural errors in the model for the Batsto River that were not apparent by looking at GoF measures alone. The implementation of the model correction on GCM-driven simulated stream flows also indicated greater uncertainty in projected stream flows for Batsto in comparison with projections for Maurice.

The extended analysis of the FDCs for GCM-driven PRMS simulations indicated that climate model-driven hydrologic simulations can be used with some confidence to assess potential changes in FDCs under climate change. In both basins, the climate change projections can be assessed with higher confidence than those from Daraio [18]. Applications in water resources and hydrology that utilize flow duration curves can use projected changes to assess the potential impacts of climate change relatively easily, including a relatively clear assessment of uncertainty. The results leave open the question as to the possibility of using FDCs based on seasonal or monthly flows to bias correct GCM-driven simulations with observed data and apply corrections to climate change projections. The use of FDCs at these time scales may allow for the timing of flows to be included in the corrections, reduce uncertainty, and improve confidence in climate change projections of stream flows using GCM-driven hydrologic simulations.

Funding: This research received no external funding.

Acknowledgments: I would like to thank my graduate student Diana Sankar for her assistance in the literature review. The simulated stream flow used is available on HydroShare, <http://www.hydroshare.org/resource/e286de0f957941a2b3dc9c230ccdd9d4>. The R source code used for analysis is available from the author upon request. I acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and I thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led the development of the software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GCM	general circulation model
BCCA	bias correction with constructed analogs
FDC	flow duration curve
PRMS	precipitation-runoff modeling system
NJ	New Jersey
CMIP	Coupled Model Intercomparison Project
CDF	cumulative density function
UC	uncorrected
BC	bias corrected
MC	model corrected
GoF	goodness-of-fit

References

1. Westerberg, I.K.; Guerrero, J.L.; Younger, P.M.; Beven, K.J.; Seibert, J.; Halldin, S.; Freer, J.E.; Xu, C.Y. Calibration of hydrological models using flow-duration curves. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2205–2227. [[CrossRef](#)]
2. Shamir, E.; Imam, B.; Morin, E.; Gupta, H.V.; Sorooshian, S. The role of hydrograph indices in parameter estimation of rainfall-runoff models. *Hydrol. Process.* **2005**, *19*, 2187–2207. [[CrossRef](#)]
3. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.* **1998**, *34*, 751–763. [[CrossRef](#)]
4. Parada, L.M.; Fram, J.P.; Liang, X., Multi-Resolution Calibration Methodology for Hydrologic Models: Application to a Sub-Humid Catchment. In *Calibration of Watershed Models*; American Geophysical Union (AGU): Washington, DC, USA, 2013; pp. 197–211, doi:10.1029/WS006p0197. [[CrossRef](#)]
5. Vaze, J.; Post, D.; Chiew, F.; Perraud, J.M.; Viney, N.; Teng, J. Climate non-stationarity – Validity of calibrated rainfall-runoff models for use in climate change studies. *J. Hydrol.* **2010**, *394*, 447–457. [[CrossRef](#)]
6. Kurkute, S.; Li, Z.; Li, Y.; Huo, F. Assessment and projection of the water budget over western Canada using convection-permitting weather research and forecasting simulations. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 3677–3697, doi:10.5194/hess-24-3677-2020. [[CrossRef](#)]
7. Qing, Y.; Wang, S.; Zhang, B.; Wang, Y. Ultra-high resolution regional climate projections for assessing changes in hydrological extremes and underlying uncertainties. *Clim. Dyn.* **2020**. [[CrossRef](#)]
8. Chen, C.; Kalra, A.; Ahmad, S. Hydrologic responses to climate change using downscaled GCM data on a watershed scale. *J. Water Clim. Chang.* **2018**, *10*, 63–77. [[CrossRef](#)]
9. Burgan, H.I.; Aksoy, H. Monthly Flow Duration Curve Model for Ungauged River Basins. *Water* **2020**, *12*, 338, doi:10.3390/w12020338. [[CrossRef](#)]
10. Chouaib, W.; Alila, Y.; Caldwell, P.V. On the use of mean monthly runoff to predict the flow-duration curve in ungauged catchments. *Hydrol. Sci. J.* **2019**, *64*, 1573–1587. [[CrossRef](#)]
11. Hrachowitz, M.; Savenije, H.H.; Blöschl, G.; McDonnell, J.J.; Sivapalan, M.; Pomeroy, J.W.; Arheimer, B.; Blume, T.; Clark, M.P.; Ehret, U.; et al. A decade of Predictions in Ungauged Basins (PUB)-a review. *Hydrol. Sci. J.* **2013**, *58*, 1198–1255. [[CrossRef](#)]

