Projected Changes in Flood Peak Discharge across Iowa: A Flood Frequency Perspective

Final report

Project # 20-SPR2-002

March 2023

Sponsored by Iowa Department of Transportation and Federal Highway Administration

Disclaimers

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. The opinions, findings, and conclusions expressed in this publication are those of the author and not necessarily those of the Iowa Department of Transportation or the United States Department of Transportation, Federal Highway Administration.

The sponsors assume no liability for the contents or use of the information contained in this document. This report does not constitute a standard, specification, or regulation. The sponsors do not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objectives of the document.

Statement of Non-Discrimination

Federal and state laws prohibit employment and/or public accommodation discrimination on the basis of age, color, creed, disability, gender identity, national origin, pregnancy, race, religion, sex, sexual orientation or veteran’s status. If you believe you have been discriminated against, please contact the Iowa Civil Rights Commission at 800-457-4416 or Iowa Department of Transportation's affirmative action officer. If you need accommodations because of a disability to access the Iowa Department of Transportation’s services, contact the agency's affirmative action officer at 800-262-0003.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. i. 20-SPR2-002	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Projected Changes in Flood Peak Discharge Across Iowa: A Flood Frequency Perspective		**5. Report Date** March 2023 **6. Performing Organization Code**
7. Author(s) Gabriele Villarini and Witold F. Krajewski		**8. Performing Organization Report No.** If applicable, enter any/all unique numbers assigned to the performing organization.
9. Performing Organization Name and Address IIHR—Hydroscience & Engineering University of Iowa 100 C. Maxwell Stanley Hydraulics Laboratory Iowa City, Iowa 52242-1585		**10. Work Unit No.** **11. Contract or Grant No.**
12. Sponsoring Agency Name and Address i. Iowa Department of Transportation, 800 Lincoln Way, Ames, IA 50010 ii. Federal Highway Administration, 1200 New Jersey Avenue, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Final Report **14. Sponsoring Agency Code** CFDA/ALN 20.205
15. Supplementary Notes		

16. Abstract

Numerous modeling studies point to an intensification of the hydrological cycle under projected climate warming, with increasing frequency of extreme events, including heavy rainfall and flooding. So, what would projected changes in flooding mean for the Iowa Department of Transportation (IDOT) and the bridges and structures that constitute Iowa's highway system? How resilient are these highway structures to different climate warming scenarios?

Addressing these questions requires flood frequency analysis. The current methodology relies on the guidelines by Bulletin 17C. However, issues related to regionalization of at-site estimates as well as accounting for the projected changes in the climate system have received little attention despite the potentially large impacts, including to the IDOT's infrastructure.

Here the authors have focused on the examination of the projected changes in flooding across Iowa using the hydrologic model developed by the Iowa Flood Center (IFC). The focus is on high-resolution and downscaled outputs from CMIP5 (Fifth Coupled Model Intercomparison Project) and CMIP6, and different scenarios. The results point to an increase in flood hazard across much of state, especially for high emission scenarios and towards the end of the 21st century. Moreover, the authors have developed a web interface - the Iowa Flood Frequency and Projections tool (IFFP) - that provides projections of flooding to the end of the twenty first century at any river reaches in the State of Iowa.

17. Key Words Flood frequency; Iowa; projections; climate warming; flooding; hydrologic modeling; Web Interface	**18. Distribution Statement** No restrictions.

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 115	**22. Price** NA

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Projected Changes in Flood Peak Discharge across Iowa: A Flood Frequency Perspective

Final report

March 2023

Principal Investigator: Gabriele Villarini, Professor
Civil and Environmental Engineering, University of Iowa

Co-Principal Investigators: Witold Krajewski, Professor
Civil and Environmental Engineering, University of Iowa

Felipe Quintero, Assistant Research Scientist
IIHR—Hydroscience & Engineering

Research Assistant: Alexander Michalek
Civil and Environmental Engineering, University of Iowa

Sponsored by the Iowa Department of Transportation and Federal Highway Administration

Table of Contents

Executive Summary

Numerous modeling studies point to an intensification of the hydrological cycle under projected climate warming, with increasing frequency of extreme events, including heavy rainfall and flooding. Recently, the occurrence of extreme flooding has become the norm rather than the exception, with the 2008 Eastern Iowa flood representing the “poster child” for this catastrophic situation: during this event, for instance, the eastern half of our state experienced the closure of a number of roads, including Interstate 80. So, what would projected changes in flooding mean for the Iowa Department of Transportation (IDOT) and the bridges and structures that constitute Iowa’s highway system? How resilient are these highway structures to different climate warming scenarios?

Addressing these questions requires flood frequency analysis. The current methodology relies on the guidelines by Bulletin 17C. However, issues related to regionalization of at-site estimates as well as accounting for the projected changes in the climate system have received little attention despite the potentially large impacts, including to the IDOT’s infrastructure.

Here the authors have focused on the examination of the projected changes in flooding across Iowa using the hydrologic model developed by the Iowa Flood Center (IFC). The focus is on high-resolution and downscaled outputs from CMIP5 (Fifth Coupled Model Intercomparison Project) and CMIP6 (Sixth Coupled Model Intercomparison Project), and different scenarios. The results point to an increase in flood hazard across much of state, especially for high emission scenarios and towards the end of the 21st century. Moreover, the authors have developed a web interface - the Iowa Flood Frequency and Projections tool (IFFP) - that provides projections of flooding to the end of the twenty first century at any river reaches in the State of Iowa.

1. Literature Review[1]

Surface water quantity presents a substantial risk to Iowa in the way of flooding. In the past few decades, floods (e.g., 1993, 2008, 2019) have devasted the state causing billions of dollars in economic damage as well as many human fatalities (e.g., Villarini and Slater 2017). Climate change is expected to exacerbate issues related to flooding as changes in extreme precipitation and temperature are likely to alter the hydrologic cycle. Changes to the hydrologic cycle need to be understood so that adequate water resources management and design of engineering structures for flood control can be implemented to help prepare Iowa for the future. For this reason, it is important to understand historical patterns in surface water quantity, expected future changes due to climate change, and what knowledge gaps are still present. In this section we assess surface water quantity conditions, primarily related to flooding and the drivers behind observed and expected changes. We primarily focus on two aspects of surface water quantity in runoff and streamflow. To provide a distinction, runoff is the overland flow of precipitated water across the landscape, and streamflow is the flow of water occurring within the river channel.

1.1 Historical Changes

Historically, much of the focus in the literature is on the use of observations at stream gauge locations to understand trends in magnitude of flood quantiles at the annual scale. The primary approach is to use annual maximum discharge time series (i.e., largest discharge value per year) to understand trends, as it is a simple indicator for flooding. For instance, multiple studies (e.g., Lettenmaier et al. 1994; Lins and Slack 1999; Douglas et al. 2000; McCabe and Wolock 2002; Lins and Slack 2005) apply the Mann-Kendall (MK) test to observations across the United States to detect monotonic patterns in different flood quantiles. For Iowa, they analyze approximately 20 stream gauges with over 30 years of data, and examine different percentiles (e.g., annual minimum, 10^{th}, 20^{th} ,30^{th}, 40^{th}, median, 60^{th}, 70^{th}, 80^{th}, 90^{th}, mean, and maximum daily discharge). A common pattern among these studies is that the detection of trends in magnitude is most prevalent in lower discharge quantiles with increasing trends. Figure 1-1 highlights how this pattern is strong not only for Iowa, but across the entire United States. The studies highlight how trends in annual maximum flood peaks are much more difficult to detect.

In recent years there have been many studies about the changes in the magnitude of annual maximum flood peaks. The methods typically assume that the annual maximum flood peaks are stationary (i.e., not changing over time), highlighting the importance of examining trends in their magnitudes. Villarini et al. (2011a) examine trends in annual flood peaks at gauges with 75 years of record for the Midwest (16 in Iowa). The study uses change-point analysis in the mean, finding only three gauges that have a significant change in the first moment (two decreases, one increase). For the variance, a change-point analysis reveals four sites (one site) with an increase (decrease) in variance. The year of the change point varies, ranging from 1940 to 1980. The MK test identifies two gauges with a significant trend, one negative and one positive. Furthermore, Villarini et al. (2009) performed an in-depth analyses for a single site along the Cedar River to detect the stationarity of annual flood peaks with multiple non-stationary modeling methods, not detecting significant trends for any of the methods. Studies such as these highlight the difficult nature of trend detection for annual maximum flows. However, even if detection is limited, it should not be viewed as evidence of a lack of trends (Villarini and Wasko 2021).

Slater et al. (2021) provide a global examination of changes in the magnitude of 1%, 2%, and 5% annual exceedance probability (AEP) floods at gauges using 1970 as reference year. They fit stationary and nonstationary models to annual peak flows and select the best fitting one at each site to assess the discharge values with a given AEP. They focus on the sites with nonstationary models as the best fit, and compare their results to values from the 1970s. For Iowa, the study shows that, whenever the authors

[1] Adapted from draft chapter of the First Iowa Climate Assessment.

detect a change, the percent change in 1%, 2%, and 5% AEP flood magnitude is over 20% or increasing for most of the area. These sites are the same as previous studies where trend analysis does not show change. The discrepancies could be due to the longer period of record in Slater et al. (2021) and/or the uses of different methods.

An important aspect related to annual maximum flood peaks is the spatial variability of changes in magnitude. For instance, Archfield et al. (2016) examine the multidimensional behavior of flood changes for stream gauges across the United States. The authors use grid cell analysis to examine trends in frequency, duration, peak magnitude, and volume; in their study, Iowa is within four cells. Of the four cells, only one shows an increasing trend in magnitude, while the remaining three do not show a significant trend. Conversely, Wasko et al. (2021) analyze trends in historical flood peaks across the globe by utilizing gauging stations with 30 years of record. The authors apply linear regression to determine the percent change over ten-year periods and analyze the median trend of all stations within 5˚x 3˚ grid cells. For the grid cells, they show that trends in the 50^{th} and 95^{th} peak flow percentiles have percent changes per ten-years ranging from 6% to 12% and -6 to 6%, respectively. The coarse nature of their analysis provides an overview of historical changes but lacks clear information for changes in flood peak magnitude, especially at the community level.

Another important characteristic of annual maximum flood peaks is the frequency of events. The detection of significant trends for frequency of floods in Iowa is overall much more frequent in comparison to magnitudes. For example, Mallakpour and Villarini (2015) examine the frequency of annual maximum daily floods from 1962 to 2011 using a peaks-over-threshold (POT) approach. The analysis reveals over 20 gauges with a significant trend in the frequency of flood events (Fig. 1-2). Furthermore, Slater and Villarini (2016) analyze the frequency of floods based on the National Weather Service (NWS) flood levels with a Poisson regression to examine trends. Their frequency analysis (Fig. 1-3) shows that the gauges mainly exhibit increasing trends in the number of annual days above the action and minor stage flood level. However, trends in days above moderate and major flood stages are much less prevalent. From a seasonal perspective, the analysis of trends in flood counts by Neri et al. (2019) shows that there are increasing trends for the spring, summer, and fall seasons. Overall, the existing literature suggests that the frequency of flood events is increasing.

The changes in characteristics of flood peak observations are likely driven by increases in precipitation and temperature, especially at the monthly scale. For instance, Lettenmaier et al. (1994) analyze hydro-climatological trends across the United States. For Iowa gauges, they find significant positive trends in streamflow and precipitation. At the monthly scale, significant positive trends in monthly peak streamflow are present for the months of January through April as well as December. Increases in winter streamflow are likely due to increases in temperature during these months, with magnitudes ranging from 2-6°F. Furthermore, Small et al. (2006) examine trends in precipitation and 7-day low flows, revealing trends exist in low flow and fall precipitation. Warm and transition season precipitation has increasing trends matching those in low-medium flow quantiles of streamflow (Lins and Slack 1999, 2005). Villarini et al. (2013) support these previous findings with a POT analysis that shows heavy rainfall and temperature have increasing trends (Fig. 1-4). From a frequency perspective, Neri et al. (2019) show that the seasonal count of flood events for Iowa gauges can be predicted with regression models containing precipitation, temperature, and/or antecedent wetness conditions (i.e., a simple proxy for soil saturation). These findings highlight how changes in climate variables translate to changes in flooding.

Increases in runoff are likely due to changes in precipitation especially for March and May. Hodgkins and Dudley (2006) investigate changes in the timing of winter-spring streamflow for the northeastern United States from 1913 to 2002, using United States Geological Survey (USGS) gauges (26 for Iowa). Only one site shows a significant change in the winter-spring center of volume dates, showing a shift to later in the year. For most of the sites, the timing of center volumes shifts between 10 to 20 days later in the winter-spring season. Furthermore, runoff in March and May is increasing for nearly all (24) gauges (Fig.

1-5), with changes typically from 1-2.5 cm. These results match patterns found in the previously mentioned studies for rainfall. The increases in runoff are facilitating increases in flood events throughout the year, as well as increasing low-medium flow magnitudes as more water is making it into streamways.

Finally, we can detect the impacts of climate change on flooding in historical events. For example, Villarini et al. (2020) quantify the impact of anthropogenic forcing on the 2008 flood in Cedar Rapids. The study forces the hydrologic model by the Iowa Flood Center (IFC) using precipitation with and without external forcing (e.g., greenhouse gasses, aerosols). The results show that anthropogenic warming leads to a bigger flood event than without it, with an increase in stage by over 2 meters and a 1.28-fold increase in flood extent. Furthermore, the results show that anthropogenic forcing causes an increase in the number of affected buildings and damage costs. This study represents one of the very few analyses that provide commentary on the impacts and risk of climate change and flooding. At a coarser scale, Merz et al. (2021) supports these findings as they show increasing trends in the amount of people impacted by disastrous floods from 1977-2019 (Fig. 1-6).

1.2 Projected changes

Projected changes in surface water quantity are important to provide basic information towards improving Iowa's preparation and mitigation efforts against flooding. Overall, the literature on flood projections is more limited compared to the analyses of historical data described above. Most of the studies focus on specifics basins, but still provide valuable information for understanding possible future outcomes and impacts.

There are two studies which utilize projections for models' part of the Coupled Model Intercomparison Project phase 3 (CMIP3) to examine changes in streamflow at three Iowa basins. First, Hay et al. (2011) analyze watershed scale response to climate change for the Clear Creek watershed. The authors utilize a hydrologic model and force it with inputs from five global climate models (GCMs) and three CMIP3 emission scenarios. The results show that the mean daily streamflow is projected to decrease by the end of the century from 2.5 to 2.0 m^3/s on average across the different models and emission scenarios. Furthermore, Quintero et al. (2018) assess changes in flood frequency due to climate change by forcing the IFC model with CMIP3 projections. The study examines changes in the 1% AEP flood for the basins of the Cedar River and Skunk River. They find that the 1% AEP flood is projected to increase for two basins, with a range of increase from 25% to 52% over historical values. A brief comparison of these two studies indicates that annual flood peaks could increase, while average streamflow decreases.

Updates to climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5; Taylor et al. 2012) have led to more studies about projections of surface water quantity. Most of the available studies focus on Relative Concentration Pathways 4.5 (RCP4.5) and 8.5 (RCP8.5), where RCP8.5 is a business-as-usual emissions scenario. The literature (e.g., Demaria et al. 2016a; Demaria et al. 2016b; Naz et al. 2016; Byun et al. 2019) with CMIP5 focuses on climate projections using the Variable Infiltration Capacity (VIC) model and on results at the basin outlets.

The study by Demaria et al. (2016a) is one the first hydrologic investigations for projected changes with CMIP5. They focus on the projected changes in streamflow characteristics for the Northwest and Midwest United States at 124 basins with USGS gauges. For Iowa, the study provides information for three basins (Fig. 1-7). At the three sites, they do not detect significant trends in annual maximum cumulative 5-day precipitation in the historical period (1951-2005). For future projections (i.e., 2028-2082), two of the three sites show a positive trend in precipitation, while the third site shows a negative trend. However, these trends in precipitation do not translate to comparable changes in streamflow. The results indicate that mean base flows for the historical period are increasing for the three sites but decreasing under each of the future emission scenarios. For 7-day low flows, the three sites show results of positive and negative trends under the historical and RCP4.5 scenario. However, significant decreasing trends are shown under

RCP8.5. Furthermore, for 3-day peak flows the study shows a positive trend for the sites in the historical period, decreasing trends under RCP4.5, and a mixture of results under RCP8.5. Percent change in the magnitude of the 1% AEP 3-day peak flows shows increases between 20-30% under RCP8.5 for two sites, while the other site shows a decrease by 10%. For the 7-day low flows, a decrease is projected for all scenarios. The projected changes are attributed to increases in evapotranspiration, lowering water levels during warm seasons, and site-specific patterns in precipitation change. Demaria et al. (2016b) analyze the same sites for the impact due to snowpack on hydrology due to climate change and find increases in the magnitude of spring streamflow under RCP4.5 and RCP8.5 (Fig. 1-8). The changes are attributed to the decreases in snow cover and more precipitation during these periods.

Byun et al. (2019) find similar results to the previous studies (e.g., Hay et al. 2011; Demaria et al. 2016a; Demaria et al. 2016b; Quintero et al. 2018). They focus on 20 watersheds throughout the region, with four watersheds in Iowa. These watersheds consist of the Des Moines, the Iowa River, the Little Sioux, and the Skunk Rivers. For the four sites, the results show monthly streamflow is projected to increase in the spring and summer months under RCP8.5 in comparison with the historical period. Furthermore, on average under RCP8.5, the 1% AEP peak daily streamflow magnitude is projected to increase (5-20%) by 2050 for the four basins. A decrease or no change in magnitude is projected under RCP4.5. For changes in the 10% AEP 7-day low daily streamflow, increases in magnitude are projected under RCP4.5 and RCP8.5 beginning in the 2020s. The increases in peak daily and low flow conditions are attributed to projected precipitation increases during the winter and spring, changes in snow water equivalent (SWE), and changes in soil moisture. Loss of snow in the winter is one of the largest changes projected, with loss of annual SWE close to 50% under RCP8.5 (Fig. 1-9).

Runoff across Iowa is projected to be broadly impacted by climate change. Naz et al. (2016) study the impacts of climate change on the regional hydrologic response for the United States based on HUC2 regions. Iowa is within regions 7 (Upper Mississippi) and 10 (Missouri). They examine the median percent changes in seasonal total, high (95^{th}), and low (5^{th}) runoff between the periods of 1966-2005 and 2011-2050 (Fig. 1-10). For the regions, total runoff is expected to increase for all seasons, with large changes expected in December-January-February (DJF) and March-April-May (MAM), with increases ranging from 11.68% to 16.28%, respectively. For the runoff extremes, high runoff is expected to increase by ~5-10%. Furthermore, the 7-day low runoff is expected to increase from 5% to 20% but is shown to be more variable in space compared to high runoff. Villarini and Zhang (2020) present similar findings. The authors analyze runoff from the newest available projections from Coupled Model Intercomparison Project phase 6 (CMIP6). They show that projected changes in annual maximum daily discharge are variable among GCMs. Projected changes for 2015-2049 are expected to decrease for five of the eight models, with mixed results for the other three. For the period of 2065-2099, three models point to an increase in annual maximum daily runoff, three indicate a decrease, and the rest show a mixture of results for Iowa. Giuntoli et al. (2018) analyze the sources of uncertainty in projected runoff indicating that GCMs and global impact models (GIM) are the largest sources of uncertainty compared to the selected scenarios and climate variability. Models part of both CMIP5 and CMIP6 highlight uncertainty in runoff projections, and do not provide irrefutable evidence of an unequivocable pattern of change in flood characteristics for Iowa.

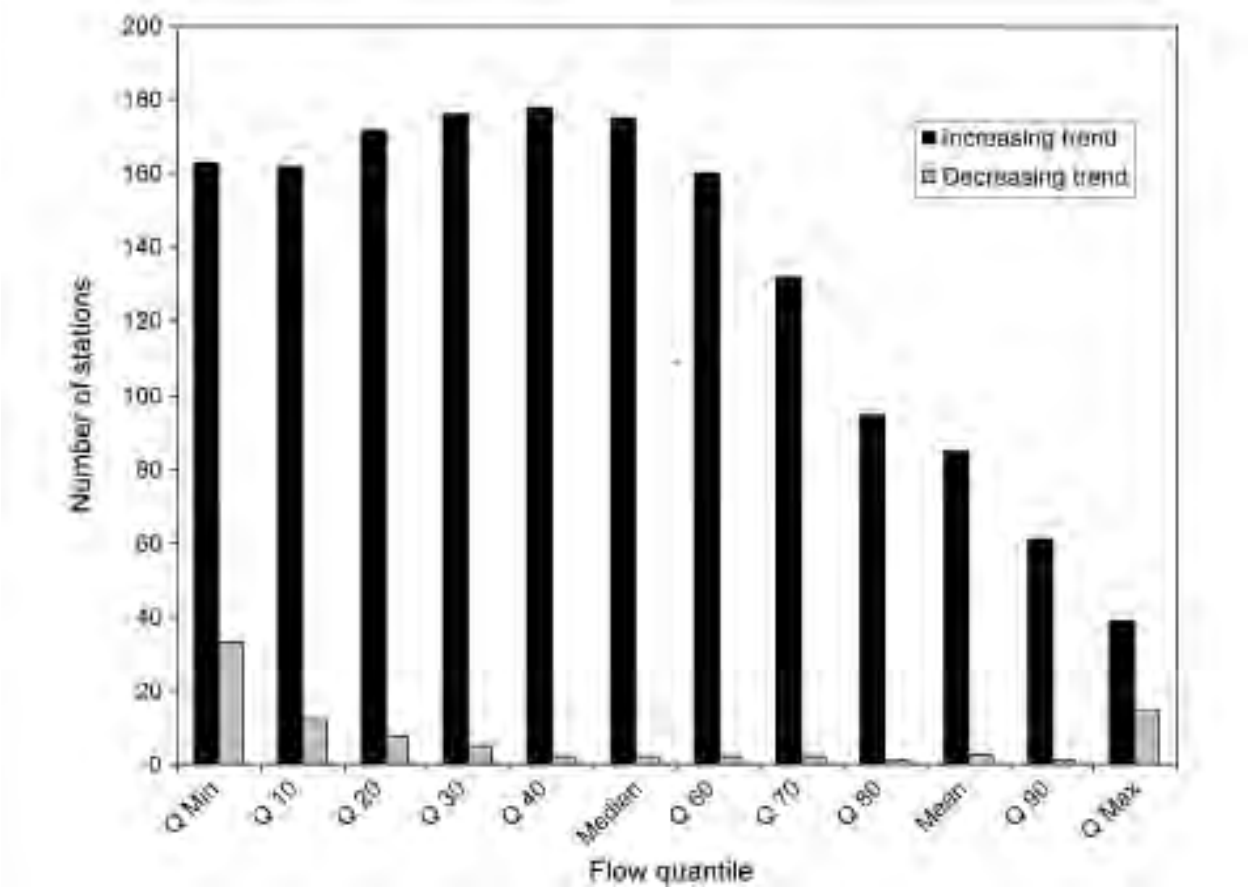

Figure 1-1. Number of U.S. stream gauges (out of 435) from 1940 to 1999 with statistically significant trends across different discharge quantiles (Lins and Slack 2005).

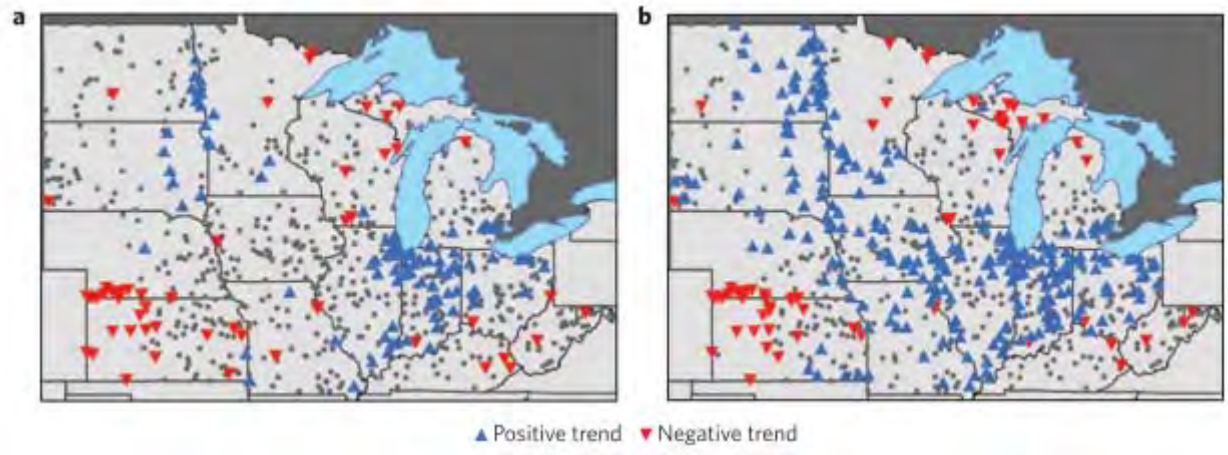

Figure 1-2. Trends in magnitude (a) and frequency (b) of flood events at the annual scale from Mallakpour and Villarini (2015).

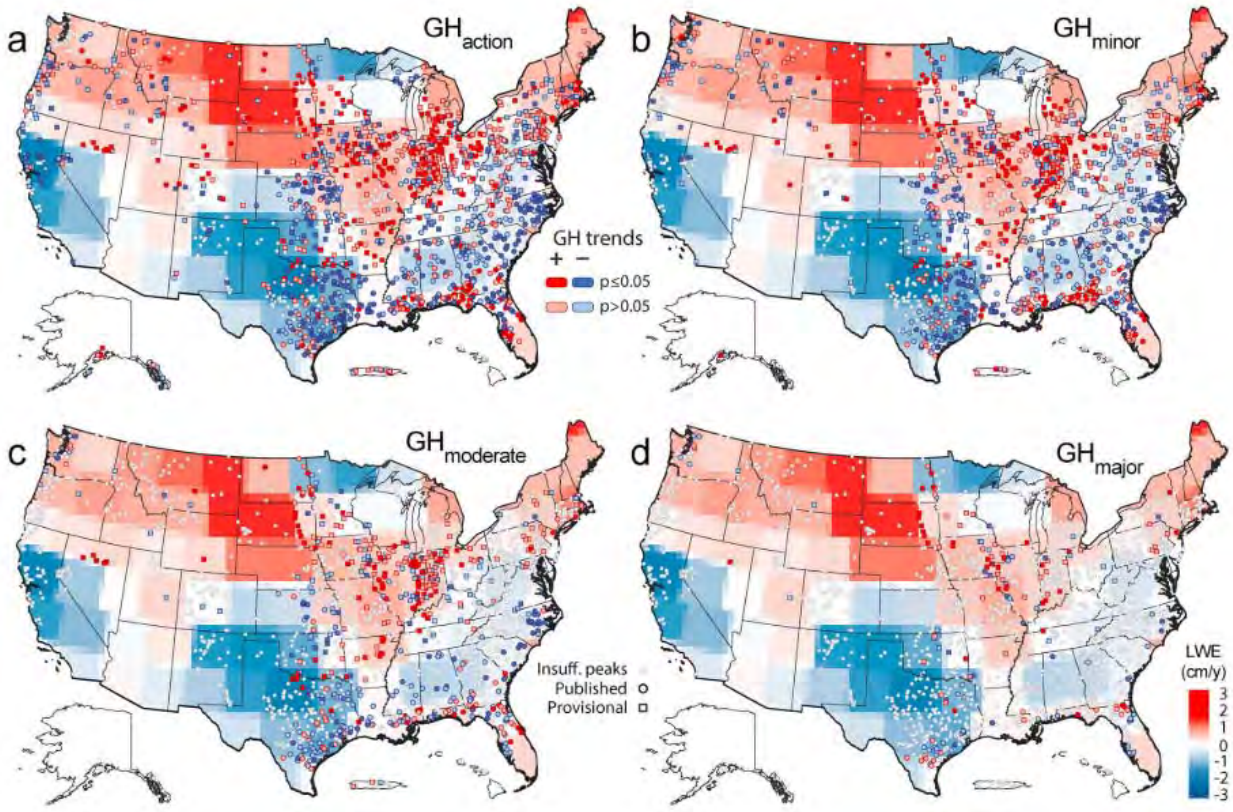

Figure 1-3. Trends in the number of days above NWS flood categories with liquid water equivalent (LWE) plotted behind (Slater and Villarini 2016).

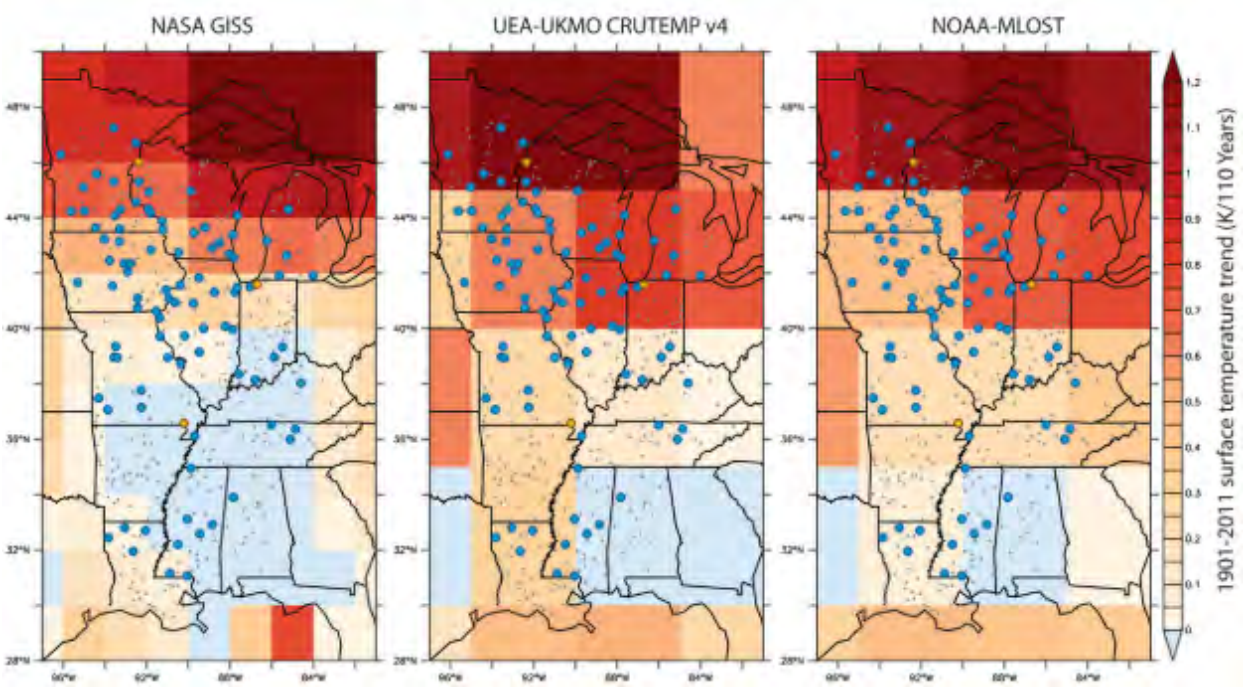

Figure 1-4. Rain gauges with statistically significant increasing (decreasing) trends represented by blue (orange) circle and trends in average temperature from March to October 1901-2011 for three datasets (Villarini et al. 2013).