12. Troin, M.; Arsenault, R.; Martel, J.L.; Brissette, F. Uncertainty of Hydrological Model Components in Climate Change Studies over Two Nordic Quebec Catchments. *J. Hydrometeorol.* **2018**, *19*, 27–46. [[CrossRef](#)]
13. Santos, C.A.; Almeida, C.; Ramos, T.B.; Rocha, F.A.; Oliveira, R.; Neves, R. Using a hierarchical approach to calibrate SWAT and predict the semi-arid hydrologic regime of northeastern Brazil. *Water* **2018**, *10*, 1137, doi:10.3390/w10091137. [[CrossRef](#)]
14. Kim, D.; Jung, I.W.; Chun, J.A. A comparative assessment of rainfall-runoff modeling against regional flow duration curves for ungauged catchments. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5647–5661. [[CrossRef](#)]
15. Vogel, R.M.; Fennessey, N.M. Flow Duration Curves II: A Review of Applications in Water Resources Planning. *Water Resour. Bull. Am. Water Resour. Assoc.* **1995**, *31*, 1029–1039. [[CrossRef](#)]
16. Muerth, M.J.; Gauvin St-Denis, B.; Ricard, S.; Velázquez, J.A.; Schmid, J.; Minville, M.; Caya, D.; Chaumont, D.; Ludwig, R.; Turcotte, R. On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1189–1204. [[CrossRef](#)]
17. Maraun, D.; Widmann, M. *Statistical Downscaling and Bias Correction for Climate Research*; Cambridge University Press: Cambridge, UK, 2018. [[CrossRef](#)]
18. Daraio, J.A. Potential Climate Change Impacts on Streamflow and Recharge in Two Watersheds on the New Jersey Coastal Plain. *J. Hydrol. Eng.* **2017**, *22*, 05017002. [[CrossRef](#)]
19. Leavesly, G.H.; Markstrom, S.L.; Viger, R.J. USGS Modular Modeling System (MMS)–Precipitation-Runoff Modeling System (PRMS). In *Watershed Models*; Singh, V.P., Frevert, D.K., Eds.; Taylor & Francis: Boca Raton, FL, USA 2006; pp. 159–178.
20. Hay, L.E.; Markstrom, S.L.; Ward-Garrison, C. Watershed-Scale Response to Climate Change through the Twenty-First Century for Selected Basins across the United States. *Earth Interact.* **2011**, *15*, 1–37. [[CrossRef](#)]
21. Viger, R.J.; Hay, L.E.; Markstrom, S.L.; Jones, J.W.; Buell, G.R. Hydrologic Effects of Urbanization and Climate Change on the Flint River Basin, Georgia. *Earth Interact.* **2011**, *15*, 1–25. [[CrossRef](#)]
22. Daraio, J.A.; Bales, J.D. Effects of Land Use and Climate Change on Stream Temperature I: Daily Flow and Stream Temperature Projections. *JAWRA J. Am. Water Resour. Assoc.* **2014**, *50*, 1155–1176. [[CrossRef](#)]
23. Fahad, G.R.; Nazari, R.; Daraio, J.; Lundberg, D.J. Regional Study of Future Temperature and Precipitation Changes Using Bias Corrected Multi-Model Ensemble Projections Considering High Emission Pathways. *J. Earth Sci. Clim. Chang.* **2017**, *8*. [[CrossRef](#)]
24. Krysanova, V.; Donnelly, C.; Gelfan, A.; Gerten, D.; Arheimer, B.; Hattermann, F.; Kundzewicz, Z.W. How the performance of hydrological models relates to credibility of projections under climate change. *Hydrol. Sci. J.* **2018**, *63*, 696–720. [[CrossRef](#)]
25. Bourdin, D.R.; Stull, R.B. Bias-corrected short-range Member-to-Member ensemble forecasts of reservoir inflow. *J. Hydrol.* **2013**, *502*, 77–88. [[CrossRef](#)]
26. Zalachori, I.; Ramos, M.H.; Garçon, R.; Mathevet, T.; Gailhard, J. Statistical processing of forecasts for hydrological ensemble prediction: A comparative study of different bias correction strategies. *Adv. Sci. Res.* **2012**, *8*, 135–141. [[CrossRef](#)]
27. Pagano, T.C.; Wang, Q.J.; Hapuarachchi, P.; Robertson, D. A dual-pass error-correction technique for forecasting streamflow. *J. Hydrol.* **2011**, *405*, 367–381. [[CrossRef](#)]
28. Hashino, T.; Bradley, A.A.; Schwartz, S.S. Evaluation of bias correction methods for ensemble streamflow volume forecasts. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 939–950. [[CrossRef](#)]
29. Li, Y.; Jiang, Y.; Lei, X.; Tian, F.; Duan, H.; Lu, H. Comparison of precipitation and streamflow correcting for ensemble streamflow forecasts. *Water* **2018**, *10*, 177. [[CrossRef](#)]
30. Yuan, X.; Wood, E.F. Downscaling precipitation or bias-correcting streamflow? Some implications for coupled general circulation model (CGCM)-based ensemble seasonal hydrologic forecast. *Water Resour. Res.* **2012**, *48*, 1–7. [[CrossRef](#)]
31. Beven, K.J. A manifesto for the equifinality thesis. *J. Hydrol.* **2006**, *320*, 18–36. [[CrossRef](#)]
32. Delignette-Muller, M.L.; Dutang, C. Fitdistrplus: An R Package for Fitting Distributions. *J. Stat. Softw.* **2015**, *64*, 1–34. [[CrossRef](#)]
33. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
34. Gudmundsson, L.; Bremnes, J.B.; Haugen, J.E.; Engen-Skaugen, T. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations—A comparison of methods. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 3383–3390. [[CrossRef](#)]

35. Gudmundsson, L. Qmap: Statistical Transformations for Post-Processing Climate Model Output; R Package Version 1.0-4; 2016. Available online: <https://cran.r-project.org/> (accessed on 18 August 2020).
36. Zambrano-Bigiarini, M. hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series; R Package Version 0.3-10; 2017. Available online: <https://cran.r-project.org/> (accessed on 18 August 2020).
37. Maurer, E.P.; Wood, A.W.; Adam, J.C.; Lettenmaier, D.P.; Nijssen, B. A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States. *J. Clim.* **2002**, *15*, 3237–3251. [[CrossRef](#)]
38. 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](#)]
39. Brekke, L.; Wood, A.; Pruitt, T. Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs. 2014. Available online: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/ (accessed on 10 August 2020).
40. Bracken, C. Downscaled CMIP3 and CMIP5 Climate Projections—Addendum. 2016. Available online: https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/Downscaled_Climate_Projections_Addendum_Sept2016.pdf (accessed on 10 August 2020).
41. Wang, S.; Wang, Y. Improving probabilistic hydroclimatic projections through high-resolution convection-permitting climate modeling and Markov chain Monte Carlo simulations. *Clim. Dyn.* **2019**, *53*, 1613–1636. [[CrossRef](#)]
42. Beven, K.J. Preferential flows and travel time distributions: Defining adequate hypothesis tests for hydrological process models. *Hydrol. Process.* **2010**, *24*, 1537–1547. [[CrossRef](#)]
43. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. *Trans. ASABE* **2015**, *58*, 1763–1785. [[CrossRef](#)]
44. Ritter, A.; Muñoz-Carpena, R. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *J. Hydrol.* **2013**, *480*, 33–45. [[CrossRef](#)]
45. Legates, D.R.; McCabe, G.J. Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [[CrossRef](#)]
46. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
47. Walker, R.L.; Nicholson, R.S.; Storck, D.A. *Hydrologic Assessment of Three Drainage Basins in the Pinelands of Southern New Jersey, 2004-06*; Technical Report 2011-5056; U.S. Geological Survey Scientific Investigations: Reston, VA, USA 2011; 145p.
48. Shi, X.; Wood, A.W.; Lettenmaier, D.P. How Essential is Hydrologic Model Calibration to Seasonal Streamflow Forecasting? *J. Hydrometeorol.* **2008**, *9*, 1350–1363. [[CrossRef](#)]
49. Clark, M.P.; Wilby, R.L.; Gutmann, E.D.; Vano, J.A.; Gangopadhyay, S.; Wood, A.W.; Fowler, H.J.; Prudhomme, C.; Arnold, J.R.; Brekke, L.D. Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Curr. Clim. Chang. Rep.* **2016**, *2*, 55–64. [[CrossRef](#)]
50. Wilby, R.L.; Harris, I. A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resour. Res.* **2006**, *42*. [[CrossRef](#)]
51. Poff, N.L.; Allan, J.D.; Bain, M.B.; Karr, J.R.; Prestegard, K.L.; Richter, B.D.; Sparks, R.E.; Stromberg, J.C. The Natural Flow Regime. *BioScience* **1997**, *47*, 769–784. [[CrossRef](#)]
52. Euser, T.; Winsemius, H.C.; Hrachowitz, M.; Fenicia, F.; Uhlenbrook, S.; Savenije, H.H. A framework to assess the realism of model structures using hydrological signatures. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1893–1912. [[CrossRef](#)]
53. Archfield, S.A.; Kennen, J.G.; Carlisle, D.M.; Wolock, D.M. An Objective and Parsimonious Approach for Classifying Natural Flow Regimes at a Continental Scale. *River Res. Appl.* **2014**, *30*, 1166–1183. [[CrossRef](#)]
54. Poff, N.L.; Zimmerman, J.K. Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. *Freshw. Biol.* **2010**, *55*, 194–205. [[CrossRef](#)]