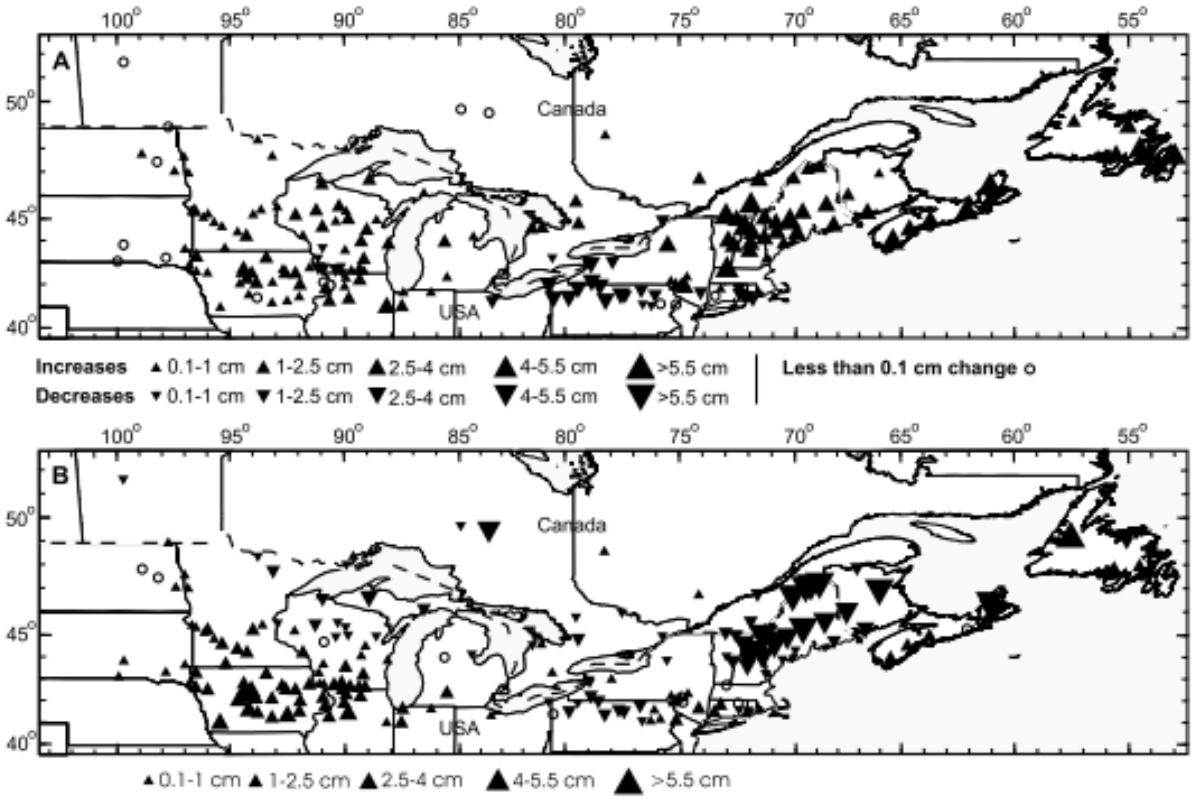

Figure 1-5. Magnitude and direction of changes in March (a) and May (b) runoff during 1953-2020 from Hodgkins and Dudley (2006).

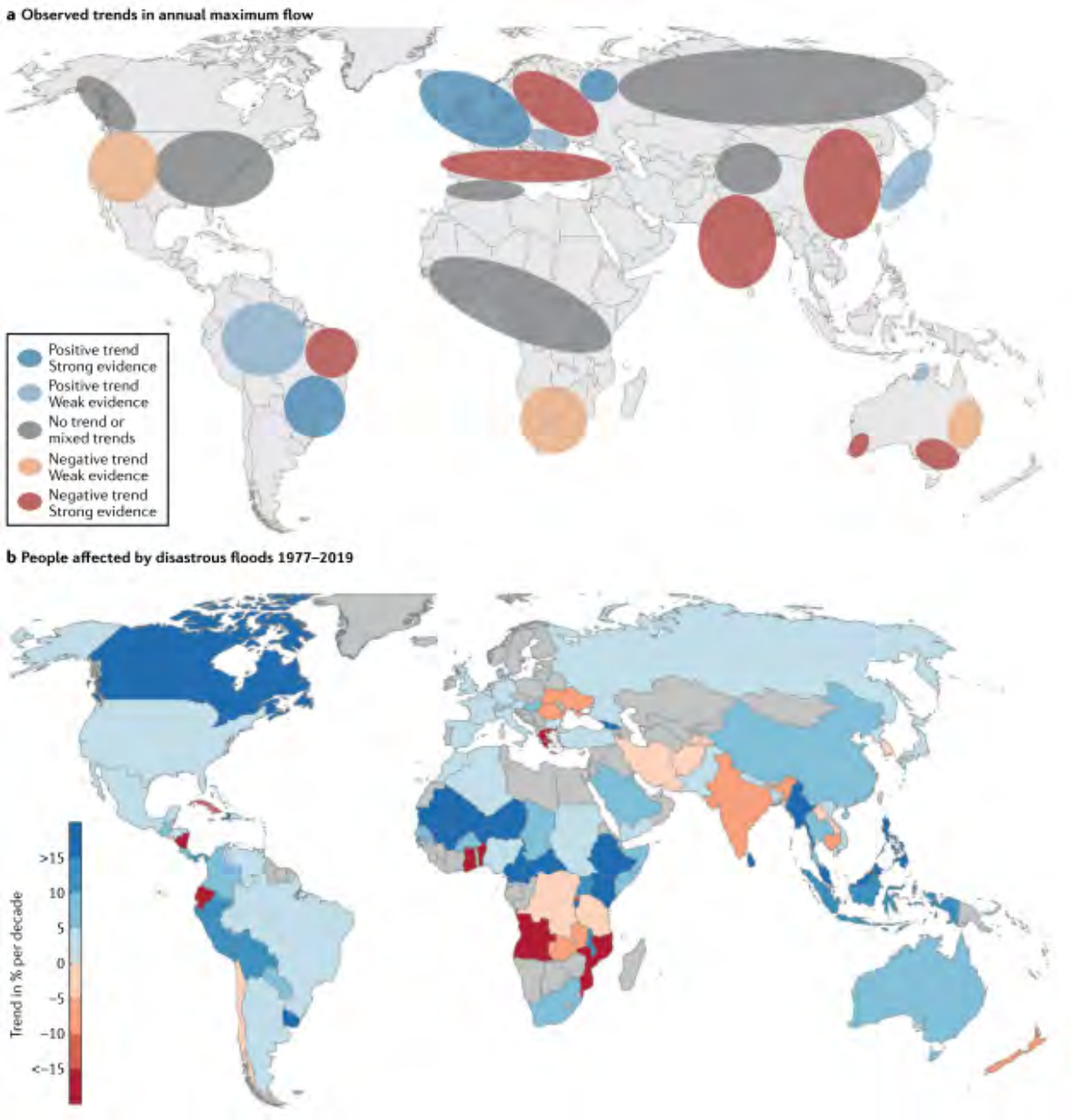

Figure 1-6. Observed trends for (a) annual maximum flow based on large-scale trend studies and trends in (b) people affected by floods from 1977-2019 (Merz et al. 2021).

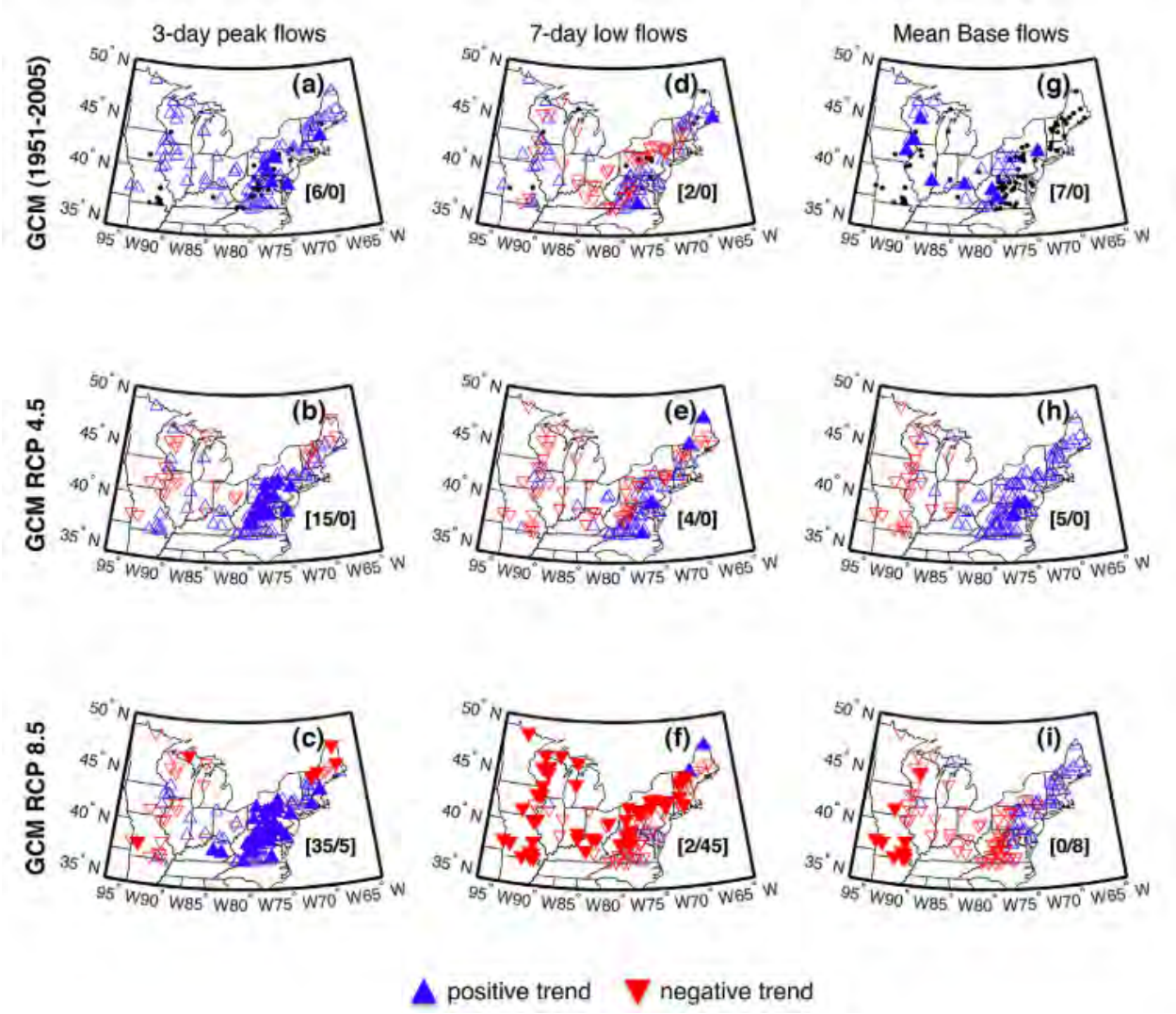

Figure 1-7. Trends in simulated 3-day peak flows, 7-day low flows, and mean base flows derived from CMIP5 GCMs for historical (1951-2005) and future (2028-2082) periods from Demaria et al. (2016a).

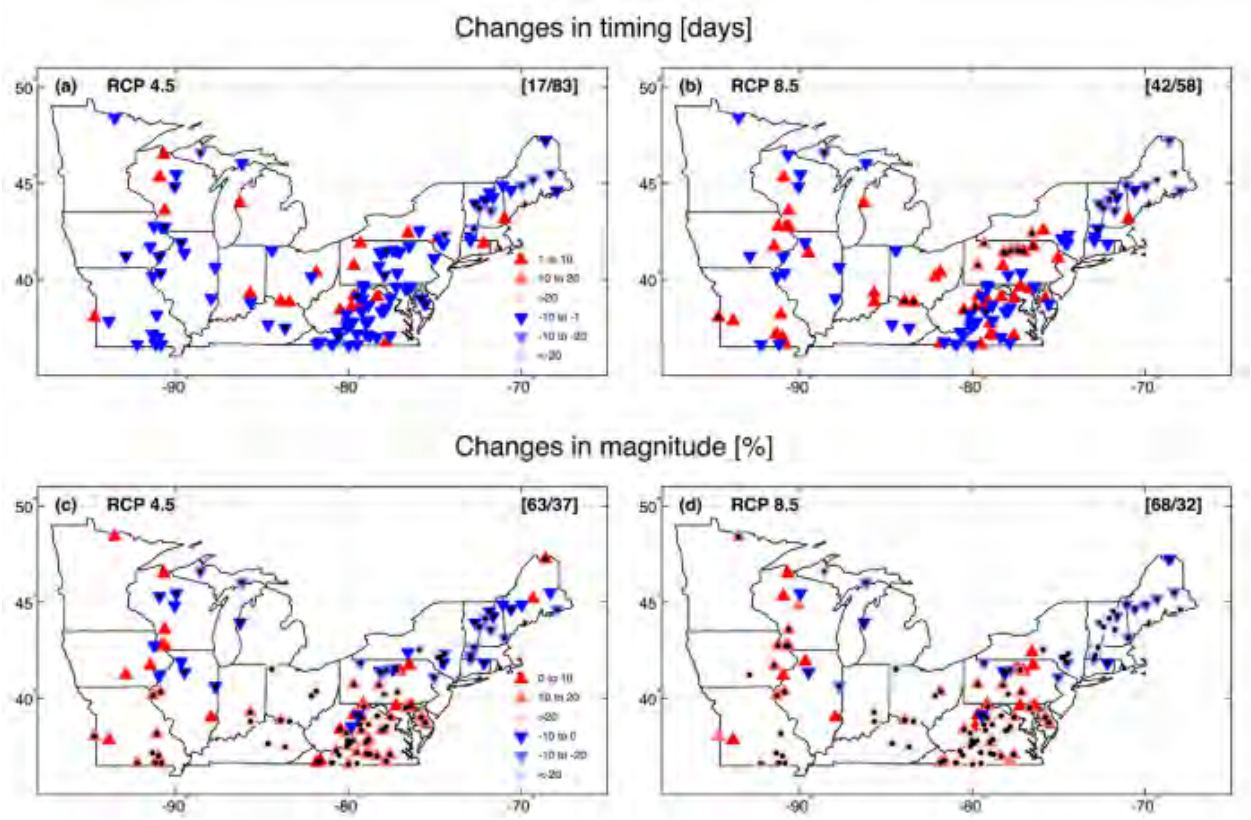

Figure 1-8. Changes in the timing (a,b) and magnitude (c,d) of maximum spring streamflow between historical (2014-2095) and future (1951-2005) periods from Demaria et al. (2016b).

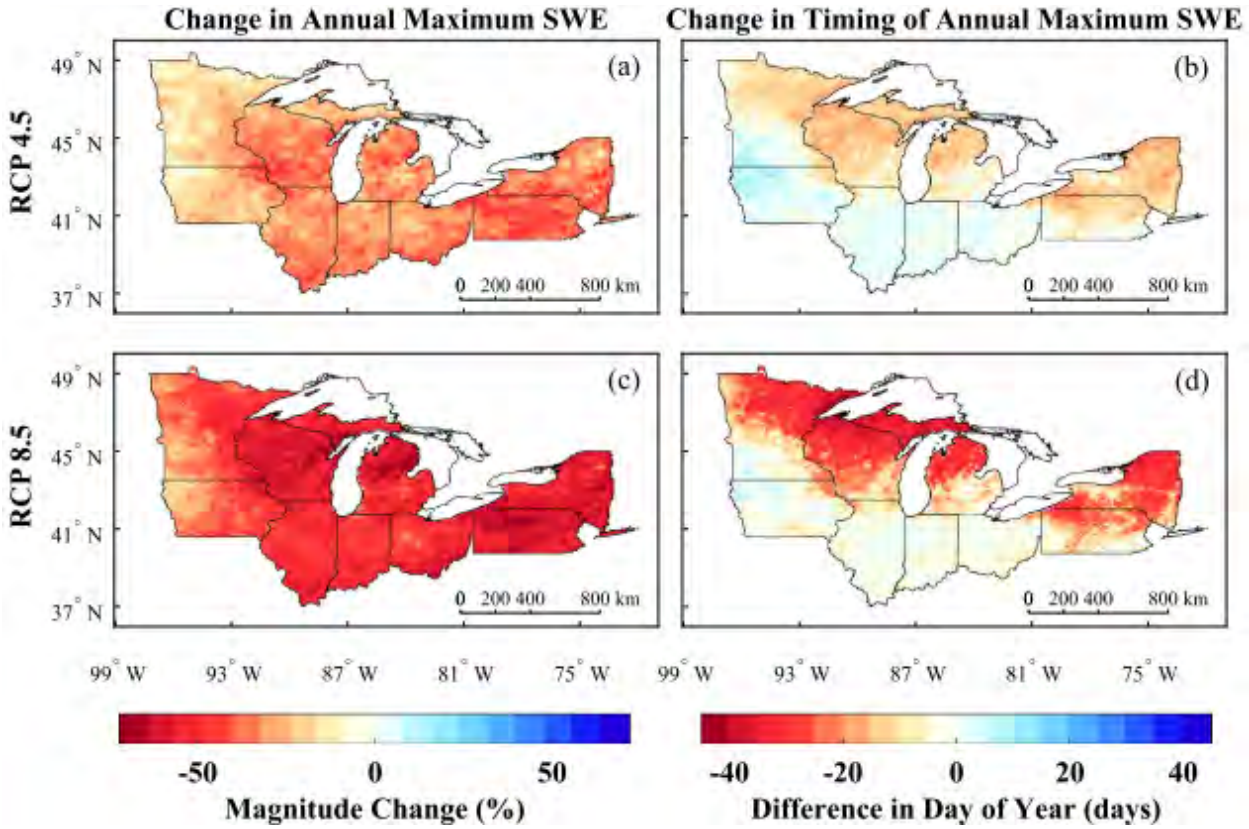

Figure 1-9. Changes in magnitude and time of annual SWE by ensemble mean based on RCP4.5 (upper) and RCP8.5 (lower) for 2071-2100 from Byun et al. (2019).

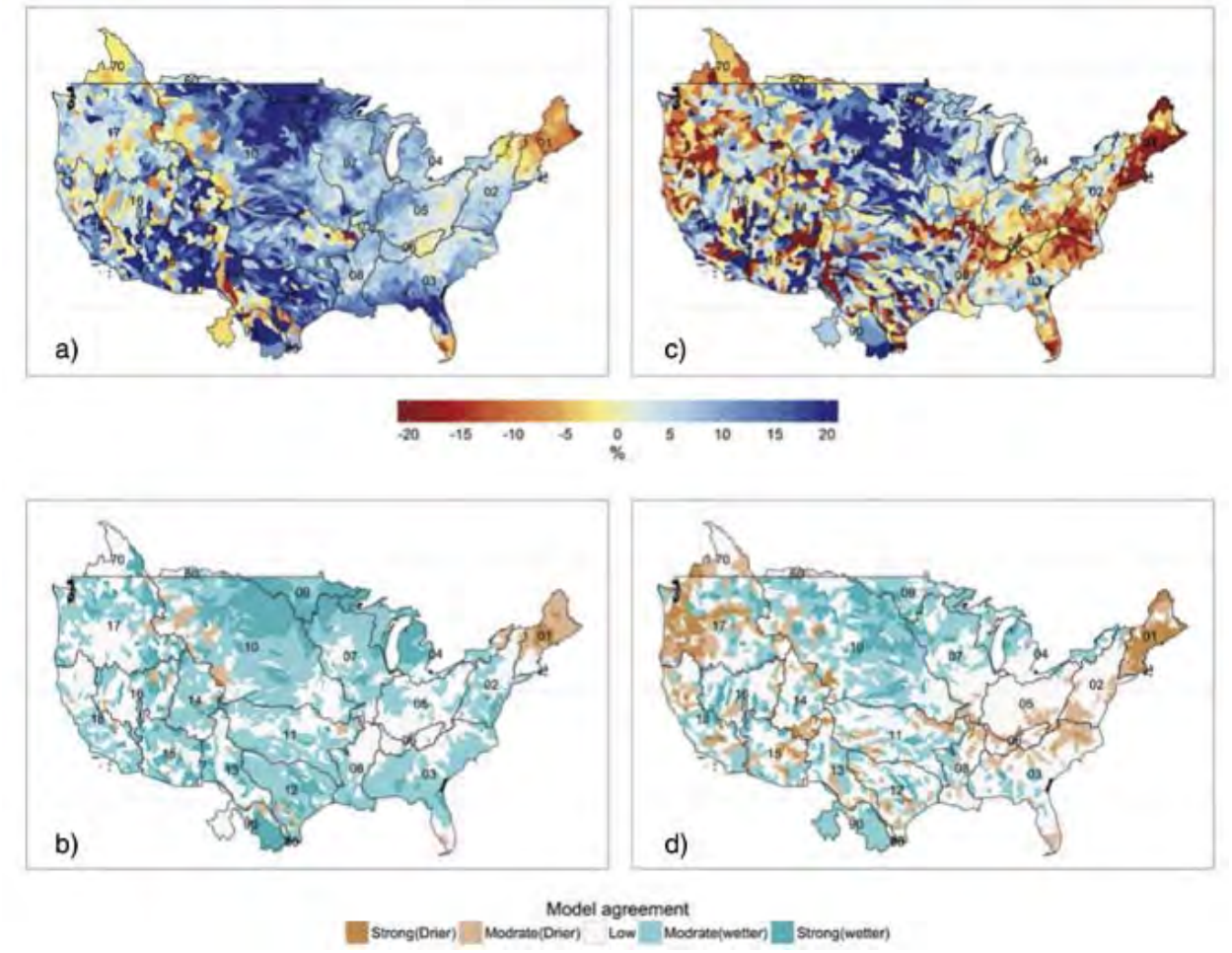

Figure 1-10. Median ensemble percentage of change on (a) 95th percentile daily runoff and (b) 7-day low runoff. Number of models predicting increase in (c) high runoff and (d) low runoff (Naz et al. 2016).

2. Advantages of Physically-based Flood Frequency Analysis with Long Term Simulations for Iowa[2]

2.1 Introduction

In the United States, a primary approach to estimate flood quantiles is to perform statistical analyses on discharge data from gauged locations following national guidelines. The guidelines, developed by the United States Geological Survey (USGS), such as Bulletin 17C (B17C) (England Jr et al. 2019), employ the log-Pearson type III distribution to determine annual exceedance probability (AEP) discharges from recorded data (Dawdy et al. 2012). However, complimentary statistical approaches (i.e., regional regression models) have been developed to aid in providing localized estimates for locations without gauge records. For ungauged locations, the development of regional flood quantile estimation equations has been a focal point since 1868 when O'Connell (O'Connell 1868) discovered a scaling power-law relationship between discharge and drainage area. This is commonly referred to as scaling theory, which applies the characteristics of self-similarity in channel networks (Gupta et al. 2007). The USGS establishes regional flood quantile estimation equations at the state level with a regression framework based on gauged location data and scaling theory. The regional regression equations can include variables related to rainfall, landform properties, or catchment physical properties (e.g., shape or slope). A notable limitation of this approach is that explanatory variables are prioritized based on the reduction of regression error residuals instead of hydrologic justification (Klemes 1974; Perez et al. 2019).

To address the concerns associated with statistical methods, the identification of governing hydrologic parameters for regional flood quantile estimation has been a primary research focus. In the past, studies such as that by Eagleson (1972) have blended statistical approaches with kinematic wave functions to capture physical processes behind flood frequency dynamics. This approach focused on AEP discharges at the outlet of a catchment, and determined that catchment area, the area contributing to direct runoff, and rainfall are important parameters for capturing flood frequency dynamics (Eagleson 1972). Ayalew and Krajewski (2017) analyzed the impact of channel network geometry by comparing two watersheds with the same drainage area located near each other; the comparison indicated that network geometry significantly impacted the AEP discharges but that current USGS regional equations largely ignore this parameter (Ayalew and Krajewski 2017). Perez et al. (2019) examined the important parameters in regional regression with respect to scaling-law for a large catchment in Iowa. Their study determined that variability between peak flow and drainage area is due to the channel network structure and spatial variability of soil moisture. Currently, many USGS regional equations at ungauged locations do not incorporate climate and stream network parameters, highlighting the need for methods which incorporate these aspects.

In recent decades, there has been a focus on applying hydrologic models for predictions at ungauged basins. This focus was started in 2003 with the creation of the Prediction in Ungauged Basins (PUB) initiative from the International Association Hydrological Sciences (Sivapalan 2003) to address inadequate predictions at ungauged locations. A review of the literature from a decade of the initiative by Hrachowitz et al. (2013) highlighted that a movement from data and calibration driven modeling methods to methods based on physical processes allowed for improved estimates of streamflow at ungauged locations. For example, physically-based hydrologic models applied in three studies (Dornes et al. 2008; Fang et al. 2013; Fang et al. 2010) from the PUB initiative showed that the application of *a priori* model parameters with flexible structures are able to provide robust estimates of streamflow across catchment scales. Finally,

[2] Adapted from: Michalek, A., F. Quintero, G. Villarini, and W.F. Krajewski, Advantages of physically based flood frequency analysis with long-term simulations for Iowa, Journal of Hydrologic Engineering, 27(12), 1-11, 2022.

the PUB decade brought about an increase in openness towards different modeling approaches and highlights the importance of continued exploration at ungauged locations (Hrachowitz et al. 2013).

In Iowa, Iowa Flood Center (IFC) has developed a continuous rainfall-runoff model called the Hillslope-Link Model (HLM); thus, there is the ability to assess regional flood frequency through physically based modeling built on major hydrologic parameters without the need for regional equations. The HLM simulates hydrologic processes at the hillslope-channel link scale, with runoff modeled through soil moisture accounting (Krajewski et al. 2017). One of the primary challenges of modeling at ungauged locations is that calibration is difficult as streamflow and stage measurements are not available (Dawdy et al. 2012). The HLM can address this issue as the model is calibration-free, i.e., a configuration of parameters is determined a priori, applied for all model inputs, and not adjusted for specific basins (Quintero et al. 2020). The strong performance of the model to reproduce observed streamflow has been illustrated in several studies (Ayalew and Krajewski 2017; Ayalew et al. 2014; Cunha et al. 2012; ElSaadani and Krajewski 2017; Krajewski et al. 2017; Quintero et al. 2016; Quintero et al. 2020). Furthermore, the HLM has shown success in performing regional flood frequency analysis as illustrated by Perez et al. (2019), providing a tool to assess flood frequency through physical processes.

Here we use the HLM to conduct a physically-based flood frequency analysis for Iowa. It should be noted that for our approach we apply a catchment modeling philosophy where continuous simulated flows represent predictions if a gauge was present with typical flood frequency guidelines applied. As discussed in Ball (2022), this is one of two primary approaches, with the other one applying catchment models that transfer the probability of rainfall to predicted flows. All subsequent discussion within this study is based on this philosophy. Our intent is to show the viability of the HLM in place of standard practices for regional flood quantile estimation analysis by conducting a comparison with current USGS regression equations at ungauged locations across Iowa

We organize this chapter as follows: Sections 2.2 and 2.3 describes the study area and HLM modeling set up with climate data used to estimate annual maximum discharges at gauged and ungauged locations. In Section 2.3 we describe the procedures used to evaluate model performance at gauged locations, the determination of exceedance discharges from modeled annual maxima, and the evaluation of regional regression exceedance discharge estimates (Eash et al. 2013) with respect to model-based estimates at ungauged locations. Section 2.5 describes the results from statistical methods and is followed by Section 2.6, where we discuss the differences between flood quantile estimation approaches and the advantages of the HLM. Finally, we conclude with an overview of our main results and closing remarks.

2.2 Study Area and Data Sources

We focus on the state of Iowa and its basins which drain into the Mississippi and Missouri Rivers. Iowa is agriculturally dominated with low relief, fertile soils, and high drainage density (Prior 1991). Its climate is seasonally variable with significant rainfall occurring during the summer. The combination of landscape characteristics and significant summer rainfall causes the state to be prone to floods (Baeck et al. 2013; Mallakpour and Villarini 2015; Villarini et al. 2011a,b; Villarini et al. 2013). Fig. 2-1 shows the three flood regions determined by Eash et al. (2013) that the authors consider as homogenous based on landform characteristics for Iowa. Each flood region was used to develop region specific regression equations for flood quantile estimates at ungauged locations. A feature of interest for regional flood quantile estimation is the Region 1, which is mainly hydrologically disconnected (Des Moines Lobe), especially for low flow regime.

We use annual maxima and daily-averaged discharges from 150 USGS stream gauges that were active within the 1981-2020 period. The stations are evenly distributed across Iowa (green circles in Fig. 2-1), with drainage areas ranging from ~10 to 36,000 km^2. For prediction of annual exceedance discharges at ungauged locations, we focused on 445 river communities uniformly distributed across the state (blue

triangles in Fig. 2-1). At these locations, we performed regional flood quantile estimation using USGS Streamstats web-based application (USGS 2021). The input data required for the regional estimates consisted of watershed, landform, and climate characteristics based on flood regions as described in Eash et al. (2010). These parameters are precalculated by the StreamStats web tool based on the watershed selected. As a quality control, we only used ungauged locations in which the watershed area from StreamStats matched the watershed area from the HLM within ±5%.

2.3 HLM Setup

The HLM was developed from scaling theory for river networks and water transport (Gupta et al. 2007). The model consists of a large system of ordinary differential equations following river network topology, solved with parallel implementation of the Runge-Kutta method. The model does not consider soil type or region-specific parameters and finds the proper configuration based on the spatial distribution of model parameters. HLM simulations are conducted at every channel link along a river network derived from a 90-m DEM. This results in a river network with 420,000 pairs of hillslopes and channel links, with an average hillslope area of 0.4 km^2 and an average channel length of 700 m (Quintero and Krajewski 2018; Quintero et al. 2020). For the experiments proposed in this chapter, we changed the HLM structure slightly compared to the configuration used in Quintero et al. (2020): instead of using a constant value of the parameter controlling the available water storage in the upper layer of soil, we made the parameter variable in time, reflecting the changes in the seasonal soil storage capacity, during colder and warmer months. Please refer to Krajewski et al. (2017), Quintero et al. (2020), and Mantilla et al. (2022) for in-depth discussion on the HLM.

In this study, we applied the HLM to simulate discharge at an hourly temporal resolution at every channel in Iowa from 1981 to 2020. The inputs required to generate hourly discharge are precipitation and evapotranspiration data. We used an hourly precipitation with a 4-km spatial resolution, hereafter referred to as PRISM-NLDAS (described below), and monthly evapotranspiration (ET) inputs taken from the NLDAS climatology (Mitchell 2004) spatially averaged across Iowa. From the simulations, we determined the simulated annual maximum discharges for every year in the simulation period (1981 to 2020).

For small catchments (<200 km^2), it has been determined that basin response is sensitive to the temporal resolution of precipitation as sub-hourly dynamics control (e.g., Berne et al. 2004; Fiche et al. 2016; Gabellani et al. 2007; Kandel et al. 2005; Krajewski et al. 1991; Nikolopoulos et al. 2011; Zoccatelli et al. 2015). We recognize that sub-hourly precipitation data could improve model performance but data availability for this study area was limited. Finally, the use of hourly StageIV precipitation data with the HLM has been shown to work well in predicting peaks across catchment sizes found in Iowa as shown in Zhu et al. (2018), Quintero et al. (2020), and Ghimire et al. (2022).

Previously, Quintero et al. (2020) evaluated model performance of the HLM using StageIV rainfall input with success. However, StageIV rainfall is only available starting from January 1st, 2002 (Lin and Mitchell 2005), limiting its utility for flood frequency analysis. For this reason, we created a new hourly rainfall product (PRISM-NLDAS) to expand the study period to 1981. We focused on the creation of a long-term hourly precipitation data set as it is a temporal resolution used in many hydrologic studies (Praskievicz and Bartlein 2014; Santhi et al. 2008; Schaake et al. 2004). For showing the viability of the PRISM-NLDAS as a precipitation input, we used hourly precipitation and monthly evapotranspiration from StageIV to simulate hourly discharge with the HLM across Iowa from 2002 to 2020.

We created PRISM-NLDAS using North American Land Data Assimilation System (NLDAS) and Parameter elevation Regression on Independent Slopes Model (PRISM) datasets from 1981 to 2020. To achieve this, we used PRISM daily precipitation sets and temporally downscaled to hourly resolution based on NLDAS hourly precipitation. We used PRISM precipitation as the base for downscaling as it has a high spatial resolution with a 4-km grid size and is based on observations using a weighted regression analysis

(Daly et al. 2002). We chose NLDAS to create the hourly distribution of PRISM precipitation as it is developed with a focus on temporal resolution using four primary datasets (Climate Prediction Center data, Weather Surveillance Radar 88 Doppler radar estimates, CPC MORPHing technique rainfall data, and North American Regional Reanalysis simulations) based on temporal disaggregation (Mitchell 2004).

The temporal downscaling method applied a simple way of redistributing the daily precipitation data from PRISM based on the hourly distribution of NLDAS. For each PRISM cell, we searched for the nearest NLDAS cell, created a ratio of the daily rainfall totals (PRISM total to NLDAS total), and multiplied the ratio by the hourly distribution from NLDAS to obtain a PRISM hourly distribution. The goal was to preserve the daily precipitation totals from PRISM. We performed these steps for each PRISM cell across the study area from 1981 to 2020. Before the downscaling process, we spatially smoothed the NLDAS data at each hour using a 15-by-15 cell size moving-window to remove the spatial influence from the 12-km NLDAS grid on the output PRISM 4-km grid.

2.4 Statistical Methods

We compared the distribution of observed and modeled annual maximum discharge values at 119 streamgauges with a 30+-year period using the two-sample Kolmogorov–Smirnov (KS) test (Massey 1951) to determine if the model is appropriate for flood frequency analysis. We set two significance levels of 0.10 and 0.05 and applied a Bonferroni correction to account for multiple hypotheses testing (Bonferroni 1936).

Annual maximum discharges for exceedance probabilities of 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2% were estimated at USGS gauge locations and for all simulation locations (gauged and ungauged) with the use of the USGS program PeakFQ version 7.3 (Veilleux et al. 2014). PeakFQ fits the log-Pearson type III distribution using the provided instantaneous annual maxima at each station according to procedures described in B17C (England Jr et al. 2019). We utilized the generalized method of moments estimator named Expected Moments Algorithm (EMA) (further described in B17C (England Jr et al. 2019) as it improves on procedures from B17B. The weighted skew option was implemented as it provides a better estimate for flood frequency analysis (England Jr et al. 2019). For Iowa, the regional skew coefficient for each station is taken as a constant of -0.4 as established by Eash et al. (2013). The determination of multiple potentially influential low flows (PILFs) or low outliers (LO) was conducted using the Multiple Grubbs-Beck test which was shown to improve the fit of small exceedance probabilities (Cohn et al. 2013). Finally, we tested for the presence of non-stationarity in the records by applying the Mann-Kendall test (Kendall 1975; Mann 1945) with a significance level of 5% with Bonferroni correction (Bonferroni 1936).

Regional estimation of AEP discharges in Iowa at ungauged locations utilizes region-specific equations determined from regional regression analysis based on least squares regression from Eash et al. (2013). Equations were developed for unregulated streams at each of the three flood regions in Iowa (Fig. 2-1.) using landform and soil parameters (Eash et al. 2013). Regional regression equations are implemented for Iowa in USGS Streamstats (USGS 2021), which provides peak flow statistics calculated according to Eash et al. (2013).

We compared regional estimates of AEP discharges from the current USGS regression equations for Iowa (Eash. et al. 2013) with estimates from flood frequency analysis using HLM simulations at 445 ungauged communities. Since the HLM simulates annual maxima that would be observed at ungauged locations we define its AEP estimates as the reference. We used the skill score, relative error, and bias of AEP discharges between the two models to highlight discrepancies in the current approach.

The skill of the regional regression to reproduce AEP discharges compared to the simulation values was evaluated using the mean square error skill score (Murphy and Winkler 1992). The mean square error skill score (SS_{MSE}) is defined as:

$$SS_{MSE} = \rho_{fx}^2 - \left[\rho_{fx} - \left(\frac{\sigma_f}{\sigma_x}\right)\right]^2 - \left[\frac{\mu_f - \mu_x}{\sigma_x}\right]^2 \quad (2\text{-}1)$$

where ρ_{fx} is the correlation of the regional regression model f and reference simulations x; σ_f and σ_x are the standard deviation of the predictions and reference simulations, respectively, while μ_f and μ_x are the mean of the predictions and the reference simulations, respectively. The SS_{mse} can be described by three terms: the first term is the potential skill of the model (i.e., the coefficient of determination), and it indicates the maximum skill that could be achieved if there were no biases; the second term is the slope reliability and measures the conditional bias; the third term is a standardized mean error, which represents a measure of the unconditional bias. A skill score of 1 corresponds to a perfect match between regional regression model and simulations, while a value of zero represents a skill that is the same as what would be obtained using the average simulation AEP discharge value across the 445 locations.

The magnitude of difference between AEP discharges from the USGS regional regression models and HLM simulation were examined by the relative error, which is defined as:

$$RE = \frac{USGS_{AEP} - Sim_{AEP}}{Sim_{AEP}} \quad (2\text{-}2)$$

where $USGS_{AEP}$ and Sim_{AEP} are the discharge value determined for a specific AEP from USGS Streamstats and simulations, respectively.

Finally, we examined the extent of overestimation or underestimation in current regional flood quantile estimation procedures with respect to the simulations using bias. Bias compares the AEP discharge from USGS regression equations to those from our simulations. We calculated bias as:

$$B = \frac{USGS_{AEP}}{Sim_{AEP}} \quad (2\text{-}3)$$

where a bias greater (smaller) than 1 indicates an over- (under-) estimation of discharge at an AEP using regional regression equations.

2.5 Results

We start by examining the validity of the PRISM-NLDAS product for use as an input for HLM simulations. Fig. 2-2 shows a comparison of monthly maximum discharges between the HLM simulations and USGS observations at 150 USGS gauges in Iowa. The results indicate the simulations forced with PRISM-NLDAS precipitation perform well as the points follow a 1:1 line. The simulations have monthly bias values ranging from 0.43 to 1.77, with a mean value of 0.80. The bias indicates that the model only underestimates monthly discharges values compared to USGS observations by an average of 20%. Quintero et al. (2020) originally found the use of StageIV precipitation data to produce reasonable results in the HLM model. As a reference, StageIV monthly maximum discharge comparisons (Fig. 2-3) from 2002 to 2019 showed that simulations overestimate maximum values for all months, with a mean bias of 1.46 indicating large differences compared to gauge observations. Overall, simulations with the PRISM-NLDAS match observations across all seasons, supporting the notion that it can be used for extended simulation period.

We tested the goodness-of-fit of the modeled annual maximum discharges compared to USGS observations using the KS test. Fig. 2-4 shows the p-values for each site across Iowa. There are 90 sites (76%) which have a p-value greater than the 5% level under Bonferroni correction, indicating there is evidence that the simulations and observations for these sites are drawn from the same distributions. Conversely, only 29 sites (24%) have a p-value lower than the 5% Bonferroni-corrected level, where the distribution of annual maxima between simulation and observation is not the same. These 29 sites had drainage area ranging from 12 to 11,222 km^2. Our results show that, for most of the sites (76%), the annual

maximum discharges from the HLM simulations and USGS observations are drawn from similar distributions. These results, together with those in Fig. 2-2, indicate that HLM is adequately simulating annual maximum discharge, providing support for use of the HLM in flood frequency analysis.

With HLM results validated, we assessed the performance of Streamstats AEP discharges compared to those obtained from the HLM for ungauged communities. Fig. 2-5 shows the relationship between HLM-based and Streamstats AEP discharges across all 445 communities. Streamstats and HLM discharge values match well for lower probabilities from 50%, 20%, 10%, and 4%, with relatively high R^2 and skill scores (Eq. 3-1) ranging from 0.92 to 0.93 and 0.83 to 0.91, respectively. As the AEP decreases and the associated discharges increase, the skill of the regression equations decreases, with skill scores decreasing from 0.78 to 0.60 from AEPs of 2% to 0.2%. One important aspect of the results in Fig. 2-5 is the alignment of points by region. In region 1, which includes the Des Moines Lobe landform, the AEP discharges between Streamstats and HLM trend along the 1:1 line across all AEPs. Regions 2 and 3 (Fig. 2-5) are rotated from the 1:1 line, with Streamstats overestimating discharges for small HLM estimates but converging to the 1:1 line for large discharge values.

To better examine these differences, we explore the relative error (Eq. 3-2) and bias (Eq. 3-3) between exceedance discharge estimates. Results obtained for relative error in Fig. 2-6 further highlight that the difference in methods varies based on flood region. The relative error is the smallest across all AEPs in region 1, where the values range from -0.6 to 0.0; this indicates that Streamstats is slightly underestimating discharge compared to HLM-derived values. In regions 2 and 3, the variability in the relative error across communities increases as exceedance probability decreases. For an AEP of 50%, region 2 has relative errors ranging from 0.0 to greater than 1.0; however, as the AEP reaches 0.2%, the relative error ranges from -0.6 to greater than 1.0 with no clear spatial pattern. The same relative error patterns are observed in region 3. For communities in regions 2 and 3, Streamstats overestimates AEP discharges compared to HLM for high exceedance probabilities (i.e., small return periods), and produces underestimation in discharges as exceedance probability decreases. The spatial variations in relative error can be explained by the upstream catchment area for each community. As shown in Fig. 2-7, the relative error decreases as the catchment drainage area increases, with an exponential pattern for regions 2 and 3. As the AEP decreases in these regions, the drainage area where the relative error converges to near or below zero also decreases (Fig. 2-7). For example, at an AEP of 50%, the relative error for region 2 decreases to a value of 1.0 at drainage area of 500 km^2, but for the 0.2% AEP it occurs at a drainage area of 50 km^2. On average, the largest relative errors are observed in regions 2 and 3 for drainage areas less than 100 km^2, with values above 2.0. For region 1, the relative error does not vary with catchment drainage area, indicating agreement between Streamstats and HLM discharge. Note that region 1 has the largest range of drainage areas compared to region 2 and 3. Another example of these patterns can be observed in the bias (Fig. S2-1) between Streamstats and HLM AEP discharge, which shows an exponential pattern in relative error and drainage area for regions 2 and 3. However, no pattern is observed between relative error and drainage area in region 1. For regions 2 and 3, overestimation (bias above 1) of AEP discharge occurs for Streamstats compared to HLM simulations for most communities. Overall, the region and drainage area impact the obtained AEP discharges for Streamstats when compared to physically derived simulations from HLM.

2.6 Discussion

The primary advantage of using a distributed calibration-free model like HLM is that it provides discharge values at any point on the stream network based on physical processes. Our statistical comparison of annual maxima highlights the good performance of the model in capturing the distribution of monthly and annual peaks important in flood frequency analysis across the state of Iowa. It should be noted that seasonal variations in Fig. 2-2 for monthly model performance are due to the model development, which focused on capturing flood characteristics during the summer season, when peak

flows typically occur in Iowa (Quintero et al. 2020). With the ability to capture the observed processes, we can assess how well current regional regression equations perform. For the 445 communities examined here, we found that the most comparable results between the two methods are observed for exceedance discharges larger than 1000 m^3s^{-1}, as indicated relative errors near zero and bias values near one. The best performance between the physically-based and the regional regression models occurred in region 1 which is part of the hydrologically-disconnected Des Moines Lobe. This region has the smallest bias and relative error compared to regions 2 and 3, where the regional regression equations vastly overestimate AEP discharge for small catchments. We note that performance of the HLM was similar across all flood regions compared to observations and attribute differences in estimates between methods to the region-specific regression equations.

To further understand the breakdown based on regions, we examine the difference in variables between the regional regression equations. For region 1, the standard errors for regional regression equations range from 31.8 to 45.2 percent (Eash et al. 2013), and the model includes drainage area, maximum 10-year 24-hour precipitation, and a constant of channel maintenance as predictors. This is the only region in Iowa in which a climate parameter (i.e., precipitation) is incorporated into the regional equations and uses the scaling-law for drainage area. Region 1 shows the smallest differences in AEP discharge between the regional regression equations and HLM simulations, possibly because of the incorporation of the rainfall parameter which is important for a physical perspective (Eagleson 1972) in a river system. In contrast, regions 2 and 3 do not include a rainfall parameter in the regional equations and place all parameters within a logarithmic scaling-law equation. For region 2, AEP discharges are determined based on drainage area, percent of Des Moines Lobe area, and a basin shape parameter. Similarly, AEP discharge for region 3 is dependent upon drainage area, hydraulic conductivity of soil, and a basin shape measure. The 50% AEP equation for region 3 is the only equation that uses a scaling-law for drainage area similar to the equations in region 1. The lack of a climate parameter, coupled with using a logarithmic scaling-law relationship for all parameters, likely leads to the large differences in discharges across all AEPs. The smaller differences shown in region 1 can likely be attributed to inherent uncertainty in the 10-year 24-hour rainfall intensity from NOAA Atlas 14 (Perica et al. 2013) used for the rainfall parameter. Finally, the average standard errors (Eash et al. 2013) for regions 2 and 3 range from 19.4 to 46.8 percent and 26.5 to 43.1 percent, respectively. Overall, the differences in comparison to flood frequency analysis with HLM highlight the shortcomings of the standard equations employed in engineering applications.

The use of HLM provides a few distinct advantages over regional regression. It allows to understand channel nesting within regions as it does not differentiate between regions. More specifically, in the current method for regional flood quantile estimation in Iowa, if a point of interest has a catchment contained within multiple flood regions, the USGS approach requires the analysis to be performed for each flood region and weight by the area within each region. This provides inconsistencies as physical processes within larger regions such as Iowa change gradually and are not defined by a strict border. Additionally, the HLM allows for in-depth modeling of channel nesting and network geometry that impact peak discharge as shown in Ayalew and Krajewski (2017) and Perez et al. (2018). Furthermore, the HLM is a continuous simulation system which is an approach that has shown promise for regional flood frequency analysis (Boughton and Droop 2003; Cameron et al. 1999; Wright et al. 2020). The method allows capturing the short-term and long-term memory in rainfall that are important for flood peaks (Blazkova and Beven 2002; Blazkova and Beven 2009; Boughton and Droop 2003; Cameron et al. 1999; Pathiraja et al. 2012). This allows for the catchment moisture, storage, and infiltration state prior to flood-inducing rainfall to be captured for a consistent mass balance (Blazkova and Beven 2002; Blazkova and Beven 2009). This is advantageous compared to event-based methods which require an assumption of losses based on a design rainfall.

Another advantage of the HLM model for flood frequency analysis is that it allows the exploration of future climate scenarios. Quintero et al. (2018) forced the HLM with climate projections to explore changes in 100-year flood for six basins and found that non-stationarity from climate change impacted the design event. With a successful framework developed for assessing future scenarios, the implementation of HLM allows the focus of flood frequency to shift from past observations to develop flood frequency guidelines based on future climate scenarios. It should be noted that any assessment on changes in flooding based on climate scenarios is limited to projected changes in precipitation and evapotranspiration. Currently, no geomorphic changes are accounted for in climate scenarios or the HLM framework and further research is necessary for more robust flood projections. The projected changes in flood frequency and its' limitation across the state will be the focus of future studies. A final advantage is that the use of physically-based models will help simplify and consolidate the regional flood frequency framework. Currently in the United States, each of the 50 states specifies several types of documentation for regional flood frequency, with some using state-developed guidelines and others implementing B17B or B17C. If models such as HLM are implemented nationwide, then the framework can focus solely on the modeling of physical processes for each area that would allow for consistency across the United States.

In this study, the configuration of the HLM model does not account for changes of land use. As the landscape shifts to impervious surfaces, the amount of available runoff from precipitation events increases, allowing for shorter time to peaks in the hydrographs and greater generation of volume (Shuster et al. 2005). This is a principal factor that should be investigated to quantify the impact of land use on annual maxima. Another limitation in our analysis is that the data of reservoir releases were not included in the model configuration, even though these could be incorporated if available (see Ayalew et al. 2017). Providing information downstream of a reservoir is important so that flood frequency analysis can be available along any point in a stream network as current regional regression equations do not apply to regulated streams. For Iowa this is not an important issue as there are only three main reservoirs within the state, but it should be investigated if the model is applied to regions that have extensive regulation.

Distributions other than log-Pearson type III, such as extreme value type I, general extreme value, lognormal and gamma, have been used for modeling annual maximum discharges (Bobée et al. 1993; Cunnane 1973). We applied the Log Pearson Type III as it is a standard approach for the United States to provide a consistent comparison with current guidelines. An additional potential avenue to improve distribution fitting for simulated discharge is to use the peak-to-volume approach (Bradley and Potter 1992), where probability distributions are estimated from runoff volume and for peak discharges conditions on runoff volumes which can provide improved consistency for extreme events. Further exploration of distributions could be considered alongside the use of a calibration-free model to provide additional improvement for flood frequency analysis.

2.7 Conclusion

This study presented the application of regional flood frequency analysis based on the results of HLM, a calibration-free physically-based hydrologic model. First, we compared HLM-based and observed annual maxima at 119 USGS gauge locations across Iowa. Our findings indicated that HLM simulations adequately reproduce the distribution of annual maxima for a majority of gauged locations. Next, we explored the use of the HLM for regional flood frequency analysis by comparing annual exceedance discharges with values from USGS regional regression equations for the state (Eash et al. 2013) at 445 communities. The best performance between the physically-based model and the regional regression model occurred in the Des Moines Lobe (region 1), where regression equations incorporated precipitation as a variable. For the other two regions (2 and 3), the largest discrepancies in exceedance discharge occurred when the discharge value is less than 1000 m^3s^{-1} where drainage area was typically less than 500 km^2. At these locations, regional equations only incorporated catchment properties, overestimating exceedance discharge compared to HLM simulations. We concluded that the use of a physically-based model like HLM

in flood frequency analysis provides a consistent method for modeling exceedance discharges necessary for engineering design as well as provides a framework to reassess current regional flood quantile estimation practices.

This chapter focused on a relevant problem for engineers within the context of climate change as well as the needs to improve resiliency of the nation's infrastructure. Our study demonstrates that the primary advantage of transitioning to flood frequency analysis with physically-based model simulations is consistency across larger spatial domains. The current standard approach for USGS regional regression equations for Iowa defines flood regions where no true physical or climate boundary exists. This provides inconsistencies in estimates as it does not allow for physical processes such as channel nesting to be adequately represented, impacting design discharge estimates. The approach proposed here allows engineers to design for the future as flood frequency analysis can be conducted for future climate scenarios. Precipitation and evaporation projections are available from sources such as the Coupled Model Intercomparison Projected Phase 5 (CMIP5) (Taylor et al. 2012) which can be used as input for the HLM to simulate future projected discharges. This would allow for the improvement of flood-related infrastructure as new design guidelines can incorporate the role of nonstationarity in flood frequency analysis. Overall, we recommend further exploration of physically-based models for flood frequency analysis as they provide much needed hydrologic justification to current statistically driven guidelines.

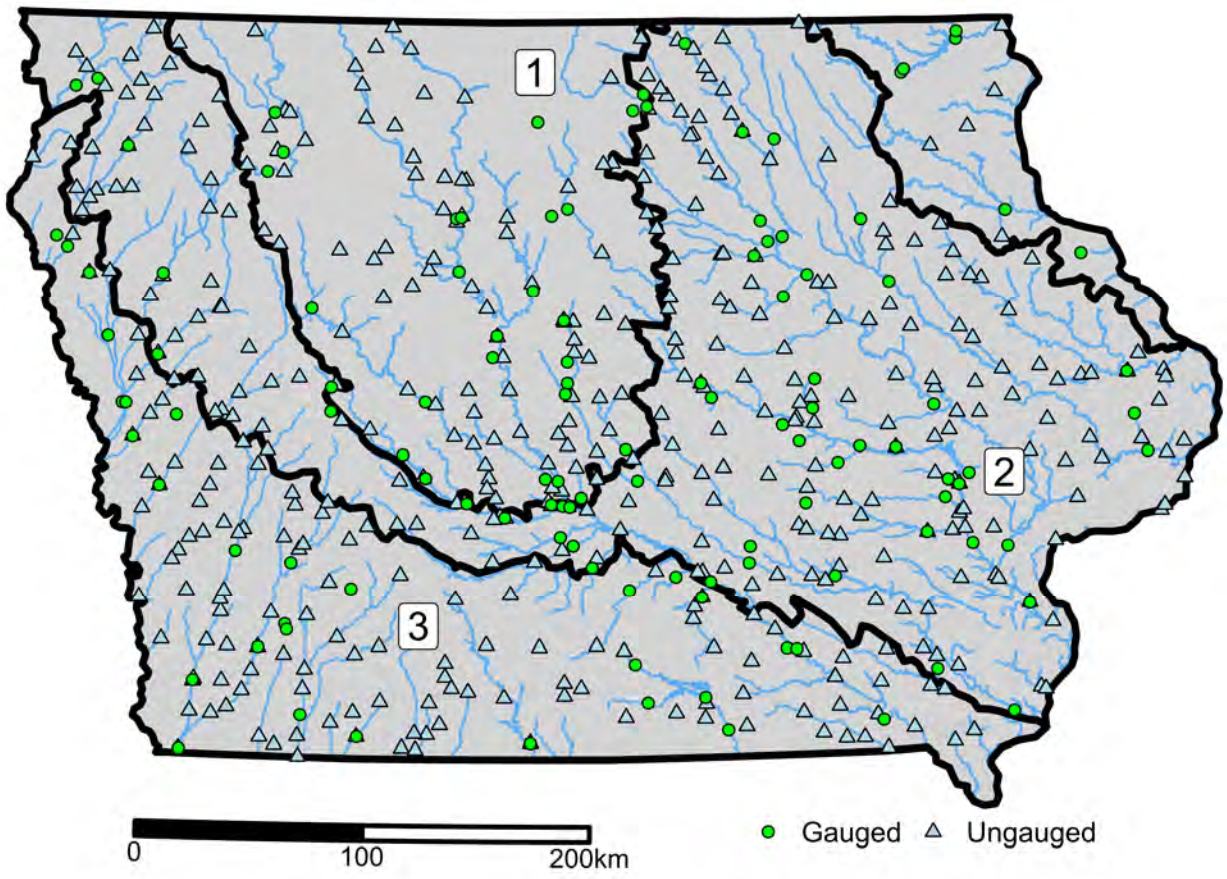

Figure 2-1. Location of flood regions: USGS stream gauges (green circles), and ungauged communities (light blue triangles).

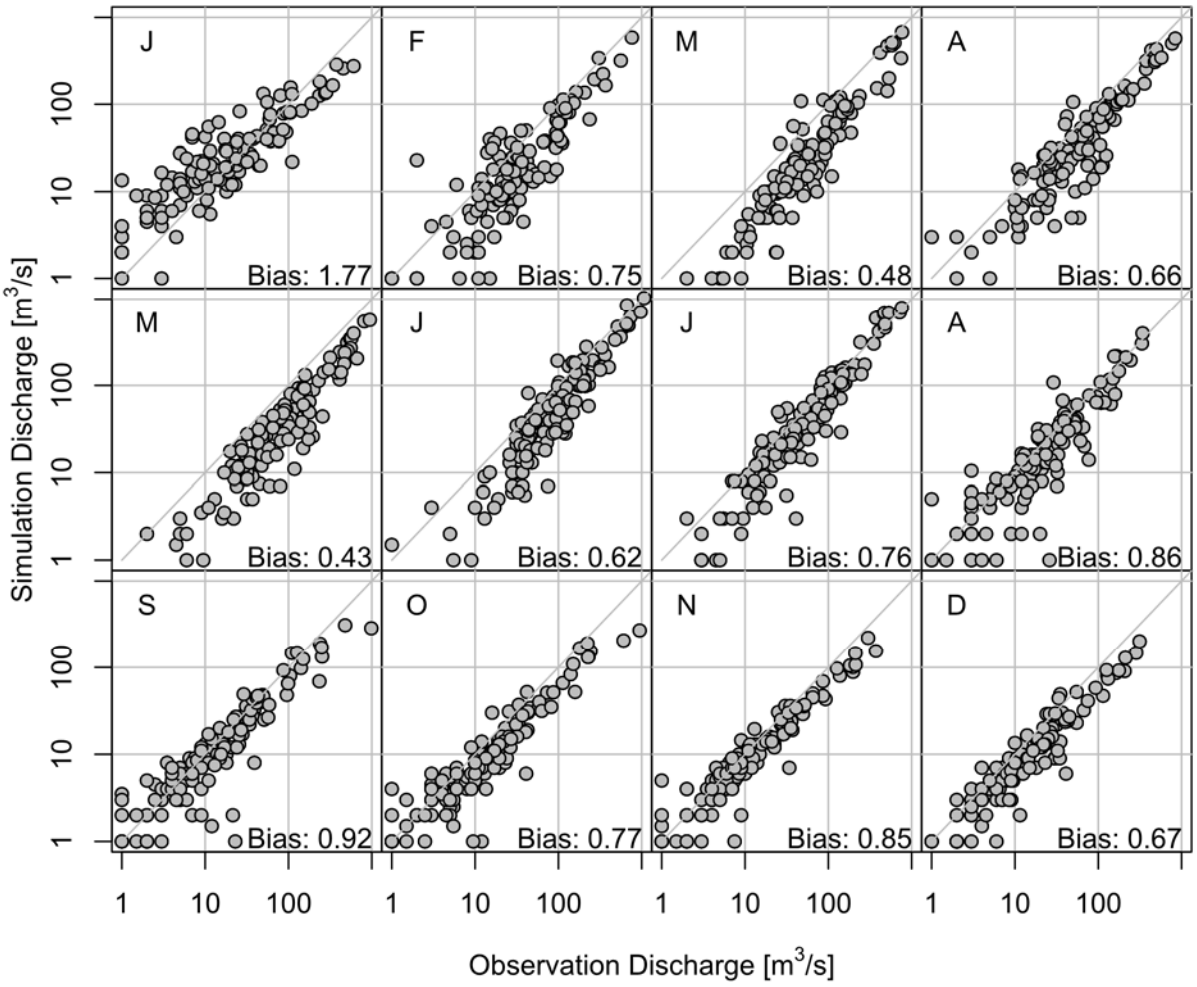

Figure 2-2. Ability of the HLM using PRISM-NLDAS as input to reproduce the observed average monthly maximum discharges at 150 USGS gauges in Iowa from 1981 to 2020. The observational data are based on the daily-averaged gauge discharge. The bias shown is the average ratio of the simulated discharge to USGS observations for all 150 gauges. A bias above (below) 1 indicates that model simulations overestimate (underestimate) discharge compared observations. Months are indicated by the letters in the upper left corner.

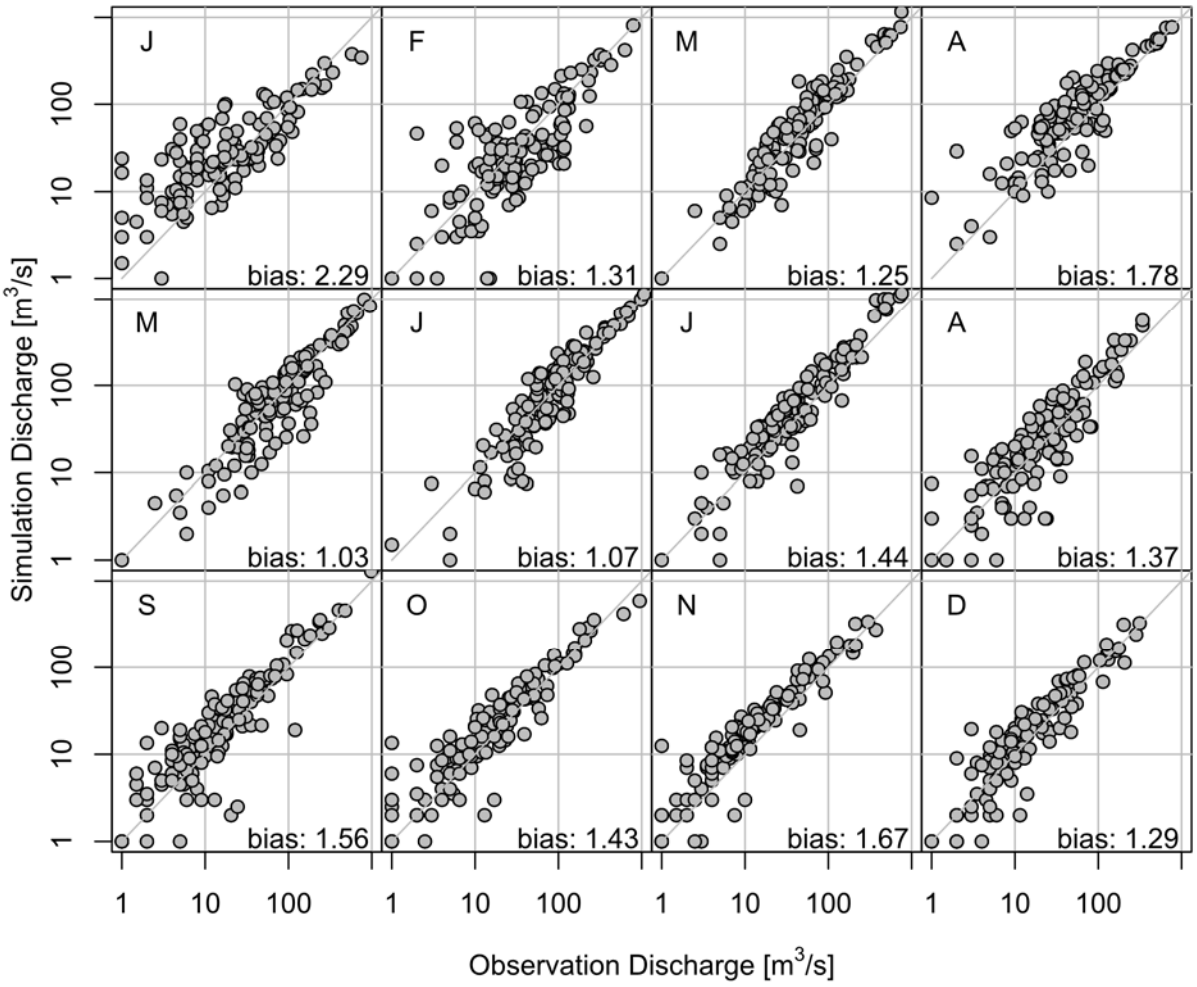

Figure 2-3. Ability of the HLM using StageIV input to reproduce the observed average monthly maximum discharges at 150 USGS gauges in Iowa from 2002 to 2019 where instantaneous gauge data are available. The bias shown is the average ratio of simulation volume to USGS observation volume for all 150 gauges. A bias above (below) 1 indicates simulations overestimate (underestimate) volume compared observations. Months are indicated by the letters in the upper left corner.

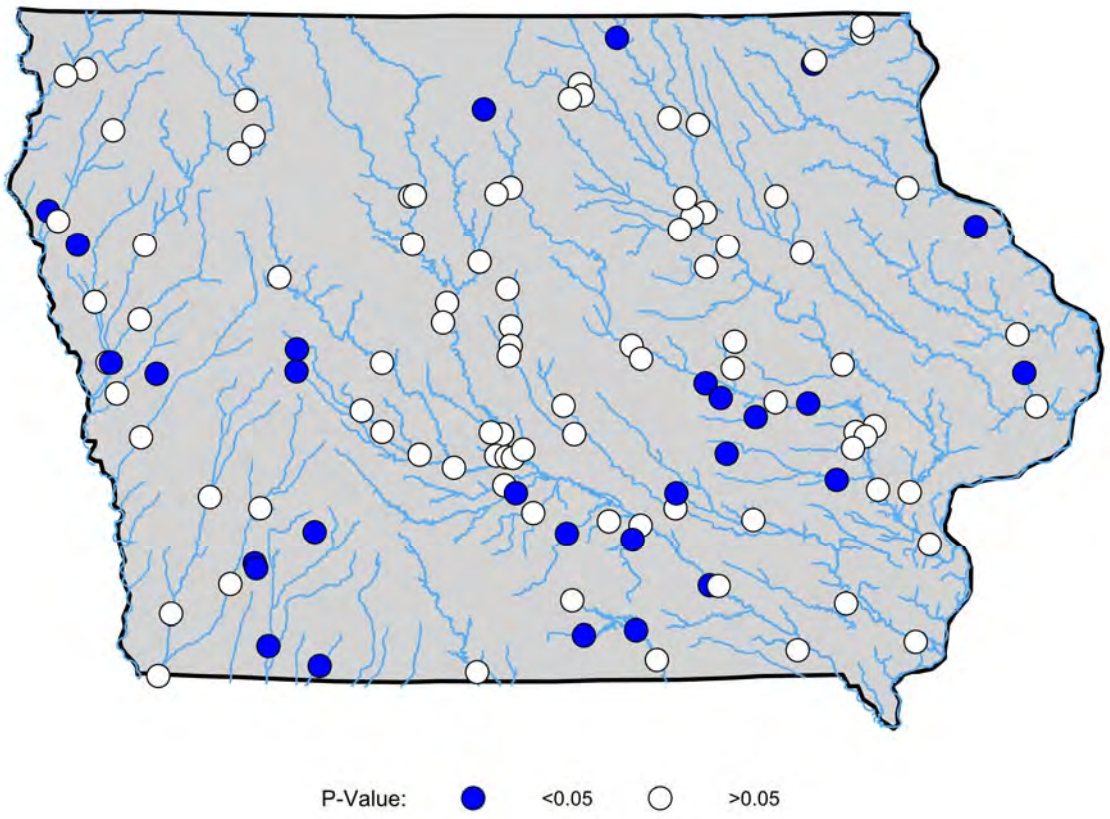

Figure 2-4. P-value for each USGS gauge location determined from the Kolmogorov-Smirnov test. The points in blue represent the Bonferroni corrected significance levels for 0.05.

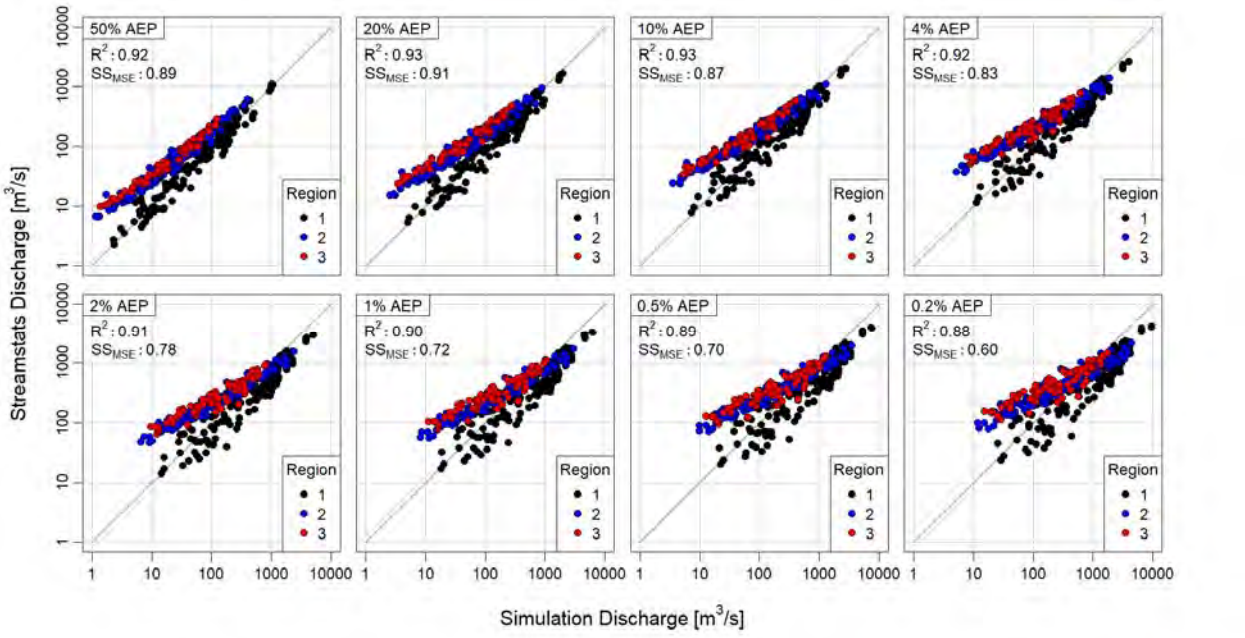

Figure 2-5. Relationship between annual exceedance discharge derived with Streamstats and HLM simulations at probabilities of 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2%. The value of R2 represents the coefficient of determination, and SSMSE is the mean square error skill score. The three regions refer to those in Figure 2-1.

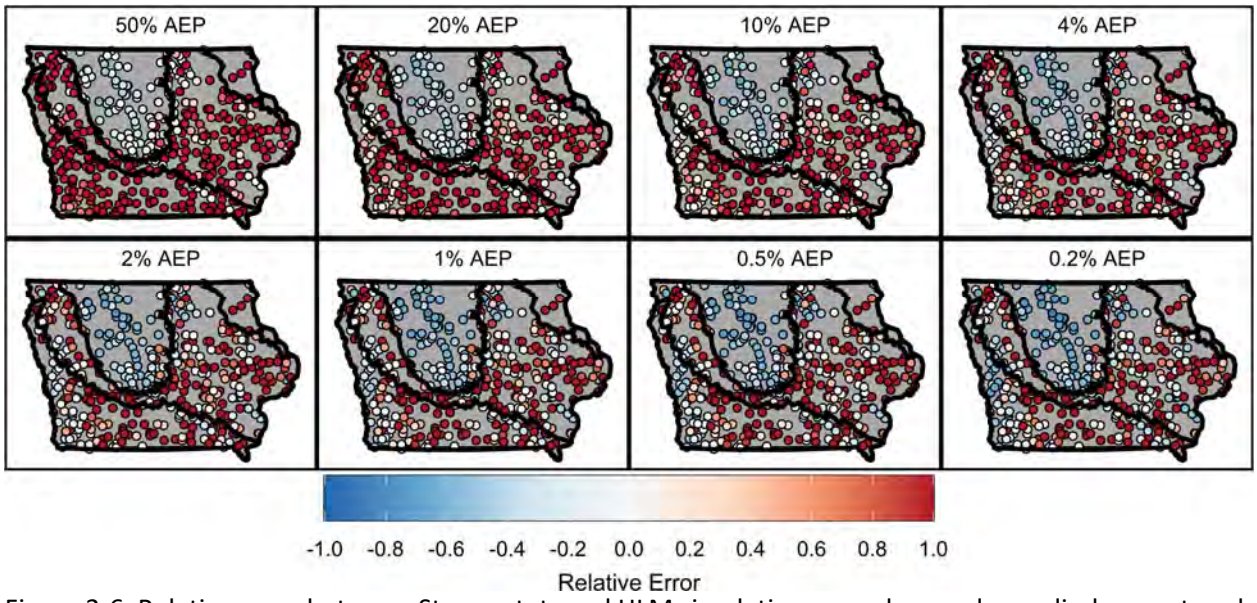

Figure 2-6. Relative error between Streamstats and HLM simulation annual exceedance discharge at each of the ungauged evaluation sties at probabilities of 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2%. Values near zero indicate identical discharge values between StreamStats-derived discharge and HLM-based discharge determined with PeakFQ.

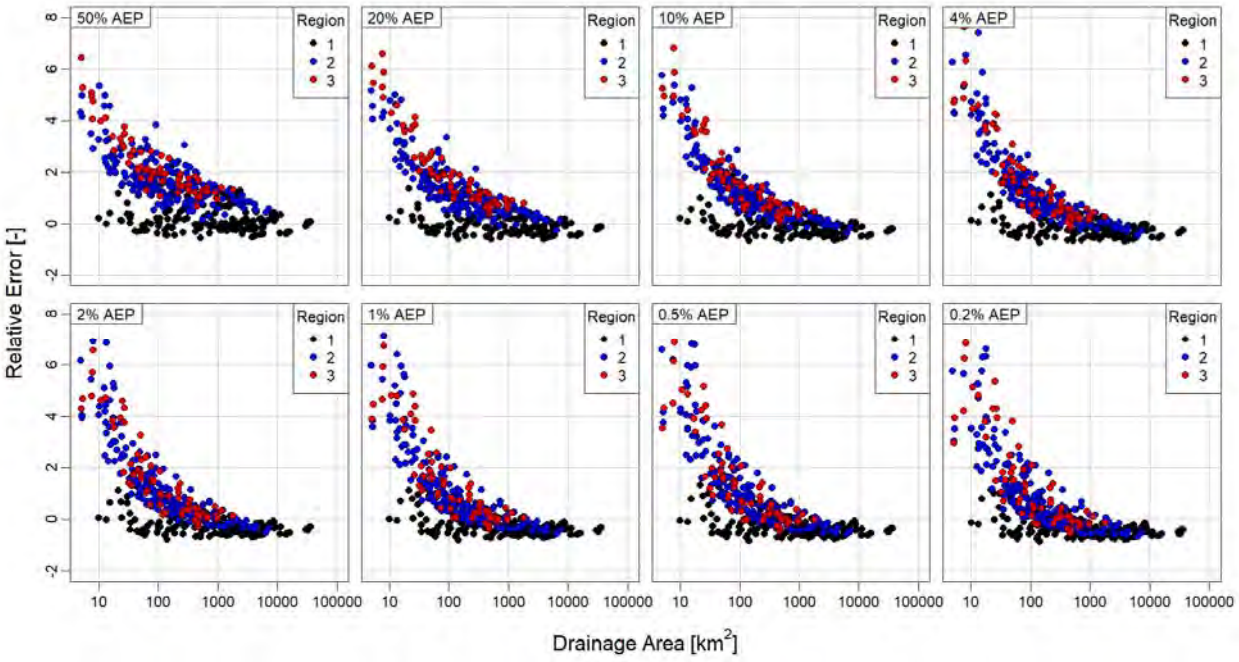

Figure 2-7. Relation between scale of the watershed and the relative error for annual exceedance discharge at probabilities of 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2%. The points are colored by the flood regions defined in Eash et al. (2013).

3. Projected changes in annual maximum discharge for Iowa communities[3]

3.1 Introduction

Over recent decades, Iowa has experienced an increasing number of flood events (Neri et al., 2019; Slater and Villarini, 2016) with large socio-economic impacts to communities (Hicks and Burton, 2008; Smith and Katz, 2013). Recently, Byun et al. (2019) examined four large watersheds across Iowa showing that monthly streamflow is expected to increase in the spring and summer months over the next century. This is expected to bring increases in the 100-year daily streamflow magnitude for those four watersheds and in the frequency of flood events. These streamflow alterations are projected to be driven by increases in precipitation during the winter and spring, changes in snow water equivalent and in soil moisture across the state. Moreover, frequent flood events will likely result in an increase in economic losses (Dottori et al., 2018), requiring the development of adaption and mitigation strategies to improve community resiliency.

To understand future flood impacts at the community scale, we need to utilize localized flood information. This is supported by Disaster Resilience: A National Imperative (National Research Council, 2012), stating this idea as a characteristic for a more resilient nation. For Iowa, two case studies already demonstrated the benefit of using localized information to deduce social and economic loss information important for resilience planning. For instance, Tate et al. (2016) used local gauge data and flood models to examine social equity for the historic 2008 flood in Cedar Rapids, Iowa; they found that property acquisition are an effective measure to reduce flood losses in the future and prioritize vulnerable neighborhoods. Furthermore, Villarini et al. (2020) used hydrologic and hydraulic modeling to investigate the role of anthropogenic forcing on socio-economic losses for the 2008 flood at Cedar Rapids. They determined that anthropogenic forcing increased inundated area by 1.28-fold, which led to a seven-fold increase in economic losses. Each of these studies demonstrates how the creation of localized flood information allows for better understanding of socio-economic impacts caused by flooding. To apply these concepts to climate change preparedness, there is a need to produce flood projections at a community level.

Currently, the focus of hydrologic projections using climate change scenarios is at the continental and regional scales. Prior studies (Byun et al., 2019; Demaria et al., 2016a; Demaria et al., 2016b) have typically applied the Variable Infiltration Capacity (VIC) model (Liang et al., 1994; Liang et al., 1996) forced with outputs from general circulation models (GCMs) obtained from the Coupled Model Intercomparison Projected Phase 5 (CMIP5) (Taylor et al., 2012). However, the available projections for communities across Iowa are limited. For instance, Demaria et al. (2016a) investigated projected changes in streamflow characteristics for the U.S. Northeast-Midwest, but only provided projections for flood peaks at six sites located in the eastern portion of the state. Other studies, such as Byun et al. (2019), examined a large spatial portion of Iowa by focusing on four watersheds covering the state; however, they only provided projections at the basins' outlet, limiting the usefulness of their results to communities. Each of these studies shows a potential for projected increases in flood peaks but lacks the granularity necessary to inform community-level action.

At the finer-scale (e.g., basin-level), understanding temporal changes has relied primarily on analyses of historical records or streamflow simulations. For instance, Mallakpour and Villarini (2015) performed peak-over-threshold analyses at stream gauges across the Midwest. They examined over 70 locations in Iowa, finding an observed increase in the frequency of flood events from 1962 to 2011. However, these results only provide insights about changes in the historical past, and not about projected changes.

[3] Adapted from: Michalek, A., F. Quintero, G. Villarini, and W.F. Krajewski, Projected changes in annual maximum discharge for Iowa communities, submitted to Journal of Hydrology, 2022.

Conversely, the work by Tohver et al. (2014) is one of the few studies which provides high-resolution future flood information that can be utilized by communities to prepare for climate change. The authors assessed flood projections with the VIC models across the Pacific Northwest to create high-resolution (i.e., 1/16-degree) 100-year flood maps. Bates et al. (2021) examined flood inundation under Representative Concentration Pathway 4.5 (RCP4.5) across the United States at a 30-m spatial resolution (~1/4-degree) with the LISFLOOD-FP two-dimensional hydraulic model. They used a framework to gather input data from across the United States and account for fluvial, pluvial, and coastal flooding. They indicated that significant changes in the land area at risk from 100-year floods will occur by 2050 under RCP4.5, with as much as 16% increases for certain areas. However, their study does not provide flood peak projections across the United States, which are important for developing flood frequency estimates for engineering design. These approaches represent a few of the high-resolution modeling efforts in the literature, providing the foundation for expanding the literature.

For Iowa, we can address the current community-level gap in the literature by leveraging the resources of the Iowa Flood Center (IFC). More specifically, we can take advantage of the physically-based hydrologic model called the Hillslope-Link Model (HLM), which can simulate streamflow, including flood peaks, anywhere in the state. Here we use the HLM to assess the impact of climate change on annual maximum discharge for 1,000 communities across Iowa under two emission scenarios, and examine the physical drivers responsible for these changes. This is particularly important for a rural state like Iowa, in which many of the smaller communities do not have access to tailored climate information that may be available for larger cities. Finally, we examine the role of variability in projections and the impact of climate model selection on community-level analysis.

This chapter is organized as follows. In Section 2 we describe the study area, model set-up, and forcing data. In Section 3.3 we describe the statistical methods used to assess CMIP5 model selection and our annual maxima analysis. Section 3.4 provides the major results and discussion of our analyses. We finish the chapter with conclusions and final remarks.

3.2 Study Site and Data Sources

We focus on projections in annual maximum discharge characteristics at 1,000 communities across Iowa, located in a close proximity to stream and rivers, through the application of hydrologic simulations. The domain includes all basins which drain into the Mississippi and Missouri Rivers with drainage areas ranging from ~10 to 36,000 km^2.

We simulate streamflow with the TETIS version of HLM as discussed in Quintero and Velasquez (2022). The theoretical background and conceptualization of the HLM is well described in Mantilla et al. (2022). The TETIS version of HLM is a fully distributed hydrologic model structure, which is based on the decomposition of the landscape into hillslopes and channels. The model estimates runoff generation at the hillslope by simulating different processes Snow accumulation and melting rates are derived from precipitation and temperature data, based on the degree-day method. The model simulates precipitation losses in vegetation and soil pore macrostructure. Surface infiltration takes into account the hydraulic soil properties of the hillslope, and the conditions of frozen ground during the cold season. Deep percolation and groundwater losses are also modelled, considering the hydraulic conductivity of soil at the subsurface. Total hillslope runoff is aggregated in the river channel from the contribution of overland flow, interflow, and base flow. A nonlinear hydrologic routing model transports flow in the channels and takes into account the geomorphologic characteristics of the river network. The model equations are written as ordinary differential equations. HLM integrates a numerical solver that takes benefit of the binary tree structure of the river network to solve the system of differential equations in an asynchronous manner, suitable for high performance computing environments (Small et al., 2013). In our implementation, the HLM outputs streamflow simulations are produced at hourly time step and annual maximum discharge is

obtained for all channel links in the river network. Projections of precipitation, temperature, evapotranspiration, and frozen ground conditions are forced into the hydrologic model. The spatial and temporal resolution of these inputs are key for the performance of the model (see Quintero and Velasquez (2022) for more details). Simulations are conducted under naturalized flow conditions where stream regulation (i.e., dams) is ignored due to the lack of available data on future operating procedures.

To study the projections of annual maximum discharge, we force daily precipitation, daily temperature, daily frozen ground, and monthly ET forcings derived from CMIP5 (Taylor et al., 2012) through the HLM. For projections, snow accumulation and melting is based on air temperature from GCMS using a threshold and melt rate for the respective processes. Additionally, the HLM is forced with an input of frozen ground conditions, that is calculated based on the GCM specific air temperature with a simplified soil temperature module (Rankinen et al., 2004). We utilize 19 climate models and three sets of simulations (i.e., Historical, RCP4.5, and RCP8.5). We use historical experiments covering the 1950-2005 period as control for our analysis, while RCP4.5 and RCP8.5 represent future midrange and high-emission scenarios, respectively (Moss et al., 2010). We utilize CMIP5 daily simulations that were downscaled statistically using bias-correction constructed analogues (BCCA; Maurer et al., 2010) to a horizontal resolution of 1/8° latitude-longitude (~12km). We use monthly ET from the original GCMs associated with the downscaled data sets. We utilize the bias-corrected and downscaled results as precipitation with a high spatial resolution is necessary to adequately model annual flood peaks and biases in native CMIP5 forcings have been shown to influence runoff generation (e.g., Lehner et al., 2019; Seager et al., 2012). Finally, we force the HLM simulations from 1981 through 2005 with daily Parameter Elevation Regression on Independent Slopes Model (PRISM) rainfall (Daly et al., 2002) to provide a reference for the GCM selection.

3.2 Methodology

Before examining the projected changes in annual maxima, we examine the performance of the HLM to reproduce observed maximum discharge. We use annual maxima and daily-averaged discharges from 150 United States Geological Survey (USGS) stream gauges that were active from 1981-2020. We examine the average bias across the 150 sites in monthly and annual maximum discharges to assess the ability of the model to reproduce observed peak flows and the suitability of the model as reference. We define the bias as:

$$B = \frac{USGS_{max}}{Sim_{max}} \tag{3-1}$$

where a bias greater (smaller) than 1 indicates an over- (under-) estimation of maximum monthly/annual discharge at a gauge location.

After model validation, we evaluate the performance of the GCMs in reproducing the observed trends in annual maximum discharge. While uncertainties in the input and the hydrologic model exist, we look for spatial consistency among the GCM derived outputs. For that purpose, we use three performance metrics to identify the GCMs which reproduce historical annual maximum discharge well. The three metrics consist of Kendall's tau matching, Mann-Whitney U test, and the Kolmogorov-Smirnov (KS) test (Massey, 1951) to ensure that simulated annual maxima in the climate models match the reference historical trend, magnitude, and distribution, respectively, at each of the 1,000 communities. For each of the tests, we compare the annual maximum peak discharge when the HLM is forced by PRISM and by the GCM outputs over the 1981-2005 period. We use these tests to select a subset of suitable climate models for our primary analysis. For the Kendall's tau matching, we compute Kendall's τ and its statistical significance in annual maximum peak discharge when the HLM is forced by PRISM and by the GCM outputs. Locations where both test statistics match in sign and are statistically significant ($p \leq 0.05$) along with points

where both test statistics are not statistically significant ($p>0.05$) are counted as good performers for model selection. For the Mann Whitney U test and KS test, we determine the p-value at each of the communities. We select models which have greater than 90% of the sites agreeing with respect to Kendall's tau matching, Mann Whitney U test, and the KS test as adequately capturing the statistical patterns in the historical experiments.

Our primary focus is to assess projected changes in annual maximum discharge characteristics. We assess nonstationarity in future scenarios using the Mann Kendall trend test applied to the entire future period (2006-2099) at significance levels of 5% and 10%. We examine changes in annual maximum discharge magnitude using the relative difference between the median historical annual peak from 1976 through 2005 and the median annual maxima for future 30-year periods (i.e., 2006-2035, 2036-2065, and 2066-2095). Finally, we apply the KS test to determine a time of emergence for which the distribution of projected annual maximum discharge is different with respect to the historical one at a 5% significance level. The reference period is based on the annual maximum discharge from 1956-2005 and the future periods are 50-year projections starting in 2006 (i.e., 1957-2006; 1958-2007; …; 2050-2099).

3.4 Results and Discussion

3.4.1 HLM Performance and GCM Selection

We start by examining the suitability of the daily PRISM product for use as an input for HLM simulations as a historical reference. Daily scale preserves total volume compared to higher temporal resolution, but it affects the simulated peak timing (delay). However, this aspect of the peak flow is not important for our analyses. Fig. 3-1 shows a comparison of monthly maximum discharges between the HLM simulations and USGS observations at 150 USGS gauges in Iowa. The results indicate the simulations forced with PRISM daily precipitation perform well as the points follow a 1:1 line (see Fig. S3-1 for volume performance). The simulations have monthly bias values ranging from 0.3 to 1.45, with a mean value of 0.76. The bias indicates that the model only underestimates monthly discharges values compared to USGS observations by an average of 24%. This monthly peak performance is similar to hourly precipitation forced in the HLM as seen in Michalek et al. (2022), Quintero et al. (2020) , and Quintero and Velasquez (2022). For the annual maximum discharge, Fig. 3-2 shows the average basin bias as a function of drainage area. For basins larger than 100 km^2, the median bias is about at 1.0 or above. For basins above 1000 km^2, the median bias is near 25% indicating an overestimation of annual maximum discharge in the reference simulations. The weakest performance is in basins below 100 km^2 where there is an underestimation of peak flows (median average basin biases around 0.50) due to the daily resolution of the precipitation. Overall, we deem that for most of Iowa, the reference simulations can adequately capture historical observations at the seasonal and annual scale and can be utilized as the reference against which to compare the GCM-driven hydrologic simulations.

Next, we evaluate which models to include for the primary assessment of future projections. We compare hydrologic simulations derived from the 19 climate models with respect to our reference simulations. Table 3-1 provides the breakdown of the percentage of sites satisfying the criteria based on Kendall's tau matching, Mann-Whitney U test, and KS test for each of the 19 GCMs. Nine of the 19 GCM-derived simulations met our 90% criteria for all three tests (Table 3-1) and are the primary focus for our subsequent analyses. Figs. S3-2 - S3-4 show the scatter and spatial plots for each of the three tests. Additionally, we perform our analyses using all GCMs as well to provide commentary on the differences in results related to model selection (see Section 4.4).

3.4.2 Projected Annual Maximum Discharge

We start our analyses by examining the presence of nonstationarities in projected annual maximum discharge. There are strong positive trends for the future period (2006-2099) under RCP8.5 (Fig. 3-3b) at

the 5% level for a majority of the state (70%). In comparison, under RCP4.5 (Fig. 3-3a) we detect fewer trends across the state, with positive trends for only 2.2% and 3.8% of the state at the 5% and 10% level, respectively. These results highlight the importance of examining community impacts across different emission scenarios. For reference, studies assessing trends in observed streamflow for Iowa (Mallakpour and Villarini, 2015; Villarini et al., 2011) indicate a lack of detectable trend at gauged locations. In relation to our results, this suggests that the lack of trends in historical data does not guarantee a lack of future trends as discussed by Villarini and Wasko (2021).

To further understand the patterns in trends for annual maximum discharge, we shift our focus to the magnitude of change with respect to the historical conditions. Fig. 3-4 presents the results of the relative difference of median discharge for each of the emission scenarios and three 30-year periods with respect to the 1976-2005 historical period. Across all scenarios, the median annual maximum discharge within each period is projected to increase with respect to the historical median value. However, the change across periods differs between scenarios. This helps explain the difference in trends between emission scenarios as the change in annual maxima is projected to increase towards the end of the century at a greater rate under RCP8.5 compared to RCP4.5. For RCP4.5 from 2036-2095, the change in median peak discharge is projected to be small (Fig. 3-4, panels c and e), with expected increases occurring in the near term (2006-2035, Fig 3-4a). These patterns can be attributed to the difference in emission projections between RCP4.5 and RCP8.5, with those for RCP4.5 that are projected to stabilize around the middle of the century. Finally, we find agreement with Byun et al. (2019) for Iowa, who showed that the percent change in 100-year return flow increases under RCP4.5 compared to RCP8.5 for the nearest 30-year period, while the highest percent increase is shown towards the end of the century under RCP8.5.

Next, we shift our focus to understanding when (i.e., which decade) the change in the distribution of historical annual maxima is projected to be first detected. This is important to understand for communities as the distribution of annual maxima is used to determine the annual exceedance probability of discharge for design purposes. If the projected distributions change compared to the historical ones, then water resource planning in those communities would need to be reassessed. In this context, we define the time of emergence as the end year of the first 50-year period in which the significance level of the KS test with respect to the historical period (1956-2005) is exceeded ($p<0.05$). Fig. 3-5 shows the results of the time of emergence for each emission scenario. Under RCP4.5, the time of emergence for half of the sites occurs before 2070, with the remaining sites having a time of emergence between 2060 and 2090. For RCP4.5, 296 sites (29%) do not show a significant change in distribution by 2100. RCP8.5 shows differing patterns as 270 sites (27%) have a time of emergence by 2070. The remaining sites under RCP8.5 have time of emergence much later in the century compared to RCP4.5, with 53% of them showing a time of emergence by 2091. RCP8.5 has fewer sites (19%) with no significant change in distribution by 2100 compared to RCP4.5. Each emission scenario indicates that, for a majority of sites, the distribution of annual maxima is expected to change in the near future. However, under RCP4.5 the change is projected to occur sooner on average compared to RCP8.5. For many communities across Iowa, current water resource designs may not be adequate as the distributions from which design storms are estimated are projected to change. This analysis provides valuable information for understanding where stakeholders and water resource managers in Iowa should focus on in reassessing current design and management plans.

3.4.3 Climate Drivers

To better understand our results, we investigated precipitation across Iowa as it is the primary driver in the HLM. We focus on precipitation in terms of the relative difference in Simple Daily Precipitation Intensity Index (SDII) and cumulative precipitation volume. The SDII represents the average daily wet day precipitation within a period. For RCP4.5 we find a percent change in SDII for 2006-2035 with respect to the historical period (1976-2005) ranging from -1.5 to 3.9% (Fig. 3-6a). For the 2036-2065 (Fig. 3-6c) and

2066-2095 (Fig. 3-6e) periods, the percent change in SDII ranges from -0.6% to 7.5%. Most of the state is projected to experience an increase in wet day precipitation under RCP4.5 that is expected not to change between 2036 to 2095. This is consistent with the changes in distribution of annual maximum discharge which are detected earlier in the century for most sites under RCP4.5 compared to RCP8.5. For RCP8.5 we find that the change in SDII ranges from -2.9% to 2.2% for the 2006-2035 (Fig. 6b) and -1.6% to 6.3% for 2036-2065 (Fig. 3-6d) periods with average values of -0.4% and 1.5%, respectively. However, by the end of the century (Fig. 3-6f), the SDII is projected to increase up to 9.6%, with an average of ~4.5%. This finding suggests that the intense precipitation events are increasing and supports the high detection of trends in the magnitude of annual maximum discharge for RCP8.5 as well as the later shift in distribution. We observed similar patterns for the relative difference in cumulative precipitation volume between the historical and future periods (Fig. S3-3 - 3-5). This is consistent with Wasko et al. (2021) who found that extreme precipitation has already increased and can be attributed to an increase in the magnitude of extreme flood events.

3.4.4 GCM Model Selection

A crucial element to consider is the impact of GCM selections on flood peak projections. For our analyses, we selected nine out of 19 models with precipitation, temperature, and ET that reproduced well the historical trends, magnitude, and distribution. However, there are many different methods used in the literature for climate model selection. For instance, Demaria et al. (2016a) selected 16 models from the bias-corrected and statistically downscaled products as we used (Maurer et al. 2012) based on Reliability Ensemble Average (REA) method. Others such as Naz et al. (2016) selected ten models from the original CMIP5 suite based on the availability of sub-daily timescale for hydrologic simulations. Each of these methods highlights the variability in selection and the importance of understanding uncertainty in modeling future projections.

To address the uncertainty in our analyses, we compared the previous results from the nine selected models against those obtained using all 19 available models. Our goal is to assess how the results changes when utilizing all available GCM simulations. Analyses using all available simulations indicate a significant trend under RCP8.5 (Fig. S6b) for the entire future period (2006-2099) at the 5% level for 66% of the sites, which is a decrease compared to using the subset of models (Fig. 3-6b). Additionally, we find that under RCP8.5 the spatial distribution of communities with significant trends changes. Finally, under RCP4.5 (Fig. S6a) we find similar lack of detected trend across most sites (98%), similar to the results shown in Fig. 3-3a.

We further investigate the impact of model selection on annual maximum discharge projections by comparing the relative difference analysis. On average, the relative difference changes using all models compared to those selected under RCP8.5 for the 2006-2035, 2036-2065, and 2066-2095 periods by +8.1%, +10.4%, and -1.1%, respectively. Similarly, for RCP4.5, the relative difference changes by +6.7%, +3.9%, and +0.6% when utilizing all climate models. The inclusion of all models increases the magnitude of the projected flood peaks for a majority of communities across Iowa compared to the models selected for our primary analysis.

For the time of emergence (Fig. S3-3 - 3-7), the inclusion of all GCMs points to a shift to earlier in the future compared to the subset of nine models (Fig. 3-3 - 3-5). Under RCP4.5, the time of emergence occurs between 2030 to 2060 for the eastern half of the state, and between 2060 to 2090 for the western portion. Under RCP8.5, a similar shift occurs, with most sites showing a shift in distribution between 2040 and 2060. The inclusion of all models shows more similarities between emission scenarios but keeps the overall pattern of later time of emergences in RCP8.5. We find this is primarily due to the increase in discharge magnitude by the middle century under both emission scenarios using all climate models.

These differences highlight the importance of understanding the impact of model selection for community flood estimation as different climate models will produce varying trends, magnitudes, and

distributions in flood peak projections. These discrepancies due to model selection must be conveyed to communities when providing localized flood projections to allow for in-depth discussion of which estimates to rely on for resilience planning. Furthermore, as discussed in Wasko et al. (2021), we must use this information to provide flexibility for flood estimate guidelines at the community-level to ensure adaptability and improved resiliency in water planning.

3.5 Conclusion

In this study we investigated the response of annual maximum discharge over time under climate change scenarios (RCP4.5 and RCP8.5) at the community level across Iowa. We utilized the physically-based HLM to conduct simulations under naturalized flow conditions from 1950 to 2100 to understand projections with respect to historical records. Since the HLM keeps track of soil moisture, using a continuous simulation mimics basin memory and the effects of antecedent moisture on runoff generation. A significant positive trend is detected for 70% of sites under RCP8.5, compared to only 2% of the sites under RCP4.5 for the selected GCMs. Under RCP8.5, a more intense increase in annual maxima is projected to occur later in the century (i.e., between 2066 and 2095) compared to RCP4.5. This is due to the projected increase in precipitation intensity during 2066-2099 under RCP4.5. This pattern impacts the time of emergence for when annual maximum discharge distribution changes. Under RCP8.5, we detected changes later in this century for ~48% of the sites compared to RCP4.5. Finally, our results showed that trends, magnitude, and distribution shifts can vary between communities even over a climatologically-similar region. Furthermore, shifts in distribution can occur even when the presence of trends cannot be detected. This illustrates the need to develop tools utilizing our projections which can improve community resilience to flooding due to climate change as well as highlights the need to develop models for the communities outside of Iowa.

The primary limitations of this study are related to the hydrologic and climate models. First, we assumed that the streams in Iowa were under naturalized flow conditions, but a few streams are regulated. This should be investigated in studies when trying to apply future streamflow simulations to engineering design as the regulatory actions are important to consider. Next, we selected the climate models based on how well they reproduced historical trends, magnitude, and distribution. However, uncertainty in CMIP5 projections due to original model spatial coarseness and bias-correction and temporal disaggregation processes exist and can contribute to the underestimation of extreme events (Demaria et al., 2016a). We explored the impact of model selection by rerunning our analyses using all available climates models, finding differences in the results over our primary analysis. We suggest that when providing flood projections for communities that a sensitivity analysis of model selection should be performed.

Our framework allows for further research into the determination of changes in annual maximum discharge important for water resources planning and design at the community scale. With the availability of updated high-resolution CMIP6 GCMs (Eyring et al., 2016), it will be important to run future analyses with the most up-to-date climate projections. Once this is completed, an exhaustive analysis should be conducted to create flood frequency guidelines which incorporate nonstationarity and localized projections. Hydrologic simulations forced with future climate projects can be combined with nonstationarity distribution fitting methods (e.g., Villarini et al. 2009; Tan and Gan 2015; O'Brien and Burn 2014) to help develop these methods. Finally, we acknowledge the omnipresent and unavoidable uncertainty in the framework, both in the future forcings as well as the hydrologic simulations.

Table 3-1. Climate model selection criteria based on comparison of historical peak flow between models and PRISM forcings. The values represent the number of sites where the statistical metrics are not significant. A 90% threshold was utilized, and all three tests must be above it for selection. The symbol "*" identifies the models used for the primary analyses.

Global Climate Model	Kendall's Tau matching (%)	Mann-Whitney nonsignificant (%)	Kolmogorov-Smirnov nonsignificant (%)
access1-0*	98	93	91
canesm2	96	94	83
cmcc-cm*	98	98	94
cnrm-cm5*	91	97	94
csiro-mk3-6-0*	98	97	92
gfdl-cm3	98	87	85
gfdl-esm2g	86	90	85
gfdl-esm2m*	92	97	90
hadgem2-cc	98	98	89
hadgem2-es*	97	98	92
inmcm4	98	93	86
ipsl-cm5a-mr	97	73	79
miroc-esm*	97	97	93
miroc-esm-chem	98	96	87
miroc5*	98	96	95
mpi-esm-lr*	98	94	94
mpi-esm-mr	98	96	89
mri-cgcm3	97	97	86
noresm1-m	98	86	84

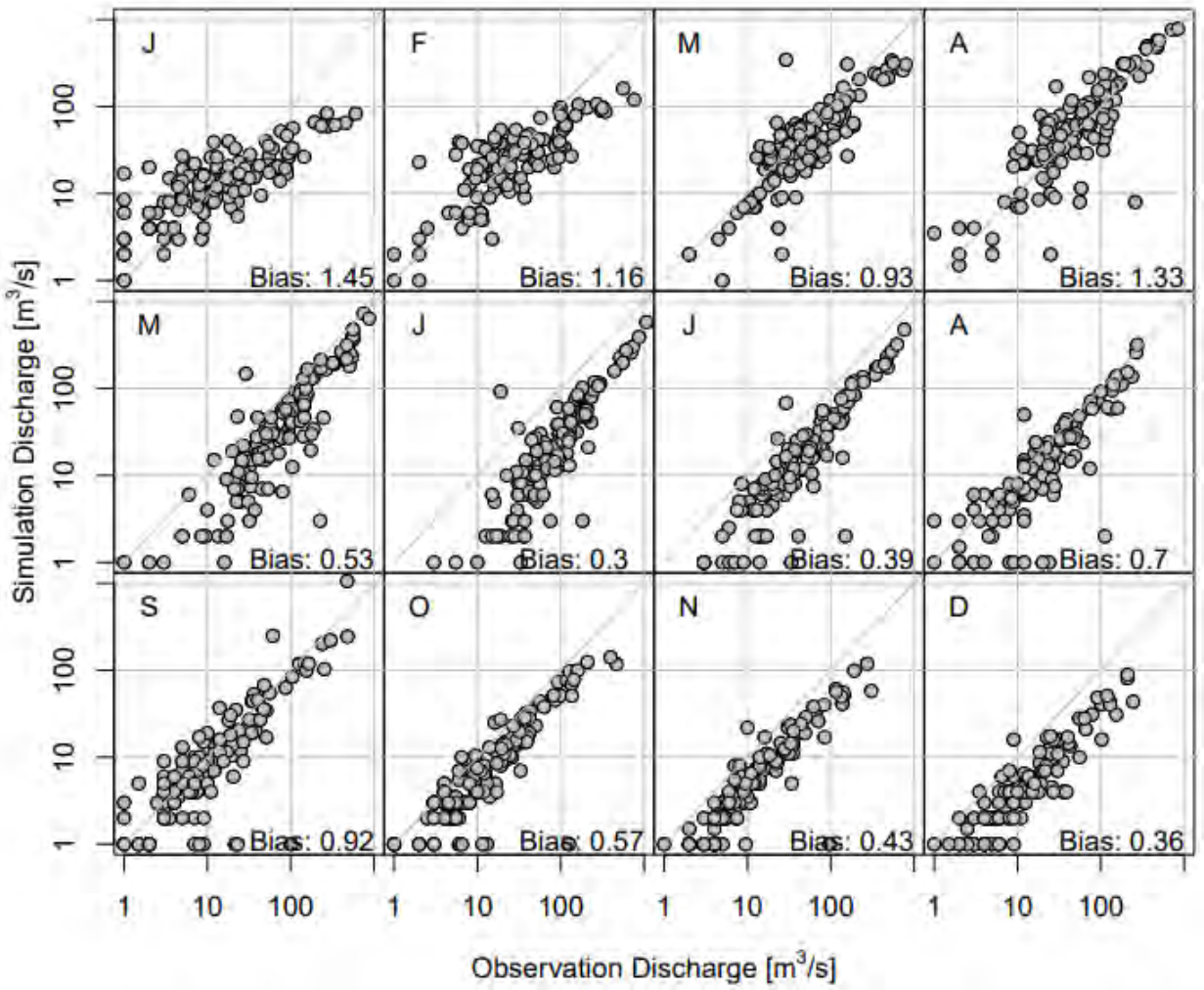

Figure 3-1. Ability of the HLM using daily precipitation, daily temperature, monthly ET derived from PRISM to reproduce the observed average monthly maximum discharges at 150 USGS gauges in Iowa from 1981 to 2020, where instantaneous gauge data are available. The bias reported is the average ratio of simulation discharge to USGS observation discharge for all 150 gauges. A bias above (below) 1 indicates simulations overestimate (underestimate) discharge compared observations. Months are indicated by the letters in the upper left corner.

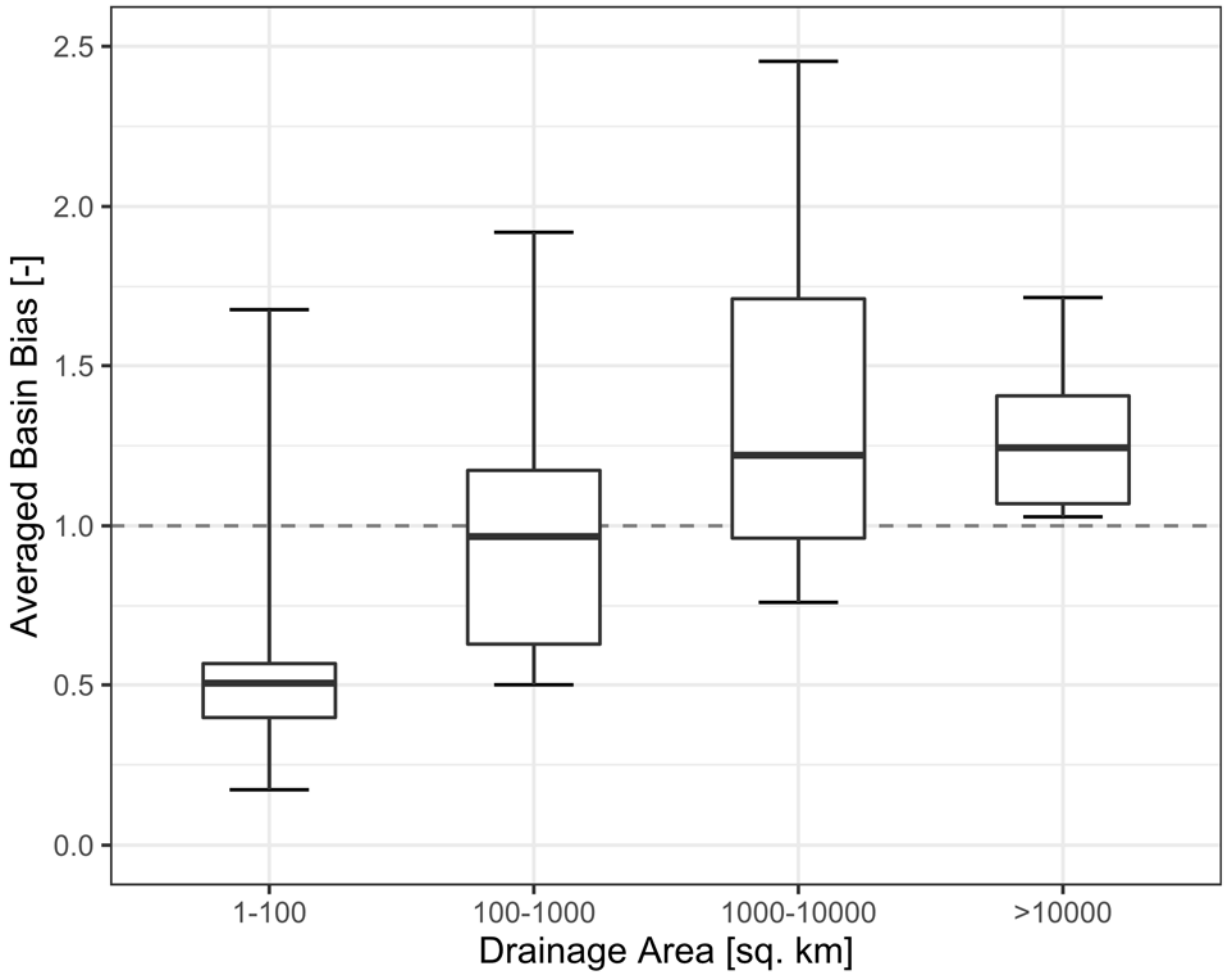

Figure 3-2. Average bias for USGS locations with over 30-years of annual maxima data to assess the HLM model's ability to recreate observed peaks stratified by drainage area. The dashed line represents a bias of 1 and the whiskers represent the 5th and 95th percentiles. A bias above (below) 1 indicates simulations overestimate (underestimate) discharge compared observations.

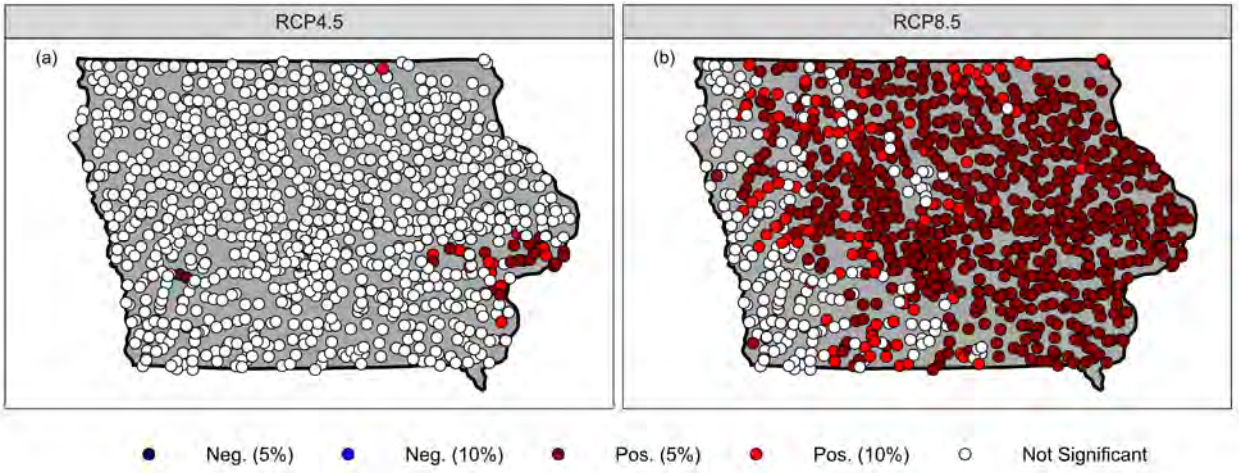

Figure 3-3. Trends in annual maximum discharge from 2006 to 2099 determined by the Mann-Kendall trend test. The left (right) panel shows the results from simulations under the RCP4.5 (RCP8.5) emission scenario. Blue (red) circles indicate a negative (positive) trend in annual maximum discharge. White circles are points where the trend is not statistically significant at the 10% level.

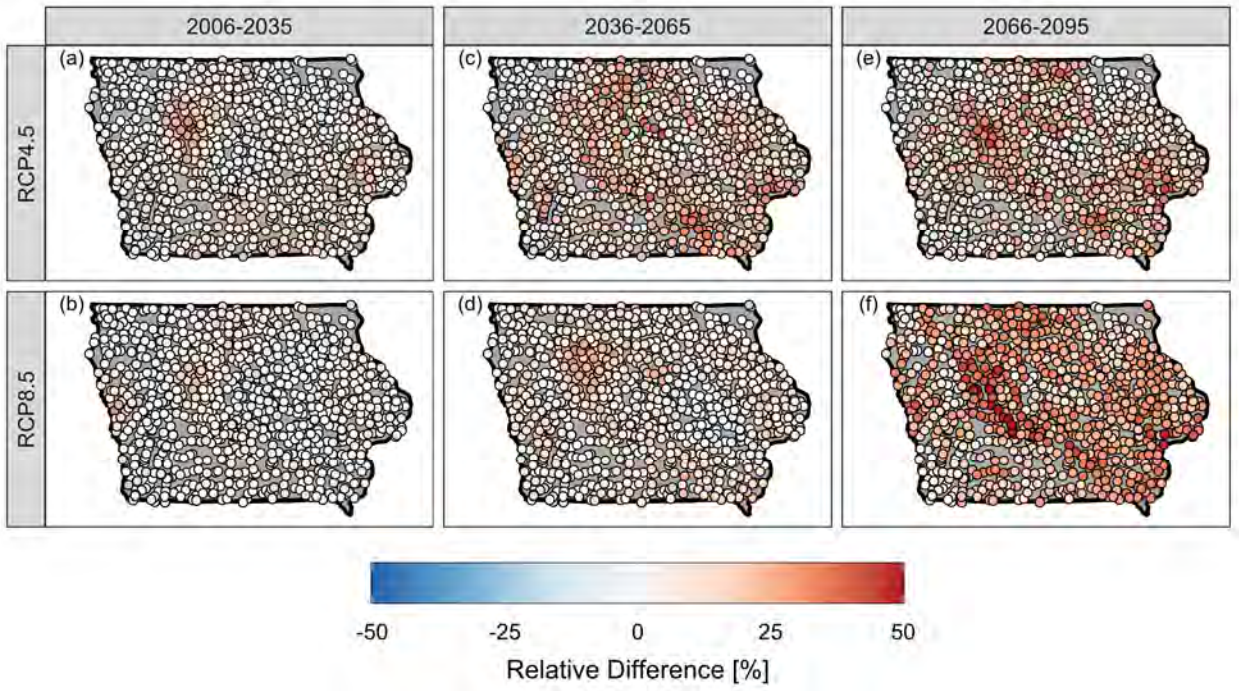

Figure 3-4. Relative difference for median annual maximum discharge between historical (1976-2005) and future 30-year periods of 2006-2035 (a, b), 2036-2065 (c, d), and 2066-2095 (e, f). Each column represents one emissions scenario (RCP4.5 in the left column and RCP8.5 in the right one). Values greater than zero indicate an increase in the median annual maxima compared to the historical median.

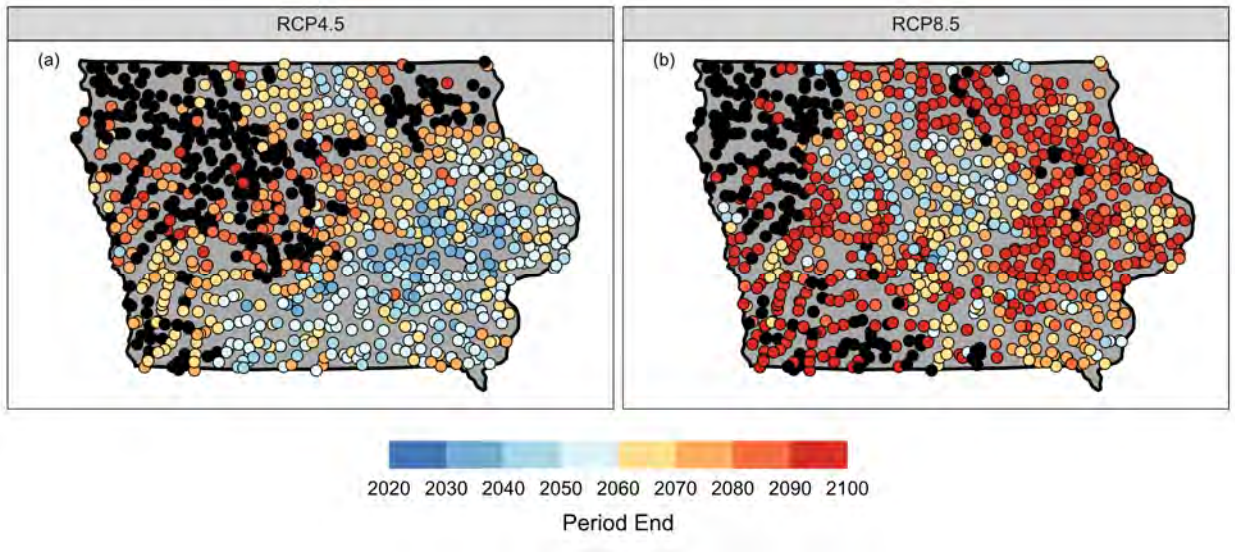

Figure 3-5. Time of emergence where the 50-year distribution of annual maximum discharge is significantly different ($p<0.05$) compared to the reference distribution from 1956-2005. The years associated with the color bars indicate the end of the comparison period (i.e., "2060" represents the period from 2011 to 2060). Black points indicate sites where the p-value never went below 0.05 for the entire study period.

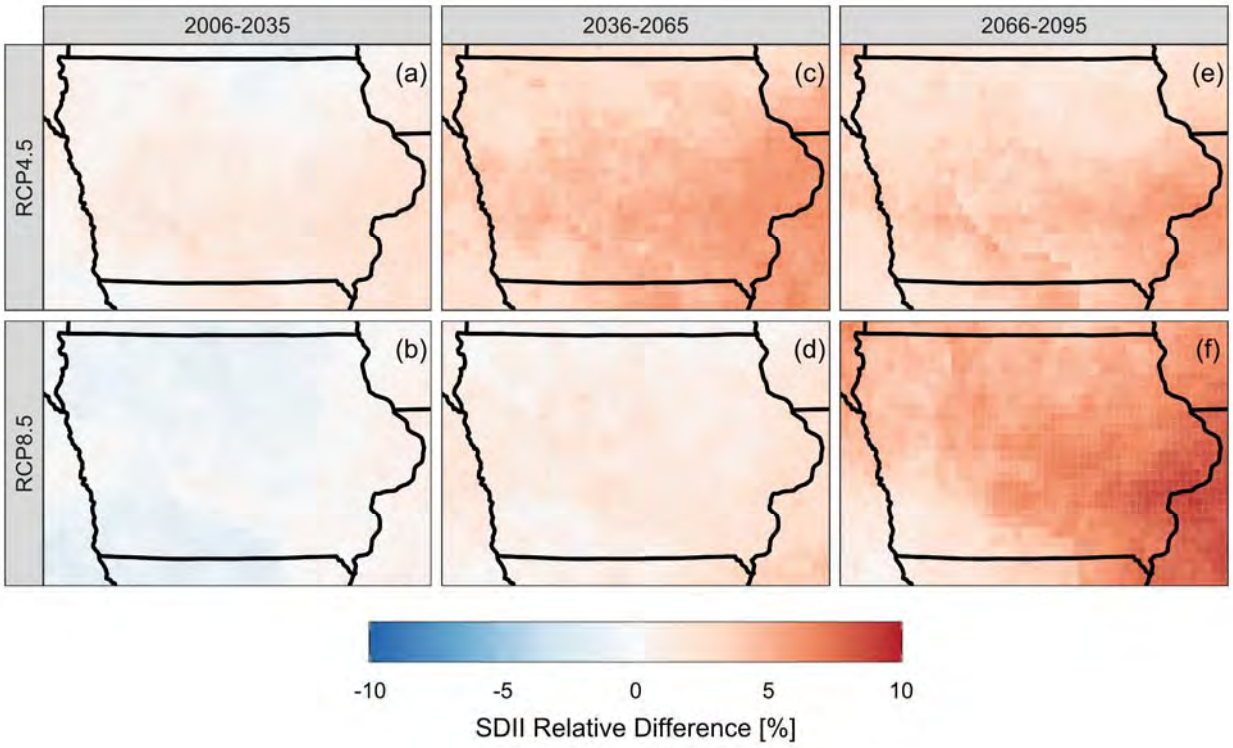

Figure 3-6. Relative differences in simple daily precipitation intensity index (SDII) between historical (1976-2005) and three future 30-year periods. Cells with values less than zero (blue) represent a decrease in the intensity of daily wet day precipitation (> 1 mm). Cells with positive values indicate an increase in the intensity of daily precipitation for wet days.

4. Evaluation of CMIP6 HighResMIP for hydrologic modeling of annual maximum discharge in Iowa[4]

4.1 Introduction

Flooding is one of the most common natural disasters worldwide, causing millions of dollars in damage and numerous fatalities every year (e.g., Centre for Research on the Epidemiology of Disaster, 2015). These damages (e.g., structural damages, affected buildings) to communities from flooding are projected to increase (e.g., Bates et al., 2021; Dottori et al., 2018) as climate change is expected to exacerbate extreme precipitation, which in turn will change the variability of flood events (Wasko et al., 2021). To improve our preparedness against this natural hazard, the most up-to-date climate projections are needed to make informed decisions to alleviate flood risks for communities. This means utilizing the newest version of climate models as inputs for hydrologic modeling, such as those from the Coupled Model Intercomparison Phase 6 (CMIP6, Eyring et al., 2016). Another aspect critical to providing hydrologic information at the community scale is the availability of high spatial and temporal resolution for the climate forcings (e.g., precipitation) required to simulate realistic hydrologic processes (e.g., Berne et al., 2004). Typically, global climate models (GCMs) are run at coarse spatial resolutions (e.g., ~1 degree or ~100 km), requiring the use of downscaling techniques to develop regional climate projections (e.g., McSweeney et al., 2014; Schmidli et al., 2006). However, there have been significant efforts as part of CMIP6 to provide climate forcings by running the GCMs at high spatial resolutions. Specifically, the CMIP6 High Resolution Model Intercomparison Project (HighResMIP, Haarsma et al., 2016) provides outputs from climate model simulations at spatial resolutions of 25-50 km, allowing for regional climate analysis to be conducted directly from the model outputs.

Recent evaluation of the HighResMIP suite has shown that increasing the spatial resolution of climate models leads to an improved representation of many processes. Multiple studies (e.g., Bellucci et al., 2021; Chang et al., 2020; Lohmann et al., 2021) have shown that the increased spatial resolution of these models also improves the multiscale climate interactions. Evaluation of precipitation indicates that the magnitude of extreme events (e.g., Wehner et al., 2021), seasonal totals (e.g., Hariadi et al., 2021; Moreno-Chamarro et al., 2021), and annual totals (e.g., Zhao, 2020, 2022) is projected to increase over much of the globe. The majority of studies evaluating HighResMIP have focused on the examination of atmospheric rivers, tropical cyclones, and mesoscale convective systems (e.g., Balaguru et al., 2020; Feng et al., 2019; Gao et al., 2020; Grist et al. 2021; Huang et al., 2021; Judt et al., 2021; Kreussler et al., 2021; Li et al., 2021; Lopez-Parages & Terray, 2022; Priestley & Catto, 2022; Vanniere et al. 2020; Vidale et al., 2021; Yamada et al., 2021; Zhang et al., 2021; Zhao, 2022). These works have indicated that the representation of precipitation related to these phenomena have improved over previous climate model simulations. However, not all processes have been shown to be improved. For instance, an analysis of temperature across Europe from the HighResMIP models by Squintu et al. (2021) found little improvement compared to past climate simulations, highlighting a potential shortcoming of the new models. Currently, the impact of improved weather processes and their associated shortcomings from the HighResMIP suite have not been explored from a hydrologic perspective. Furthermore, hydrologic models typically require even finer spatial resolution (<12 km) of precipitation data (Quintero et al. 2022) than the resolution from the HighResMIP suite. This highlights an additional gap in the literature to assess bias-correction and downscaling methodologies on the HighResMIP precipitation products and their applications in hydrology.

[4] Adapted from: Michalek, A., G. Villarini, T. Kim, F., Quintero, W.F. Krajewski, and E. Scoccimarro, Evaluation of CMIP6 HighResMIP for hydrologic modeling, submitted to Water Resources Research, 2022.

Beside modeling the rainfall to runoff transformation processes using meteorological variables from GCMs (e.g., potential evapotranspiration, precipitation, temperature) through a hydrologic model, another way to obtain discharge from GCMs is through the routing of their runoff. The HighResMIP suite provides runoff from the models at a high spatial resolution, a topic that has received little attention in the literature. Most studies (e.g., Byun et al., 2019; Giuntoli et al., 2018; Naz et al., 2016; Villarini & Zhang, 2020) utilize runoff from climate model suites as a proxy for flooding to assess relative changes, trends, and uncertainty in the future. Review of past climate models from the Coupled Model Intercomparison Phase 5 (CMIP5) project indicate that runoff outputs have large uncertainties and are not advisable to be applied directly to water resource applications (Lehner et al., 2019). With improved representation of climate processes in the HighResMIP project, there is a need to assess if the hydrologic component has also improved regarding these concerns. Finally, it is important to assess if the generation of flood peaks for the historical and future periods are similar when utilizing runoff versus precipitation forcings to gain more confidence in the robustness of results.

The objective of this study is to assess the suitability of the HighResMIP forcings for hydrological applications by utilizing their outputs to simulate annual maximum discharges at the 1,000 river communities across the state of Iowa. We generate simulations in two ways. First by routing only runoff through our hydrologic model, and second by forcing our hydrologic model with precipitation, temperature, and potential evapotranspiration. We assess the ability of the native GCM forcings to match a reference simulation dataset that captures observations well. We then apply nine bias-correction and downscaling methods to improve the spatial resolution of precipitation for hydrologic application and quantify the improvements of performance in hydrologic simulations due to the corrections. Finally, we analyze the projections in annual maximum discharge based on their changes compared to the historical period and their respective interannual variability. We organize the chapter as follows. In the next section we define the study region with relevant information on landform features and gauge locations. In Section 4.3 we describe the data from the HighResMIP climate models, our hydrologic model, the methods used to bias correct and downscale precipitation, and statistical methods utilized for model assessment. In Section 4 we describe the results related to the GCMs' performance compared to historical precipitation and hydrologic simulations as well as analyze projected changes and variability in flood peaks. Section 4.5 summarizes the results and presents a discussion of the findings.

4.2 Study Region

We focus on the state of Iowa and all basins draining into the Missouri and Mississippi Rivers. The study area is agriculturally dominated with low relief and can be divided into seven distinct landforms (Fig. 4-1). The landscape has a high drainage density in the Southern Iowa Drift Plains, Loess Hills, Iowan Surface, Northwest Iowa Plains, and Paleozoic Plateau, each of which consist of a loess cover (Prior, 1991). The Des Moines Lobe has the poorest surface drainage (Prior, 1991). Iowa has a seasonally variable climate, with the largest amounts of rainfall occurring in summer and snow in winter. The combination of landscape and climate features has caused the region to be prone to significant flooding in the past few decades (e.g., Smith et al., 2013; Mallakpour & Villarini, 2015; Flanagan et al., 2020).

For our analyses, we utilize 119 U.S. Geological Survey (USGS) gauge streamgauges for model validation (described below) (Fig. 4-1). The drainage areas of the gauges range from 10 to 36,000 km^2.Table S4-1 provides the USGS gauge identification numbers along with the drainage area and a description of the nearby community and river. We select gauges which have at least 30 years of annual and daily data from 1981 to 2020. For our primary analysis we focus on annual maximum discharges at 1,000 river communities that have been the focus of the Iowa Flood Center (IFC) (Fig. 4-1).

4.3 Methodology

We use 15 high-resolution ensemble members from CMIP6 HighResMIP tier (Haarsma et al., 2016) to conduct hydrologic simulations to evaluate their possible use as forcing variables. The GCMs are selected based on the availability of specific forcing variables at a daily temporal scale. First, we consider daily runoff from ten ensemble members to examine its suitability for routing in our hydrologic model. Next, we choose 11 ensemble members with precipitation, temperature, and potential evapotranspiration outputs for historical and high-resolution future experiments as required for our hydrological model. The GCMs have spatial resolutions of either 25 or 50 km, with daily temporal scale for precipitation and temperature and monthly temporal scale for potential evapotranspiration. The historical experiments represent atmospheric simulations from 1950 to 2014. The future experiments represent a period from 2015 to 2050, following the RCP8.5 emission scenario. Table 4-1 provides the model names, ensemble members, and resolution for data used in this study as well as references for further information.

For a historical reference, we use precipitation, temperature, and potential evapotranspiration derived from the 4-km daily Parameter Elevation Regression on Independent Slopes Model (PRISM; Daly et al., 2002); we use the hydrologic simulations derived from PRISM inputs as reference. Additionally, we use PRISM-derived forcings from 1981 to 2014 as reference to conduct a bias-correction (precipitation, temperature, potential evapotranspiration) and downscaling (precipitation) on the GCM forcings. Only precipitation is downscaled, while temperature and potential evapotranspiration are taken as the statewide average. This is because of the current setup of the TETIS HLM (Quintero & Velasquez, 2022) as this version of the model is very computationally expensive to run with spatially distributed temperature and potential evapotranspiration, especially for 70 years of simulations. The IFC is currently working to improve the model so all inputs can be managed spatially (Krajewski et al., 2017). However, temperature and potential evapotranspiration do have spatial gradients across the Iowa and readers should note this modeling limitation.

To downscale and bias-correct precipitation, we employ nine methods consisting of (1) detrended quantile matching (dqm, Cannon et al., 2015); (2) empirical quantile mapping (eqm, Amengual et al., 2012); (3) generalized quantile mapping (gpqm, Gutjahr & Heinemann, 2013); (4) local intensity of scaling (loci, Schmidli et al., 2006); (5) mean variance adjustment (mva); (6) parametric quantile mapping (pqm, Piani et al., 2009); (7) power transformation (ptr, Leander & Buishand, 2007); (8) quantile delta mapping (qdm, Cannon et al., 2015); and (9) simple scaling. We utilize functions developed by the Santander Meteorology Group (2015) to accomplish this task. To achieve the bias-correction and downscaling we use the following procedure:

1. For a given day *i*, we pull together the precipitation *R* all the days that fall within a time window of ±15 days (i.e., a 31-day window) for all the years (*n*=34 because we focus on the 1981-2014 period). Therefore, we obtain a vector $\mathbf{R_i}$ that has 1054 values (i.e., 31×34) values for the i^{th} day (training dataset). We also collect the i^{th} day of the years in the historical and future periods (test dataset). This is done for PRISM and each GCM.
2. The training dataset is used to derive statistical differences between PRISM and each GCM and implement the nine bias-correction and downscaling methods above, and the test dataset is used to perform the respective corrections.
3. We use a leave-one-out cross validation (looCV) on the historical period, by training the bias-correction and downscaling approaches for all the values in $\mathbf{R_i}$ minus the 31 values for the j^{th} year, to examine the performance of each bias-correction and downscaling method.
4. These steps are applied to all ensemble members and bias-correction and downscaling techniques.

Next, we bias-correct temperature and potential evapotranspiration at the state level from the GCMs using the bias calculated with respect to PRISM from 1981 to 2014. For temperature, we use a two-week

moving window to take the average temperature representative of each day within a year and then average across all years from 1981 to 2014. Then we take the temperature bias as the difference between the GCM value and the PRISM value for each day. The value for each day is used to bias-correct the daily temperature from 1981 to 2050. For potential evapotranspiration, we take the average monthly potential evapotranspiration from the GCMs and divide it by the average PRISM-derived potential evapotranspiration from 1981 to 2014 to compute the bias. The potential evapotranspiration bias for each month is used to bias-correct the potential evapotranspiration values for the GCMs accordingly. The temperature and potential evapotranspiration bias-correction is computed on a per-climate-model basis.

For our hydrologic assessment, we focus on annual flood peak characteristics at 1,000 river communities across Iowa (Fig. 4-1). For the initial runoff-only analysis we utilize an HLM structure with only one storage, which represents the amount of water stored in the channels. The model is forced with daily runoff inputs from the GCMs and is routed using the nonlinear routing equation as described in Mantilla et al. (2006). The daily runoff values for a hillslope are taken from the nearest GCM grid cell.

For the primary analysis, we simulate streamflow with the TETIS version of HLM as presented in Quintero and Velasquez (2022). The theoretical background and conceptualization of HLM is comprehensively summarized in Mantilla et al. (2022). The HLM is a fully distributed hydrologic model, which is based on the decomposition of the landscape into hillslopes and channels. The model estimates runoff generation at the hillslope by simulating different processes. In the TETIS version of the model, snow accumulation and melting rates are derived from precipitation and temperature data, based on the degree-day method (Martinec & Rango, 1981; Rango, 1995). The model simulates precipitation losses in vegetation and soil pore macrostructure. Surface infiltration considers the hydraulic soil properties of the hillslope, and the conditions of frozen ground during the cold season. Deep percolation and groundwater losses are also modelled, considering the hydraulic conductivity of soil at the subsurface. Total hillslope runoff is aggregated in the river channel from the contribution of overland flow, interflow, and base flow. A nonlinear hydrologic routing model transports flow in the channels and considers the scaling of the geomorphologic characteristics of the river network. The model equations are written as a system of ordinary differential equations. A numerical solver that integrates the HLM benefits from the binary tree structure of the river network to solve the system of differential equations in an asynchronous manner, suitable for high performance computing environments (Small et al., 2013). In our implementation, the HLM outputs streamflow simulations at a daily time step and annual maximum discharge can be determined for all channel links in the river network. Projections of precipitation, temperature, potential evapotranspiration, and frozen ground conditions are forced into the hydrologic model. The spatial and temporal resolution of these inputs are key for the performance of the model (see Quintero and Velasquez (2022) for more details).

We perform four sets of hydrologic simulations. Our reference simulations are conducted from 1981 to 2014 using PRISM forcings (described previously). For our climate model analysis, we run an initial set of simulations that uses the runoff directly from the GCMs and routed through our model to generate annual maximum discharge; next, we generate annual maximum discharge using native forcings (e.g., potential evapotranspiration, precipitation, temperature) from the GCMs from 1981 to 2014. Additionally, we run the hydrologic simulations using bias-corrected and downscaled GCM forcings to assess the improvements associated with the different downscaling methods. Finally, we conduct future simulations from 2015 to 2050 based on runoff routing as well as for selected bias-correction and downscaling precipitation methods with bias-corrected temperature and potential evapotranspiration.

For the evaluation of hydrologic simulations, we use the following analysis techniques. For the primary PRISM reference simulations, we compare the simulated annual maximum discharges to observed data from USGS stream gauges (Fig. 4-1) to highlight the suitability of the HLM for this application. We accomplish this by comparing the empirical cumulative distribution functions (ecdfs) between observed and simulated annual maximum discharges utilizing the two-sample Kolmogorov-Smirnov (KS) test (e.g.,

Massey, 1951). The KS test is a non-parametric test with the null hypothesis that the ecdfs are the same, while the alternative hypothesis is that the ecdfs are statistically different from each other.

Furthermore, we assess the bias between the simulated and observed annual and monthly maximum discharges from 1981 to 2021. We define bias as:

$$Bias = \frac{Q_{USGS}}{Q_{PRISM}} \quad (4\text{-}1)$$

where a bias greater (smaller) than 1 indicates an over-(under-) estimation of annual (monthly) maximum discharge compared to observations.

To examine the performance of the HighResMIP runoff and precipitation simulations (native and bias corrected) during the historical period, we apply the KS test as previously described, with the PRISM-forced HLM simulations at the 1,000 river communities as reference. This statistical analysis allows us to examine the entire distribution of annual maximum discharge from the HighResMIP outputs with respect to the reference.

To examine the projected changes in annual maximum discharge, we apply the Mann-Whitney U-test (Mann & Whitney, 1947) at the 1,000 river communities at a 5% level. The test is a nonparametric version of the two-sample t-test and allows determining if the central tendencies in the two populations are equal. More specifically, this test allows us to determine if the central tendencies of the distributions of annual maxima shift in the future (2017-2050) compared to the historical period (1981-2014). We examine the relative difference (percent change) to determine the direction of shift in the central tendency at each location compared to the historical period. We define the relative difference as:

$$RD = \frac{QM_{FUT} - QM_{HIST}}{QM_{HIST}} \quad (4\text{-}2)$$

where QM_{FUT} and QM_{HIST} are the median annual maximum discharge for the future and historical periods, respectively, at a river community for the selected climate model.

Finally, we examine the change in interannual variability between the historical (1981 to 2014) and future (2017 to 2050) simulations of annual maximum discharges for each ensemble member. We apply the F-test of equality of variances for each river community to determine if variability is changing in the future compared to the past. The null hypothesis for the F-test is that the two populations (historical and future) have the same variance. The test statistic is the ratio of the sample variances, which follows an F-distribution. We perform the test at a 5% significance level and present statistically significant results with an indication of an increase or decrease in future variability.

For the three statistical tests described above, we apply the false discovery rate (FDR) to control for type I in errors in our multiple hypothesis testing. Specifically, we apply the generalized FDR approach by Benjamini and Yekutieli (2001), which does not require assumptions to be made regarding spatial dependence. It should be noted that the test is conservative in its correction and results should be interpreted accordingly (Ventura et al., 2004).

4.4 Results

Before we present our analyses, we first want to highlight the ability of the PRISM reference simulation to reproduce observed USGS flood peaks from 1981-2021 at 119 locations by means of the KS test (Fig. 4-2). We find 90 sites (76%) have a p-value greater than 0.05, indicating that there is not enough evidence to reject the null hypothesis that the two distributions are the same across most of Iowa, and only 29 sites have a statistically significant p-value (accounting for FDR). Examination of the bias in the

annual maximum discharges stratified by drainage area is provided in Fig. S4-1. The average annual maximum discharge bias for watersheds greater than 100 km^2 ranges from 1 to 1.25 (median line), meaning that there is an overestimation in the simulations compared to the observations. For watersheds smaller than 100 km^2, the median bias is 0.55, pointing to an underestimation for small watersheds. For monthly peaks, the average monthly bias (unitless) ranges from 0.6 in June to 1.9 in September but is close to 1.0 for most months (Fig. S4-2). The performance of our model setup is similar to simulations with hourly precipitation in Quintero et al. (2020) and Michalek et al. (2022) for Iowa. Therefore, these results indicate that the reference simulations are generating distributions of flood peaks comparable to observations, supporting their suitability as a control for our experiments, while at the same acknowledging that there are biases in the magnitude of the annual flood peaks especially for small (<100 km^2) and large (>10,000 km^2) basins.

After model validation, we ask if the runoff available from the GCMs can be used directly as input into a routing model to generate annual maximum discharges. Fig. 4-3 shows the results of the KS test when comparing the distribution of annual maximum discharges from 1981 to 2014 between the reference simulations and the simulations using runoff. As shown in Fig. 4-3, the four HADGEM3-GC31 ensemble members and CNRM-CM6-1-HR-r1i1p1f2 can capture the distribution of the reference simulation well for over 99% of river communities at the 5% level (accounting for FDR). The remaining CNRM-CM6-1-HR ensemble members along with the EC-EARTH3P-HR ensemble members do not perform as well but still have over 80% of the communities with nonsignificant p-values. The BCC-CSM2-HR model performs the worst across all communities, with no ability to capture the distribution of annual maximum discharges anywhere in the state, likely because the magnitude of runoff values routed through the HLM directly from this GCM is too small to generate an appropriate hydrologic response. Based on our analysis, three of the five climate models (5/10 ensemble members) can generate annual maximum discharges comparable to our reference across Iowa, while another four ensemble members exhibit some skill in reproducing an adequate hydrologic response. No generalization of the viability of runoff from the HighResMIP suite beyond Iowa can be determined and performance should be examined on a per-model basis.

While routing runoff directly is one method to generate annual maximum discharges, we also want to assess the ability of utilizing other HighResMIP forcings (precipitation, temperature, and evapotranspiration) to estimate the discharge values. We calculate the p-value of the KS test between hydrologic simulations modeled with native climate model inputs and PRISM inputs (reference) from 1981 to 2014. All of the 11 total ensemble members (seven climate models), have zero communities with a significant p-value (<0.05), which indicates that there is not enough statistical evidence to say that they are not able to capture the reference distribution of annual maximum discharges for the study sites. For this reason we do not show a figure of the results. In comparison with the runoff simulations, the precipitation-based simulations are better at capturing the distribution of annual maximum discharge across Iowa. However, the forcings from HighResMIP still have biases that are masked within this hydrologic simulation set. We find that the raw precipitation from the GCMs before bias correction and downscaling is overestimated in the winter and fall seasons and underestimated in the summer (Fig. 4-4, left column). Furthermore, seasonal biases in temperature (Fig. S4-3) and potential evapotranspiration (Fig. S4-4) exist across the study region. For temperature, the bias on average ranges from an overestimation of 2°C to 6°C in the summer season and an underestimation of 4°C to 6°C in the winter season (Fig. S4-3). For potential evapotranspiration, the bias ranges from 0.5 in the summer months to 1.5 in the spring months (Fig. S4-4). The combinations of these biases are likely providing “the right answer for the wrong reasons” in terms of model performance and need to be removed.

To improve the suitability of the HighResMIP forcings for hydrologic simulations, we correct the bias in the GCM-derived forcings and downscale the precipitation to a spatial scale more suitable for hydrologic simulations. First, we examine the performance of the bias-correction and downscaling methods on the

precipitation of the 11 climate models/members based on leave-one-out cross validation (looCV). Fig. 4-4 shows the results for eqm-corrected precipitations in terms of daily bias for the GCMs. Fig. S4-5 to S4-12 provide the same information for the other eight downscaling methods. The eqm method centers the median daily bias around 1.00 for the entire year and removes the seasonality observed in the original data, leading to more unbiased conditions (Fig. 4-4). The same performance can be observed for mva (Fig. S4-5) and ptr (Fig. S4-6) methods. Furthermore, their performance is also better than the other six techniques (Figs. S4-7 - S4-12) and this is why we focus on these three methods in the rest of the chapter. Overall, the average daily bias for each of the three methods across GCMs ranges from 1.08 to 1.12 compared to values from 0.92 to 1.32 for the original outputs, highlighting the ability of these methods to remove daily precipitation biases.

Next, we focus on the performance of our hydrologic simulations forced with the bias-corrected and downscaled precipitation, bias-corrected temperature, and bias-corrected potential evapotranspiration. Fig. 4-5 shows the results of the KS test between historical (1981-2014) simulated annual maximum discharges with the bias-corrected HighResMIP forcings and the PRISM reference stratified by bias-correction method. The plot indicates the percentage of the 1000 river communities which have a nonsignificant p-value at the 5% level. Our results highlight that the dqm, eqm, mva, ptr, and qdm bias-correction and downscaling methods translate well hydrologically, with all climate models performing nearly identical to the reference simulations. The method that performed the worst is gpqm, with only three ensemble members having above 75% of the communities with a nonsignificant KS test result. A final takeaway from this comparison is that the FGOALS model performed the worst, with its simulations not reaching the 95% threshold for three downscaling methods.

Next, we assess the impact of the bias-correction and downscaling methods (eqm, mva, and ptr methods) on projections in annual maximum discharge as well as compare results between precipitation- and runoff-based projections. To begin, we apply the Mann Whitney U-test to compare the central tendencies of the distributions of annual maximum discharges between the historical (1981-2014) and future (2017-2050) periods (34 years) for each community based on the eqm, mva, and ptr downscaling. We find no communities with a statistically significant ($p<0.05$) shift in the central tendencies for the future compared to the historical period for any of the three downscaling methods (figure not shown). The same results are found for the runoff-based simulations across all respective ensemble members (figure not shown).

To highlight the results of flood projections obtained from HighResMIP forcings, we now examine future changes in the median (Eq. 4-2), even though they are not statistically significant. Fig. 4-6 provides the relative difference between future and historical medians for eqm-based simulations with mva- and ptr-based simulations provided in Figs. S4-13 and S4-14, respectively. For eqm simulations, 65% of communities on average have a relative difference in the median greater than zero indicating an increase in median future flood peaks compared to the historical experiments. MPI-ESM1-2-XR-r1i1p1f1 had the least number of communities with a future increase (354 sites) while HADGEM3-GC31-HH-r1i1p1f1 had the most (921 sites). On average 42% of the communities show greater than 25% increase in the median flood peak across Iowa. Only 25% of the communities show a projected decrease in the median and only 6% have a negative difference greater than 25%. Furthermore, the average relative difference per ensemble member ranges from -0.09 (MPI-ESM1-2-XR-r1i1p1f1) to 0.31 (HADGEM3-GC31-HH-r1i1p1f1). When taking the yearly median flood peak across the first ensemble member of each model (Fig. 4-6, GCM Median panel), 930 sites have a relative difference greater than zero, but the average projected increase across communities is only 16%. For this approach, the signal is more consistent across Iowa compared to other members as the range in projected differences is smaller. We find similar patterns in the projected changes of the median as well for mva (Fig. S4-13) and ptr simulations (Fig. S4-14). Our results indicate a projected increase in the median annual maximum discharge for most of Iowa according to the precipitation-based simulations.

For a comparison of future changes with runoff and precipitation derived simulations, we utilize the three ensemble members of the HADGEM3-GC31-HM suite. We present the relative difference (Eq. 4-2) for eqm-corrected precipitation-based and native runoff-based projections in Fig. 4-7. A comparison of the results for each modeling approach indicates that the runoff-based simulations show less of an expected increase in annual maximum discharge across Iowa. For member r1i1p1f1 (Fig. 4-7, left column), the runoff simulations show, on average, a ~25% decrease in the median whereas the eqm-precipitation simulations show a 25-50% increase (Fig. 4-7, top panel). A similar pattern can be observed in member r1i2p1f1 as well (Fig. 4-7, 2nd column). The r1i3p1f1 member is the most similar between the two methods, with similar projected changes observed in the southeastern portion of Iowa. The median across members (Fig. 4-7, 4th column) highlights the stark difference in projections with runoff-based results showing a less drastic expected increase for most of the communities compared to a larger projected increase for the precipitation-based projections. Projected changes for runoff-based simulations for all available climate models can be found in Fig. S4-15.

For our final analysis, we examine how variability is expected to change in the future. Fig. 4-8 shows the results of the F-test for eqm-based simulations. On average, across all models and ensemble members, 19% of the communities have a statistically significant F-test result, corresponding with an increase in future flood peak variance compared to historical one. Specifically, four of the eleven ensemble members (4 out of 7 models) have greater than 25% of the communities with a projected variance increase ($p<0.05$). These members include CMCC-CM2-VHR4-r1i1p1f1, FGOALS-F3-H-r1i1p1f1, HADGEM3-GC31-HH-r1i1p1f1, and HADGEM3-GC31-HM-r1i2p1f1. Only the ensemble members of EC-EARTH3P-r2i1p1f1, EC-EARTHP-r3i1p2f1, and MPI-ESM1-2-XR-r1i1p1f1 have less than 10% of the communities with a statistically significant projected increase in variance. When we conduct the F-test on the ensemble medians (Fig. 4-8, bottom right panel), 24% of the communities are projected to have statistically significant increases in discharge. No ensemble member has more than 0.5% of the communities with a projected decrease in interannual variance.

In comparison with mva and ptr-based simulations, we find similar projected changes in variance. Figs. S4-16 and S4-17 present the results of the F-test similar to Fig. 4-8 for these simulations. On average, 17% and 24% of the communities have a larger future variance compared to the historical one for mva and ptr, respectively. For eqm, mva, and ptr simulations, CMCC-CM2-VHR4-r1i1p1f1 is the only ensemble member with greater than 50% of the communities having a projected increase ($p<0.05$) in variance. A projected increase in variance for over 25% of the communities is found for 5 and 6 ensemble members (5 out of 7 models) with mva and ptr simulations, respectively. The similar ensemble members consist of CMCC-CM2-VHR4-r1i1p1f1, EC-EARTH3P-HR-r1i1p1f1, HADGEM3-GC31-HH-r1i1p1f1, and HADGEM3-GC31-HM-r1i2p1f1. When examining the projected variance based on the GCM median yearly values, 4% and 10% of communities have a projected increase in variance ($p<0.05$) for mva and ptr simulations, respectively. This is lower than what found with eqm-based simulations. Neither method has an ensemble member with more than 1% of the communities showing a projected decrease ($p<0.05$) in the variance. Finally, we provide the results of the F-test for the runoff-based simulations in Fig. S4-18. For all ensemble members, at least 84% of the communities do not indicate a significant change in the variance for the future compared to the historical period. Overall, our results indicate a projected increase in the interannual variability of flood peaks is detected for more of the communities across Iowa with eqm-based simulations compared to mva, ptr, or runoff.

4-5 Discussion and Conclusions

We evaluated the potential for the CMIP6 HighResMIP suite to be utilized for hydrologic simulations. We examined two approaches for the generation of annual maximum discharge data, consisting of runoff-based and precipitation-based hydrologic modeling. As a reference, we utilized simulations at the 1000

river communities based on PRISM forcings. The reference simulations were shown to capture the empirical cumulative distribution function at 76% of the gauge locations based on the KS test. We believe this is adequate for the study as the model performs well across Iowa and its performance is similar to previous studies (e.g., Michalek et al., 2022; Quintero et al., 2022). However, there are still biases in the model for small (<100 km^2, underestimation) and large (>10000 km^2, overestimation) watersheds that should be noted. Another limitation of our model set up is that only precipitation is a spatially varying input. Due to the model complexity and computational costs, the use of temperature, potential evapotranspiration, and frozen ground as spatially dependent is not currently feasible. Future work will be conducted to expand the capability of the model towards the modeling of larger spatial domains.

We started our analysis by evaluating if runoff directly from GCMs could be routed through a hydrologic model to estimate annual maximum discharges. Our results indicated the runoff-based simulations captured the reference annual maximum discharge distribution for all of Iowa for only five of the ten ensemble members assessed. Furthermore, the areas of Iowa in which poor performance occurred were not spatially consistent. For instance, one of the CNRM-CM6-1-HR ensemble members performed poorly for communities in the northern portion of the state, whereas for the EC-EARTH3P-HR members performance was poor in southeastern Iowa. Despite the good performance based on the KS test, we should interpret the results through the lenses of the results of hypothesis testing: we can say that we do not have enough statistical evidence to reject the null hypothesis that the distribution of annual maximum discharge based on the PRISM-forced model and the runoff routing simiulations are the same, which should not be misinterpreted as meaning that the two distributions are the same. This is especially true in light of the abovementioned biases Another limitation of this analysis with respect to runoff simulations is the size of the grid cell (50-km), which might not be suitable for small watersheds. To improve upon our limitations, we suggest the bias-correcting and downscaling of the runoff from the GCMS. Additionally, we did not assess the differences in runoff outputs among GCMs, suites, and past experiments (i.e., CMIP5 vs CMIP6) to determine if hydrologic performance has improved due to the higher-resolution climate models and comparative analysis should be performed in the future.

After examining the performance of runoff routing, we wanted to evaluate how well the native HighResMIP precipitation, temperature, and potential evapotranspiration worked in our precipitation-based hydrologic model (HLM). Our results showed that the estimated annual maximum discharges from 1981 to 2014 compared to PRISM reference simulations were able to recreate the empirical cumulative distributions across Iowa for all 11 ensemble members. Our results highlight that the climate forcings directly from the HighResMIP suite have the potential to be directly utilized for hydrologic purposes, although risk of biased estimation should not be ignored. Furthermore, the same careful interpretation of the KS test results as mentioned before should be applied. Additionally, we did not examine the role of convective precipitation from HighResMIP as we were looking at all precipitations events through a year to obtain the largest flood peak at a location. Recent literature (e.g., Judt et al., 2021; Moreno-Chamarro et al., 2022; Moreno-Chamarro et al., 2021; Vidale et al., 2021; Voldoire et al., 2019; Zhao, 2022) suggests that HighResMIP can represent mesoscale convective processes with some skill; however, we suggest an analysis of convective processes if event-base hydrologic modeling is conducted. Finally, we recommend that the discrepancies in climate model biases between precipitation and runoff be examined to improve runoff for hydrologic application.

A further investigation into input biases indicated that they exist in precipitation, temperature, and potential evapotranspiration for the native HighResMIP climate forcings compared to the PRISM references. Precipitation was overestimated in the models for the winter and underestimated for the summer, while temperature and potential evapotranspiration were found to be overestimated in the summer and underestimated in the winter. The results about the bias match similar results in other studies (e.g., Moreno-Chamarro et al., 2022; Moreno-Chamarro et al., 2021; Squintu et al., 2021; Tsartsali et al., 2022; Zhao, 2020, 2022). We did not investigate the sensitivity of the biases on hydrologic

simulations as the current set up is primarily precipitation-driven. The accuracy of the climate change signal in temperature and potential evapotranspiration is not known and we urge caution when interpreting the results. Therefore, we suggest that temperature and potential evapotranspiration be bias corrected and downscaled as well and used as spatially distributed inputs into hydrologic models.

Due to the precipitation biases, we decided to explore the viability of exploiting nine statistical bias-correction and downscaling methods to try to match our reference forcings and improve precipitation spatial resolution. Our results indicated that equivalent quantile mapping (eqm), mean variance adjustment (mva), and power transformation (ptr) performed the best for precipitation compared to PRISM. This translated directly to annual maximum discharge simulations, where the performance for the three methods across all 11 GCMs/members was nearly identical to the PRISM reference simulations. According to results from the KS test, the 11 ensemble members captured the reference distribution for all Iowa's river communities across the bias correction methods. Conversely, we found that the gpqm, loci, and scaling downscaling methods performed the worst. For these downscaling methods, poor performance was typically observed for communities along the Des Moines and Cedar River. As previously stated, careful interpretation of the KS test results should be applied. Finally, one of the primary assumptions of our modeling efforts, as previously noted, was that HighResMIP captured stratiform and convective precipitation to determine the annual maximum flood peak across Iowa. We did not examine how statistical bias correction and downscaling impacted precipitation for these two processes as it was outside the scope of this study.

Next, we explored the changes in annual maximum discharge projections based on the eqm, mva, and ptr corrected precipitation as well as compared to a subset of runoff-based results. We applied the Mann Whitney U-test to determine if the future annual maximum discharges are expected to differ from the historical simulations. Our results indicated that no communities are detected to have significant future changes when utilizing eqm, mva and ptr based simulations. Despite the lack of detected changes in the future based on the Mann Whitney U-test, we should interpret the results in a similarly to the KS test results previously discussed. To explain, we can say that we do not have enough statistical evidence to reject the null hypothesis that the two populations (future and historical) of annual maximum discharge are equal, which should not be misinterpreted as meaning the two populations are the same. For this reason, we examined the relative difference in the medians of the future and historical flood peaks for each community. This is a limited metric but allowed us a simple way to visualize projected changes in flood peaks based on HighResMIP models. We found that the median in flood peaks from most communities (65%) is projected to increase (not statistically significant) by ~20% on average. The river communities across Iowa which should be focused on by stakeholders are those along the Des Moines and Cedar River, with Des Moines and Cedar Rapids that are two major cities in Iowa. These areas showed a consistent increase in the median values across ensemble members and downscaling methods. Additionally, communities in southwestern Iowa within the Missouri River Basin and Middle Iowa River basin also showed this consistent increase. For runoff projections, we also did not detect future changes in annual maximum discharges, with less than 10% of Iowa river communities showing a significant change. Furthermore, a comparison of relative difference in the median flood peak values showed less of a change compared to precipitation-based simulations. The discrepancies between methods should be explored further. Finally, the lack of detection in future changes across both methods might be due to only examining the immediate future (2050) where the climate change signal might not be strong. We recommend a comparison of short-term results with other experiments that provide information to the end of the century to determine if this is the case.

Finally, we investigated how interannual variability was projected to change through the application of the F-test. For eqm-based simulations, four of the 11 members (4 out of 7 models) showed greater than 25% of the communities with a projected increase in variance compared to the historical period. Furthermore, at least 10% of the communities are projected to have a larger interannual variability in the

future for nine of the 11 members (5 out of 7 models). Simulations with mva-corrected precipitation detected less changes in variance in the future. However, ptr-based simulations produced changes in interannual variability similar to those of eqm results. We found that the choice of downscaling methods can impact the variability of flood peak projections and should be examined to determine the impact on the detection of the climate change signal. Projected increases in flood peak variance were commonly found in northern Iowa communities as well as for communities along the main reaches of the Cedar, Des Moines, Iowa, and East and West Nishnabotna Rivers. Lack of detected changes were found in southern central Iowa along the Thompson and Chariton Rivers. Additionally, we found the change in variance for runoff simulations was less prevalent across Iowa compared to precipitation-based projections. We suggest further exploration of the discrepancies in the changes interannual variability between simulations methods to determine if they are caused by the climate model forcings or the hydrologic modeling approach.

In conclusion, our study is one of the first examining the viability of CMIP6 HighResMIP forcings for hydrologic application. We recommend that hydrologic modeling with HighResMIP projections utilizes precipitation forcings instead of runoff outputs. Furthermore, bias-correction and downscaling of precipitation led to a good hydrologic performance in reproducing the distribution of annual maximum discharges across Iowa's river communities. Projected changes in the magnitude of annual maximum discharge are not detected for the immediate future across the state, but significant changes in interannual variability for the future are, with results differing between precipitation and runoff simulations. Investigation of the discrepancies between approaches should be the focus of future endeavors to quantify uncertainties in projections of flooding.

Table 4-1. Breakdown of climate models and ensemble members used in this study with forcing spatial resolution to generate annual maximum discharges. The runoff simulations highlight which models and ensemble members are used to force the HLM with daily runoff only. The forcing simulations refer to the models and ensemble members with daily precipitation, daily temperature, and monthly potential evapotranspiration to force the HLM. A "Yes" indicates an ensemble member was used in the respective simulation type.

Model	Ensemble Members	Reference	Spatial Res. (km)	Runoff Simulations	Forcing Simulations
BCC-CSM2-HR	r1i1p1f1	Wu et al. (2021)	50	Yes	No
CMCC-CM2-VHR4	r1i1p1f1	Scoccimarro et al. (2022)	25	No	Yes
CNRM-CM6-1-HR	r1i1p1f2	Voldoire et al. (2019)	50	Yes	No
	r2i1p1f2			Yes	No
	r3i1p1f2			Yes	No
EC-EARTH3P	r3i1p2f1	Haarsma et al. (2020)	50	No	Yes
EC-EARTH3P-HR	r1i1p2f1	Haarsma et al. (2020)	50	Yes	Yes
	r2i1p2f1			Yes	Yes
	r3i1p2f1			No	Yes
FGOALS-F3-H	r1i1p1f1	An et al. (2022)	25	No	Yes
HADGEM3-GC31-HH	r1i1p1f1	Roberts et al. (2019)	50	Yes	Yes
HADGEM3-GC31-HM	r1i1p1f1	Roberts et al. (2019)	50	Yes	Yes
	r1i2p1f1			Yes	Yes
	r1i3p1f1			Yes	Yes
MPI-ESM1-2-XR	r1i1p1f1	Gutjahr et al. (2019)	50	No	Yes

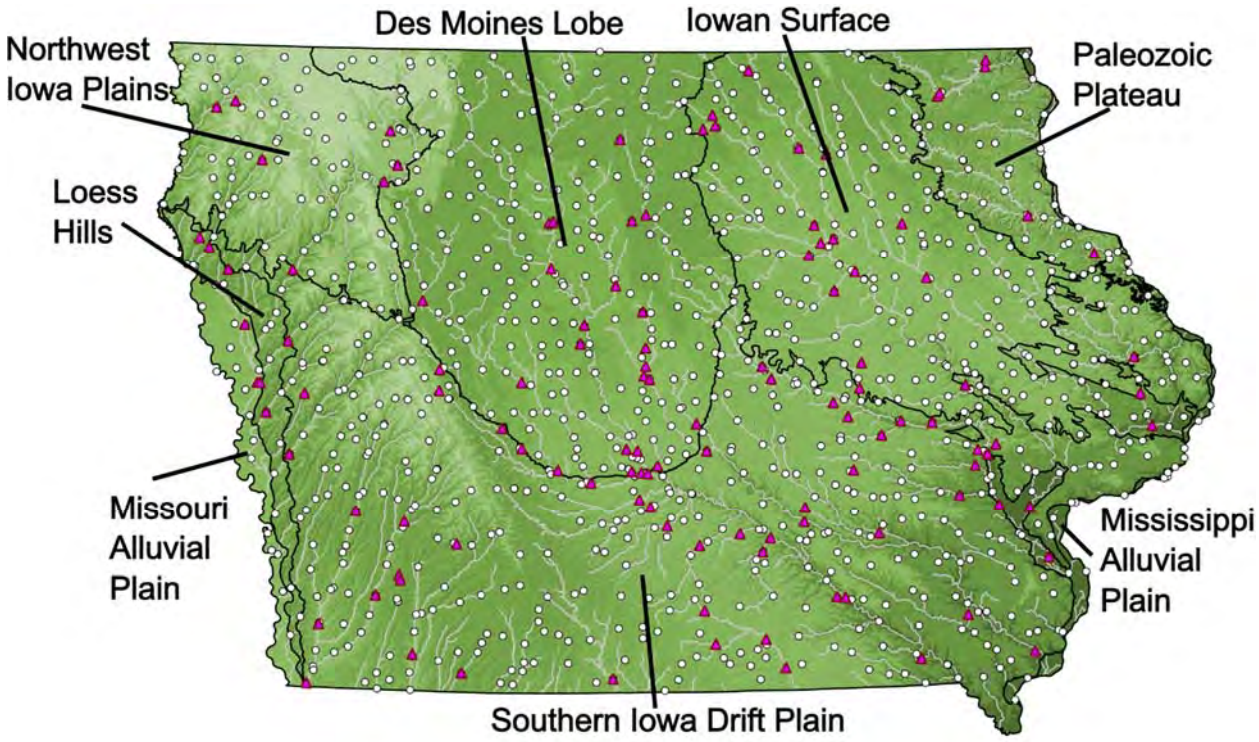

Figure 4-1. USGS gauge locations (triangles), river communities of interest (circles), major rivers (grey lines), and landform features for the study area.

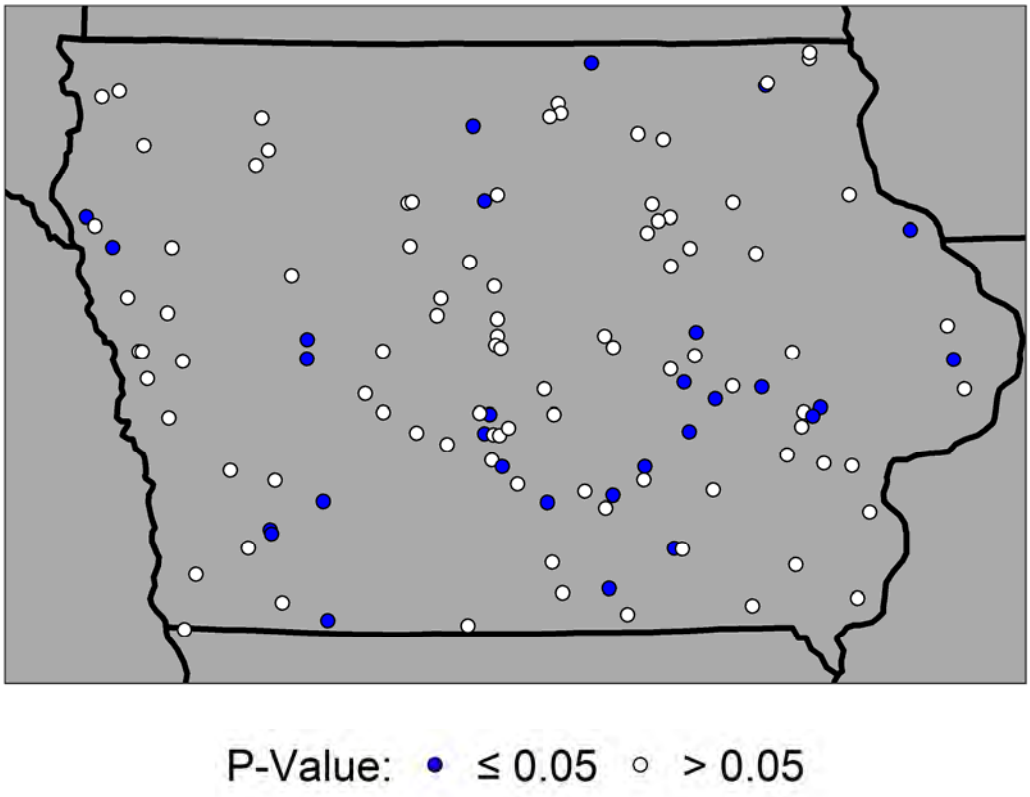

Figure 4-2. P-values of the Kolmogorov-Smirnov test for annual maximum discharge during the historical period from 1981 to 2020 between PRISM simulations and USGS observations. The blue circles are for locations in which the Kolmogorov-Smirnov test is rejected at the 5% level accounting for false discovery rates.

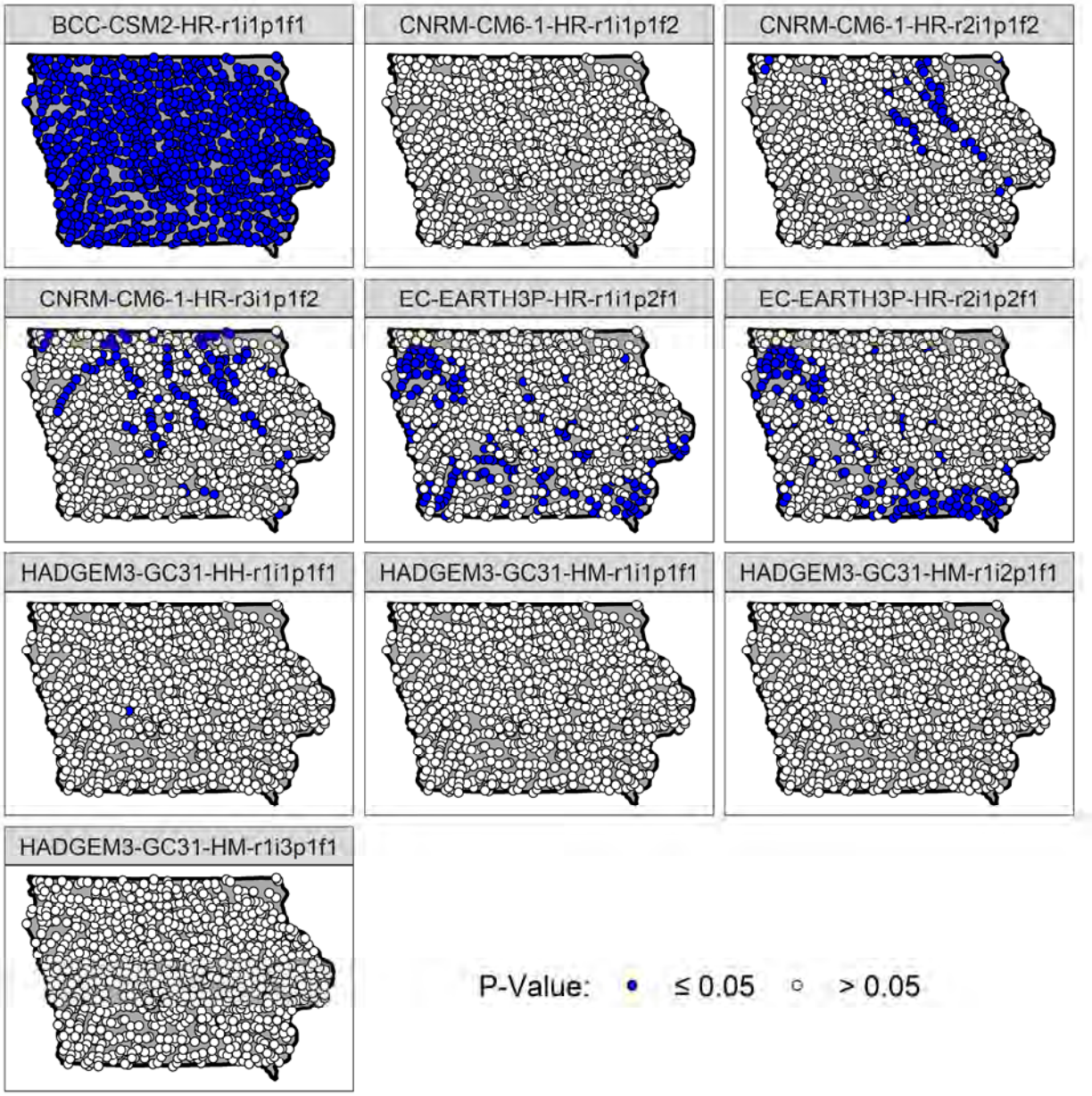

Figure 4-3. P-values of the Kolmogorov-Smirnov test for annual maximum discharge during the historical period from 1981 to 2014 between simulations using native HighResMIP runoff and PRISM reference. The blue circles represent locations in which the distributions of the reference and HighResMIP simulations are statistically different at the 5% level accounting for false discovery rates.

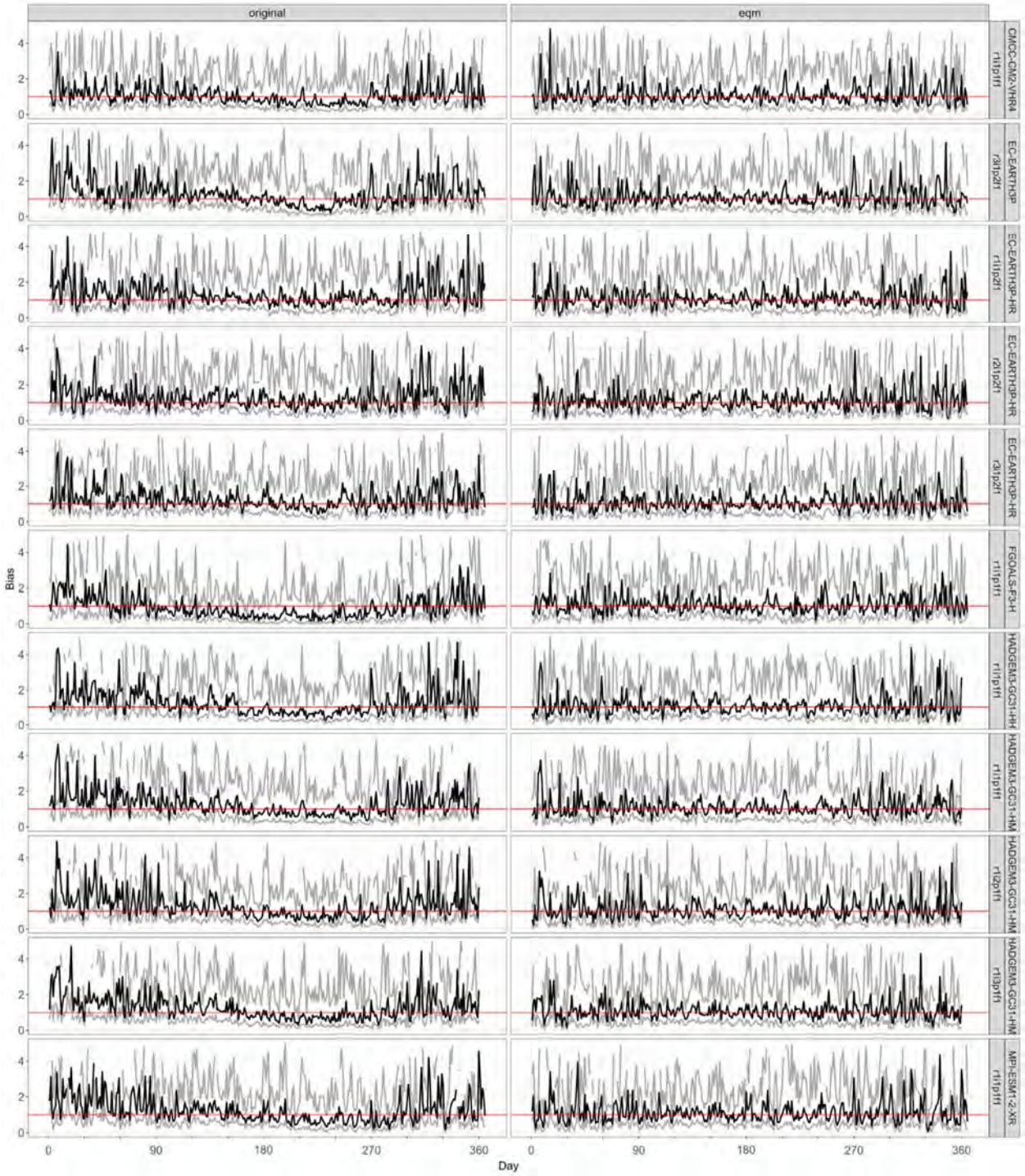

Figure 4-4. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and eqm corrected HighResMIP precipitation, respectively. The rows represent each of the climate models/members. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of 1 or perfect match with respect to PRISM precipitation.

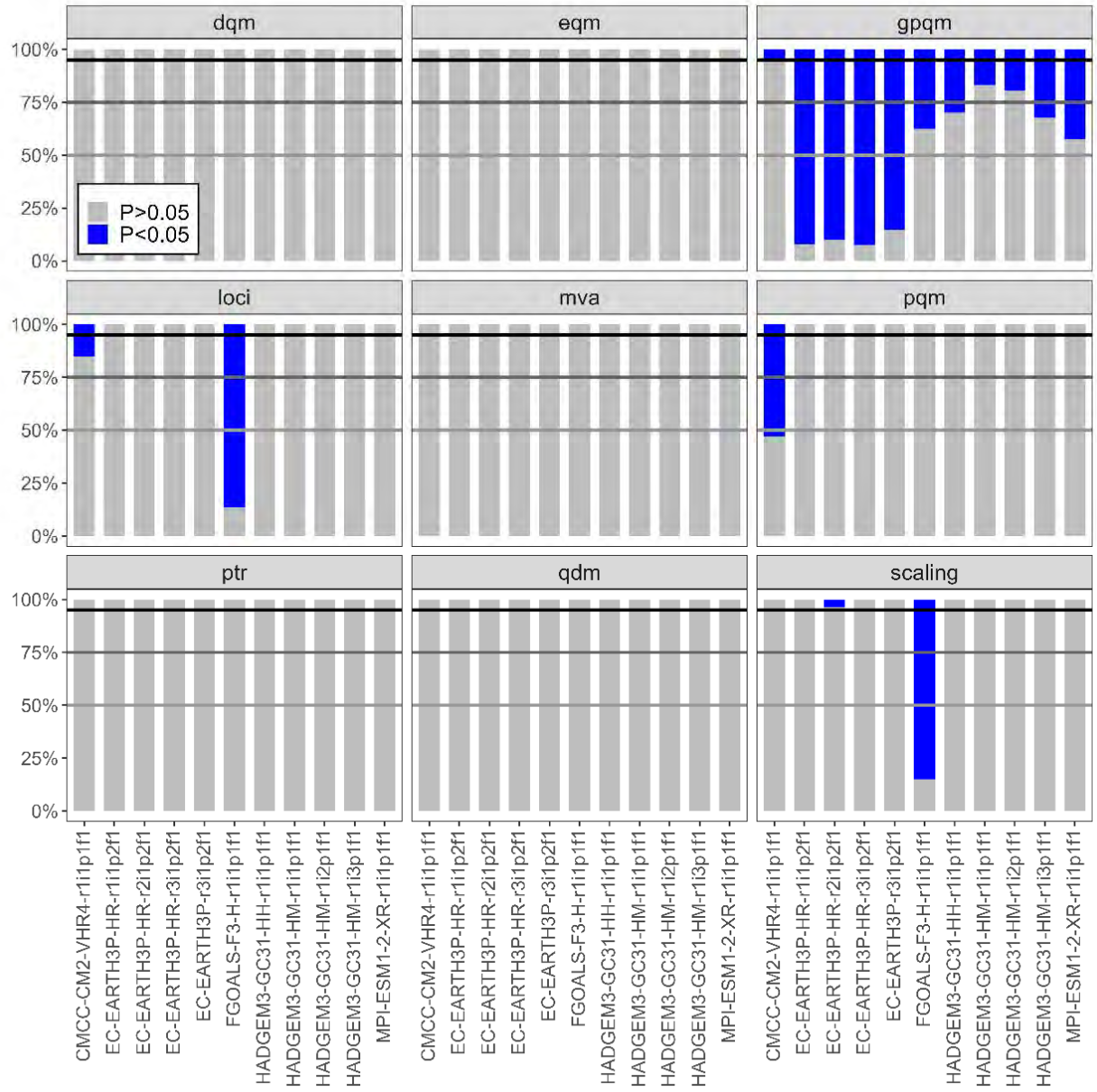

Figure 4-5. Percentage of 1000 river communities with a nonsignificant Kolmogorov-Smirnov test p-values for annual maximum discharge during the historical period from 1981 to 2014. The statistical test is conducted between the simulations from bias-corrected and downscaled HighResMIP precipitation and PRISM reference for each downscaling method. The blue bar represents statistically significant test results at the 5% level. The black and grey lines represent 90%, 75% and 50% of the river communities, respectively. The test results account for false discovery rates.

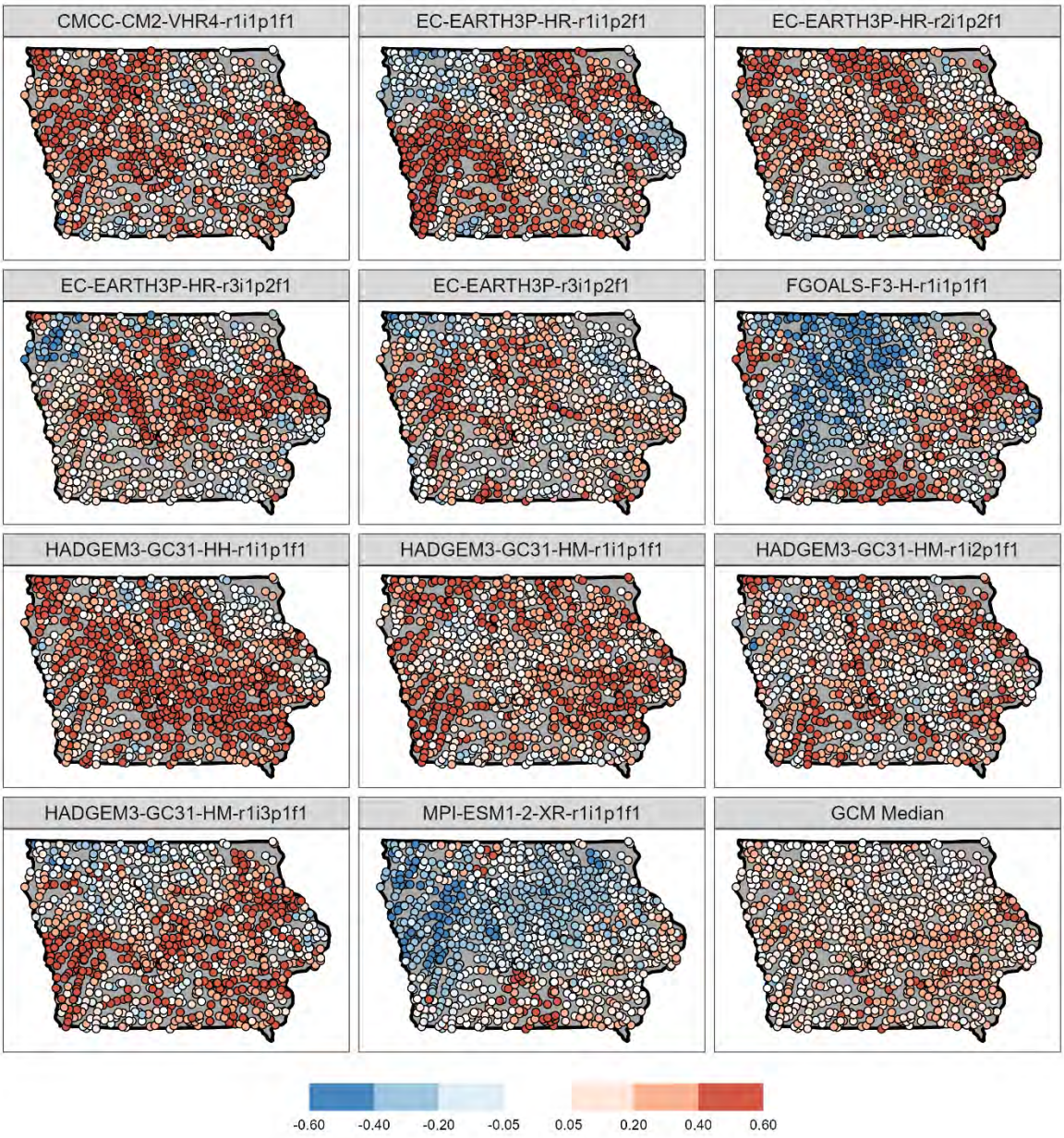

Figure 4-6. Relative difference (Eq. 4-2) in median annual maximum discharges between historical (1981-2014) and future (2017-2050) experiments from eqm corrected simulations. The GCM Median represents the relative difference based on the median of the first member (r1) of the three GCMs.

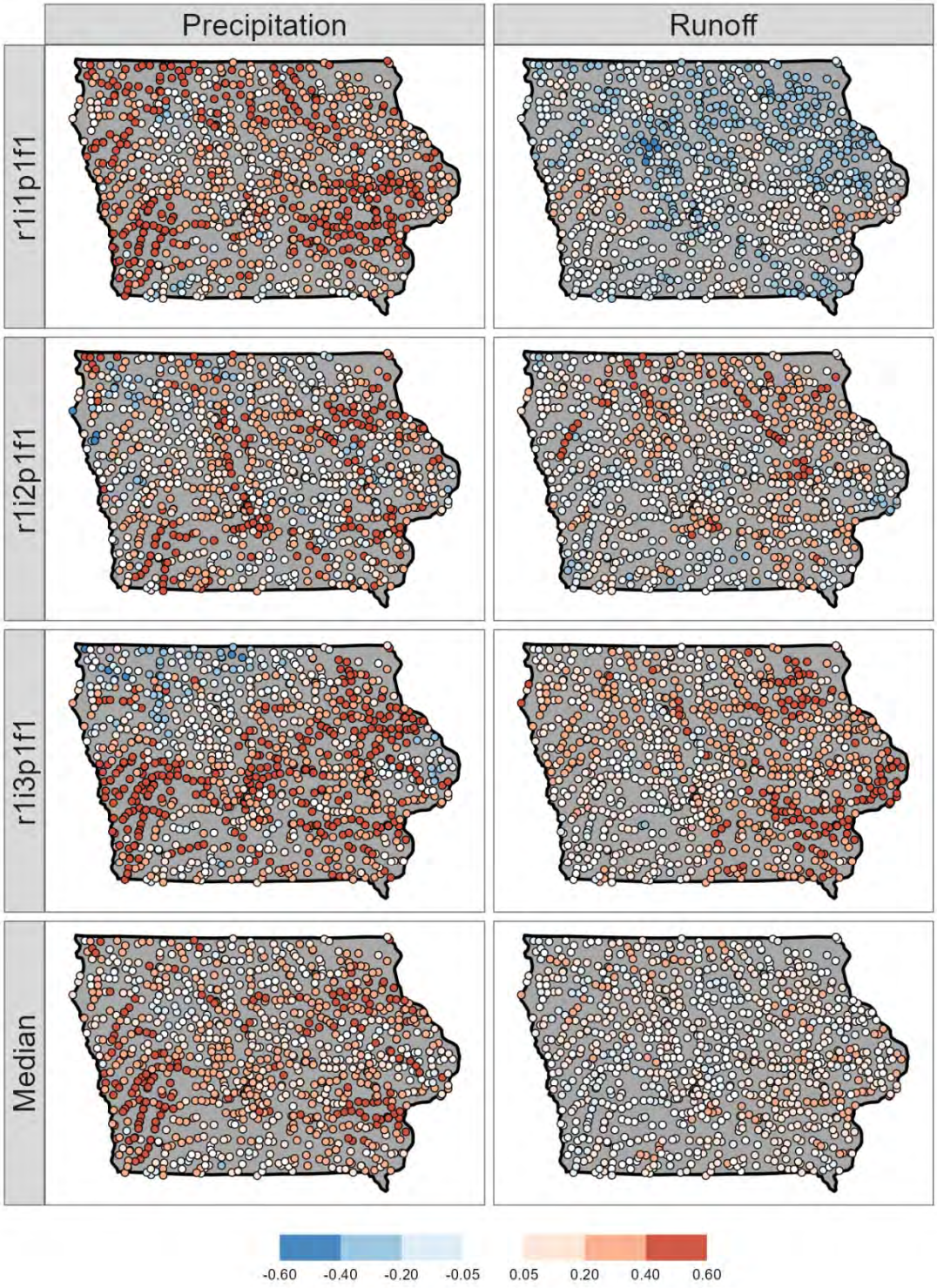

Figure 4-7. Relative difference (Eq. 4-2) in median annual maximum discharges between historical (1981-2014) and future (2017-2050) experiments for the HADGEM3-GC31-HM ensemble members. The columns represent annual maximum discharge generated with eqm-corrected precipitation and native runoff, respectively. The GCM Median represents the relative difference based on the median of the three ensemble members.

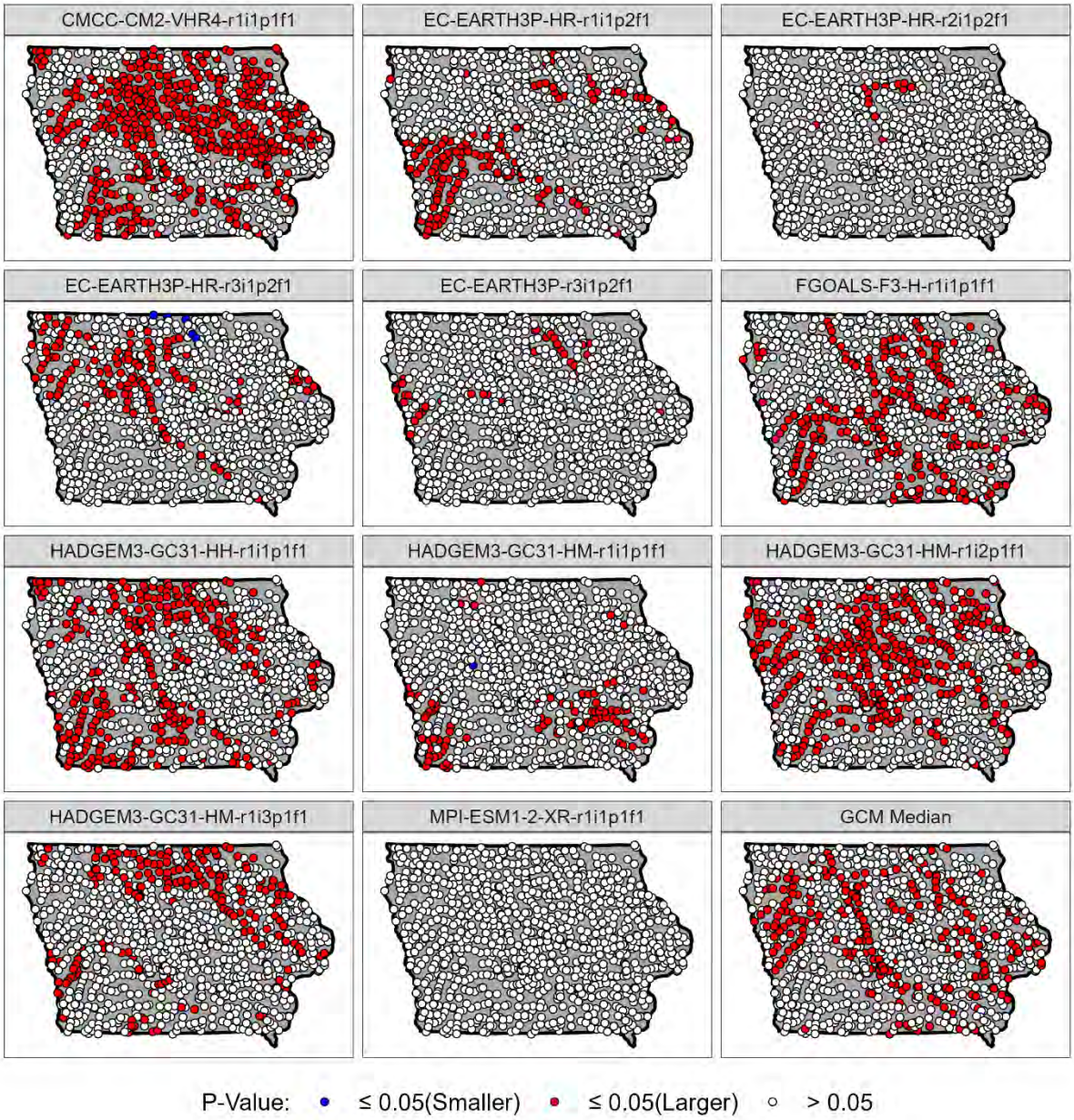

Figure 4-8. F-test p-values for annual maximum discharge between historical (1981-2014) and future (2017-2050) experiments for eqm-based simulations. The red (blue) represent a larger (smaller) variance in future simulations compared to the historical period. The test results account for false discovery rates.

5. Iowa Flood Frequency and Projections Tool (IFFP): A tool to project flood frequency estimates across Iowa[5]

5.1 Introduction

Climate change presents a significant challenge to water resources engineers and planners as it is expected to increase the likelihood of hydrometeorological extremes, including extreme precipitation and temperature (IPCC 2021). For many watersheds within the United States, this means that the annual maximum discharge (i.e., flood peaks) is expected to increase over time, leading to nonstationary conditions (Milly et al., 2008). Currently, federal guidelines for flood frequency analysis (see England Jr et al. (2019)) do not provide guidance on how to account for projected changes in the climate system, and new tools and approaches are needed to improve the resilience of water resources infrastructure (IPCC, 2022; Wasko et al., 2021).

In Iowa, the need for understanding changes in design flows due to climate change is especially important as increases in temperature and precipitation are expected to increase the magnitude of annual maximum discharges (references). Currently, the U.S. Geological Survey (USGS) provides design flows for ungauged locations in Iowa through a webtool called Streamstats (USGS, 2021). In Streamstats, the design flows are calculated from regional flood estimation equations established by a relationship between maximum discharge and drainage area through statistical regression (Eash et al., 2013), and the parameters of these models related to catchments' physical properties are selected based on minimization of residuals. However, only one-third of the developed regional equations have a climate-related parameter (i.e., rainfall), which is an important driver of hydrologic processes (Michalek et al., 2022). Therefore, the current methods cannot provide projections for design flows, highlighting limitations of available tools.

Due to these reasons and lack of available online resources, we have created the Iowa Flood Frequency and Projections Tool (IFFP; https://iowafloodfrequency.iihr.uiowa.edu/), which incorporates climate change projections in flood frequency estimates. We use the Iowa Flood Center (IFC)'s Hillslope-Link Model (HLM) to estimate discharge projections utilizing forcings derived from global climate models (GCMs) part of the Coupled Model Intercomparison Projected Phase 5 (CMIP5, Taylor et al., 2012) and Phase 6 (CMIP6; Eyring et al., 2016). This tool addresses two specific objectives. First, we provide flood frequency estimates based on default settings (*IFFP page*). Specifically, the approach provides estimates by fitting a subset of discharge projections to a nonstationary generalized extreme value (GEV) distribution. These settings cannot be edited by the user and reflect our best assessment of the projected changes in flood frequency for a given scenario and CMIP suite. The second objective is to provide another service, which allows users to explore how the selection of the CMIP suite, emission scenario, climate models, time period, and distribution type (*For Research page*). For this chapter, we briefly summarize the hydrologic dataset used for flood frequency analysis, highlight the features of the main and for research pages, and give a brief discussion on the tool.

This chapter is organized as follows. In Section 5.2 we describe the study area, simulation data, statistical methods, and the web framework. In Section 5.3 we provide a case study highlighting the ability of the IFFP tools. Section 5.4 provides some discussion and concluding remarks.

[5] Adapted from: Michalek, A., R. Amorim, G. Villarini, F. Quintero, and W.F. Krajewski, Iowa Flood Frequency and Projections Tool (IFFP): A tool to project flood frequency estimates across Iowa, 2023 (in preparation)

5.2 Materials and Methods

5.2.1 Study Area

IFFP focuses on providing information for the State of Iowa, located in the central United States. Information is available at every stream within the state that drains into the Mississippi and Missouri Rivers with a drainage area greater than 400 km^2 as shown in Fig. 5-1. Iowa is an area dominated by agriculture, with low relief and a high drainage density (Prior, 1991). It is characterized by seven landforms' features, with the Des Moines Lobe being the most hydrologically disconnected. The climate of the region is seasonally variable, with most of the rainfall occurring in the summer. The culmination of the landscape and climate characteristics has made Iowa prone to significant flooding (Flanagan et al., 2020; Mallakpour & Villarini, 2015; Smith et al., 2013).

5.2.2 Web Interface

The IFFP web tool is a simple interface providing four main pages. First, the web tool opens up to the *About* page (Fig. 5-2). This page is populated with information on the web tool discussing the functionality, related publications, and information on the authors. Additionally, there is a link to a *wiki* page that provides in-depth information on the hydrologic model, climate models utilized as a data source for this tool as well as equations for the distribution types incorporated in the tools. The second page is the *IFFP* page, which provides our best estimates of projected changes in flood frequency and is the main tool for users to interact with. The third page is titled *For Research* and provides a secondary tool for in-depth and customizable analyses. More information on the tools can be found below. The final page is a *Contact* page that allows users to send questions and information to the research team.

5.2.3 Data Sources

The data included in the IFFP tool is based on simulations of annual maximum discharges from 1981 (1950 for CMIP5) to the end of the 21st century at every stream segment in Iowa using the HLM developed by the IFC. Specifically, we utilized the TETIS-version of the HLM (see Quintero and Velasquez (2022) for more details). The theoretical background and conceptualization of the HLM is comprehensively described in Mantilla et al. (2022). The TETIS version of HLM is a fully distributed hydrologic model structure, which is based on the decomposition of the landscape into hillslopes and channels. Simulations are conducted under naturalized flow conditions where stream regulation (i.e., dams) is ignored due to the lack of available data on future operating procedures. The simulations are conducted offline as they require high-performance computing, and the final results are stored in a PostgreSQL database on our server. All calculations with the tool are conducted with these simulations.

For the primary function of the tool (*IFFP page*), annual flood peak projections are developed from forcings from 20 climate models within the CMIP6 suite. The CMIP6 suite provide climate projections for four emission or shared socio-economic pathways (SSPs). Each SSP represents a corresponding emission and land-use scenario. We incorporate SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 into our tool, which represent end-of-century radiative forcings of 2.6, 4.5, 7.0 and 8.5 W/m^2, respectively.

For the research page of the tool, we also include flood peak projections from 1950 to 2100 utilizing 19 CMIP5 GCMs. The CMIP5 suite contains two emission scenarios or representative concentration pathways, RCP45 and RCP85. RCP45 represents a scenario in which the peak emissions occur near 2040 and then decline towards the end of the century. RCP85 is a scenario in which emissions continue to rise through 2100. The values 45 and 85 represent the radiative forcing level in 2100 (W/m^2) indicating 4.5 and 8.5 W/m^2, respectively.

5.2.4 Statistical Methods

The primary function of this tool is to return flood frequency projections for given river segments based on flood peak projections. To achieve this, we utilize the distribution fitting of the time series of projected flood peaks based on stationary and nonstationary procedures to determine values at annual exceedance probabilities (AEPs) of 0.5, 0.2, 0.1, 0.04, 0.02, 0.01 and 0.004 as utilized by the USGS. For the main function (*IFFP page*), we determine AEPs by fitting a GEV distribution to the annual flood peak projections for a scenario of interest. We select the "best" GEV distribution model with respect to the Bayesian Information Criterion (BIC) among different model configurations: 1) no parameters varying on time (if no trend is detected); 2) location varying on time; 3) scale varying on time; and 4) location and scale varying on time. Trend in the projected flood peak is determined based on the Mann-Kendall trend test (Kendall, 1975; Mann, 1945) at the 5% significance level. We utilize the GEV distribution for the main page of our tool as it is rooted in extreme value theory and has been shown to be flexible for estimating AEPs of nonstationary flood peak time series (O'Brien & Burn, 2014; Slater et al., 2021; Villarini et al., 2009). For the research section of our tool, this same procedure can be applied utilizing stationary and nonstationary versions of the Gumbel and Lognormal distributions. The distribution fitting is conducted with the GAMLSS (Rigby & Stasinopoulos, 2005) and extRemes libraries in R. Additionally, on the research page we provide the ability to determine the percent change between a past and future period at a stream link. This function returns the percent change of the medians between each user-selected period.

5.2.5 Architecture and System Structure

5.2.5.1 Architecture

The IFFP tool consists of two main toolboxes or pages as previously discussed. Each follows a similar front-end/back-end system architecture as shown in Fig. 5-3. For the front-end, we utilize HTML and JavaScript for dynamical visualization of content such as graphs, tables, and maps. We utilize the Google JavaScript library (Charts, 2023) to provide line plots. We use Leaflet's JavaScript library (Leaflet-js, 2023) to create an interactive map for users. For the back-end processes, we utilize PHP to query our hydrologic dataset from the SQL databases (PostgreSQL, 2023) and to execute an R script with the data for distribution fitting (when applicable). The primary differences are related to the data flow, and the next two subsections describe the individual system structure for the *IFFP* and *For Research* pages.

5.2.5.2 System Structure IFFP Page

Figure 5-4 provides the schematic overview of the data flow for the *IFFP* page tool set. The main steps of the system are as follows:

1. **Step 1:** The user navigates to a river segment of interests and clicks on it to send a request to the back-end of the server.
2. **Step 2:** The request utilizes the ID for the river segment selected and queries the PostgreSQL database for all climate model flood projections from 1981 to 2099 (18 values per year, 2160 in total for a segment) for the CMIP6 suite. The default setting is for SSP1-2.6 scenario. The data are queried utilizing PHP, and a JSON with the results is returned to the client side.
3. **Step 3:** The queried flood projections for the river segment are averaged across climate models for each year. This step is completed within JavaScript.

4. **Step 4:** The average data time series is fit with a GEV distribution and a JSON file of AEPs over time is returned. This is completed by sending the time series to a PHP script, which executes an R script on the server side.
5. **Step 5:** The final step is to take the JSON file of AEPs and generate a line chart in JavaScript for the user to visualize.

The requests sent between the server and client side are executed in the background and cannot be intervened by the user. Furthermore, the user can select a different emission scenario (SSP) with a radio button on the chart created in step 5, which reruns through steps 1 to 5.

5.2.5.3 System Structure For Research Page

Figure 5-5 provides the schematic overview of the data flow for the *For Research* page tool set. This is similar to the *IFFP* page but allows the users more flexibility in the analyses. The main steps of the system are as follows:

1. **Step 1:** The user is allowed to select multiple options to inform their request consisting of: a) selection of stream segment of interest; b) either distribution fitting or ratio of historical and future flood peak medians; c) either CMIP5 or CMIP6 climate suite; d) suite specific climate models; e) the emission scenario of interest (RCP or SSP); and f) analysis specific parameters. For part f), if distribution fitting is selected, the user specifies only the start and end time along with the distribution of interest. If the ratio of annual maxima is selected, the user selects the start and end times for a historical and future period.
2. **Step 2:** The request utilizes the user specific information (i.e., ID) and queries the PostgreSQL database. The data are queried utilizing PHP and JSON, with the results that are returned to the client side.
3. **Step 3:** The queried flood projections for the river segment are averaged across climate models for each year and combined into a JSON array, with the individual climate model results based on the user selection. This step is completed within JavaScript.
4. **Step 4a:** If the user selects distribution fitting for the analysis, the data from step 3 call to the server side to run this analysis. In this step the distribution type is utilized as an input to the server commands as well (i.e., PHP call to R). The distribution fitting is done for each individual climate model as well as the average time series. All of the distribution fitting results return the AEP in JSON format, which are stored in a JavaScript array.
5. **Step 4b:** If the ratio of annual maxima is called, the JSON array of flood projections is utilized to run a simple function for all models and the GCMs' average, which are stored for step 5.
6. **Step 5:** If step 4a is utilized, a line chart of AEPs is returned, displaying the distribution fitting results for the average GCMs' flood projections. A drop down is made available to toggle between different GCM results. If step 4b is returned, an output box is populated with the calculated ratios by GCM.

5.3. Case Study

To demonstrate the capability of the IFFP tools, we highlight the outputs of the primary and secondary analyses for the City of Cedar Rapids. Additionally, we provide brief commentary on how a prospective users might utilize these results.

To begin, we examine a case study of the primary tool page (*IFFP*). Figure 5-6 shows the results when a user selects the river link (highlighted in yellow) for Cedar Rapids. A display is returned providing a line

chart of AEP projections for SSP1-2.6 scenario for the 1981-2095 period. The box displays a unique id (LinkID: 367814) for the river segment selected and the drainage area of the basin (6492 mi^2). The plot indicates that the annual maximum discharges have a trend, as the AEP values (y-axis) are changing over time (increasing). The user can then select a different emission scenario based on the available radio buttons. Figure 5-7 shows the results if a user selects SSP5-8.5. The new AEPs are different from those in Fig. 5-6. This example shows the AEP discharge values are projected to increase more throughout the 21st century for the higher-emission scenario. This allows users the flexibility to toggle between emission scenarios with minimum input and allows for easy comparisons.

If the user selects the "Build Report" button, a table of the AEP values appears along with a map of the selected point and the ability to download the simulation and AEP estimates (Fig. 5-8). For the 50% AEP, a stakeholder would be able to see that from 1981 to 2095 the discharge value are projected to increase from 41,506 ft^3/s to 47,396 ft^3/s under SSP5-8.5. This change could be incorporated into a design or plan to help account for future conditions. Figure 5-8 shows the top and bottom screenshot of the report that can be built from this page; it provides the AEP estimates in tabular format as well as the ability to download the AEP estimates and the simulation discharge data to csv.

We now shift our focus to the *For Research* page tools. The tools on this page are for stakeholders, engineers, and scientists who would want to examine different scenarios and time spans. We start with a demonstration of the distribution fitting analysis available. Figure 5-9 shows a user case where the stream link (yellow) near Cedar Rapids is selected for a distribution analysis. The selected climate model suite is CMIP5, with RCP4.5 and three models (i.e., access1-0, canesm2, cmcc-cm). The start and end years for the period of interest can be specified and for this example we have selected 1950 and 2020, respectively. Finally, the user selects a Lognormal distribution to be fit over time using imperial units. The output provides the AEP discharges over time as on the first page, except the user can toggle through each individual model; however, these results are based on a subset of different GCMs and on a different distribution compared to the *IFFP* page. In Fig. 5-9 the access1-0 climate model simulation results are displayed and no trend in the simulation is detected ($p>0.05$), and hence a stationary Lognormal distribution is fit. The "Results" boxes highlight the respective parameters selected, trend test results, BIC, and distribution parameters. In the final line of the "Results" box, the "-999" indicates parameters not used in the distribution fit. Here the log-mean (mu1) is 9.63 ft^3/s with a variance (sigma1) of 0.68. These statistics are displayed when selecting a different chart to display as well.

Next, we display the capability of the second tool in this page, which allows the users to see the ratio of future to present median annual maximum discharges to obtain a scaling factor for the analyses. Figure 5-10 shows an example where a user selects the same climate model scenario as in Fig. 5-4 and is interested in the ratio between the median annual maximum discharge computed during the 1950-1979 and 2070-2099 periods. The "Results" box provides the ratio (future over reference) of the medians of the periods for each model and their average. The results show that, for the selected models, the annual maximum medians are expected to increase by 9% on average, with values ranging from -19% to 49% in this case. This tool allows users to explore a simple metric to understand how flood peaks are expected to change and how they vary between simulations with difference climate models.

5.4 Discussions and Conclusions

We have presented the IFFP web-based tool, which provides flood frequency projections for Iowa and allows the exploration of future changes in flood peaks. We have provided simple use cases to demonstrate its capabilities, which can be efficiently utilized by stakeholders across the state. It fills a current software gap for stakeholders as current tools do not provide flood projections for end-users. The main function of the tool provides users with AEP discharge estimates over time to the end of the 21st century by fitting a GEV distribution with parameters that can vary on time. It provides users the ability to

download the simulation data, AEP data, as well as a pdf report for a specific river link anywhere in the state. Furthermore, it easily allows users to search a location with a simple geographical interface. The research page expands on the ability of the main tool by allowing users to explore how climate model selection, time period, emission scenario, and distribution type impact flood frequency projections. Finally, it provides users the ability to see the median change between two periods based on the selected scenario. To generalize the web-tool for the entire United States, further development of web-tools for flood frequency projections should be developed and integrated.

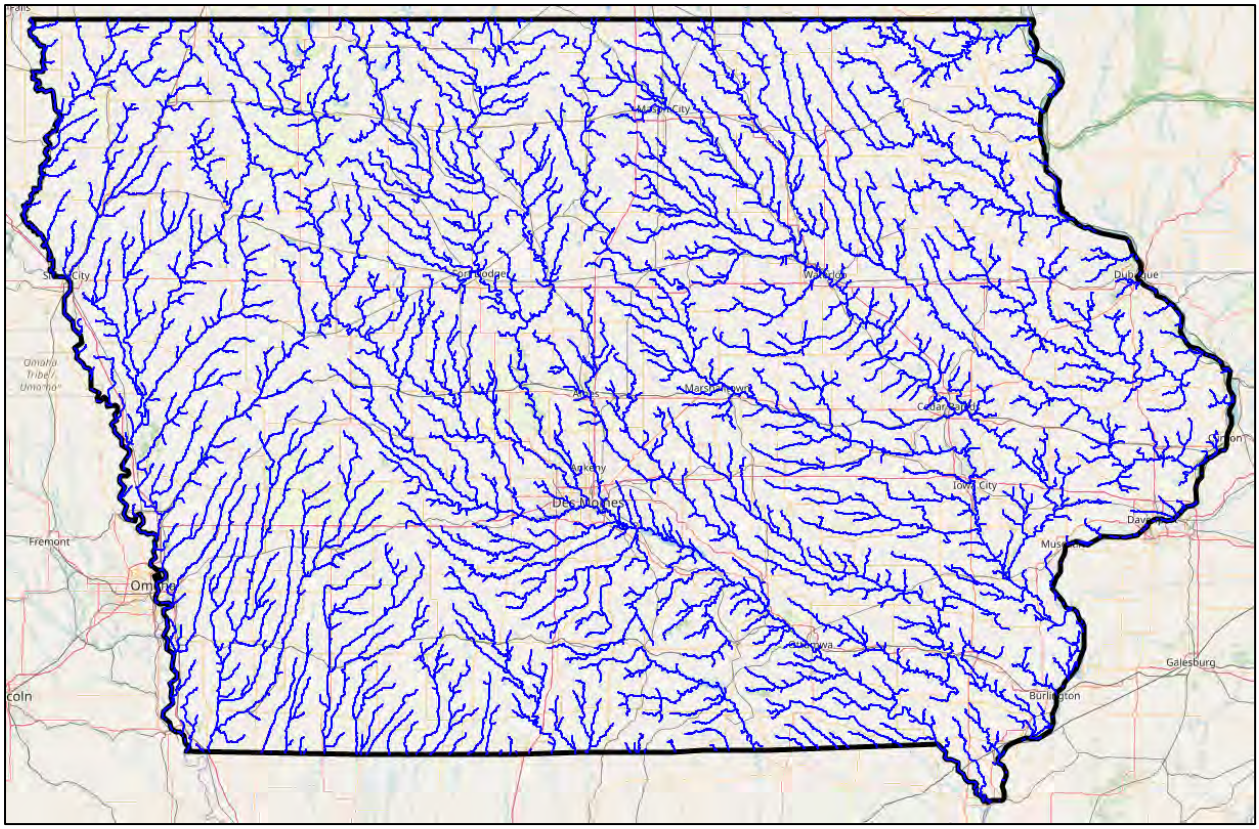

Figure 5-1. Stream segments in Iowa utilized in IFFP.

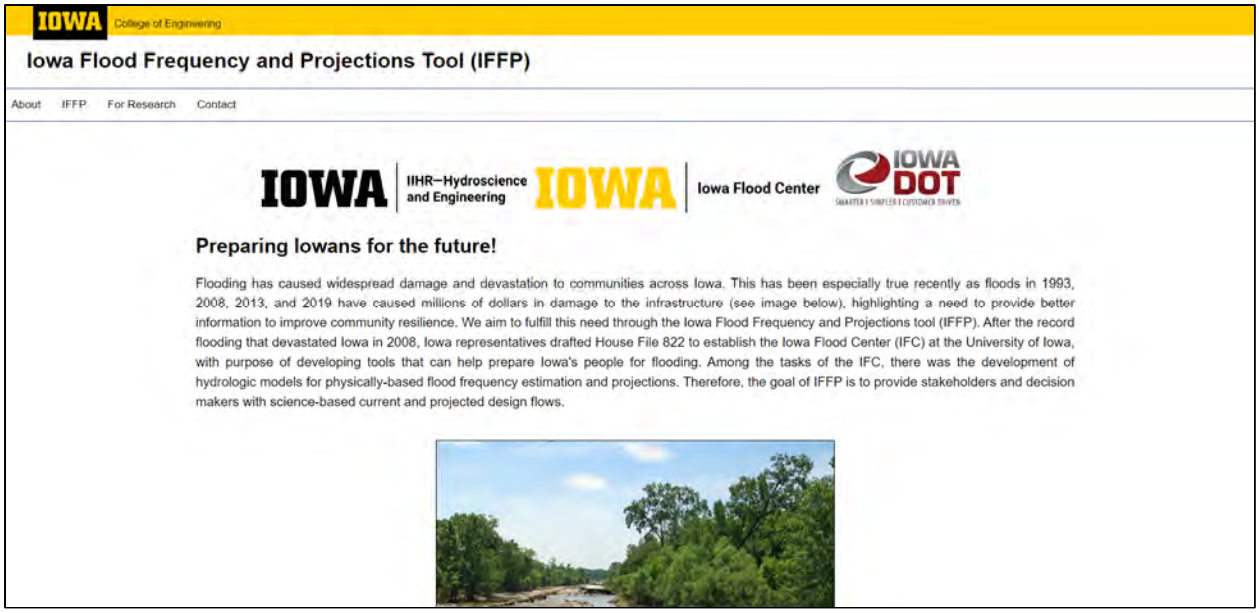

Figure 5-2. Home interface of the IFFP tool.

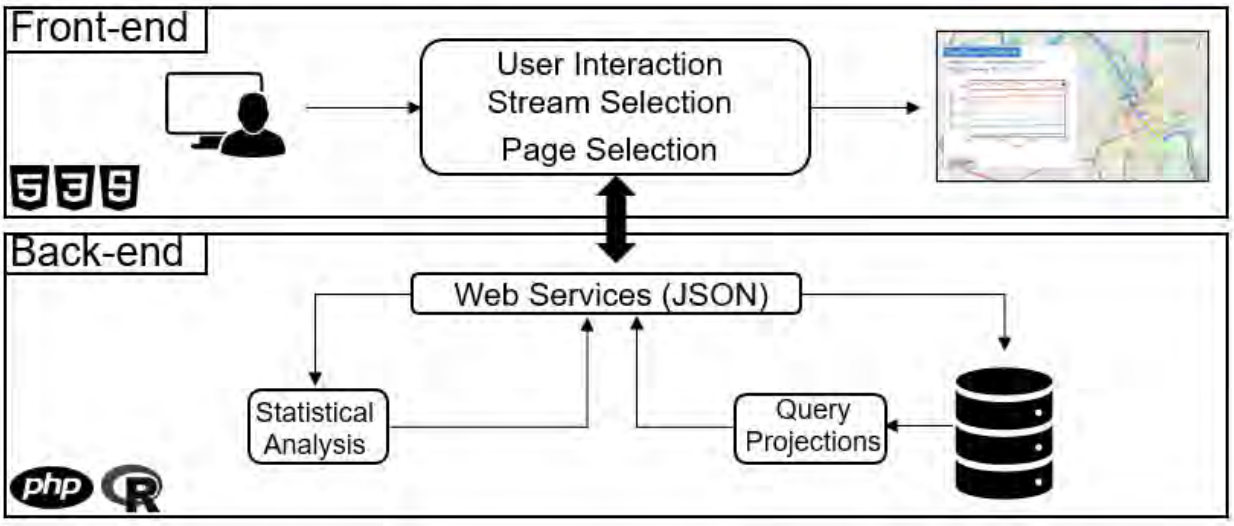

Figure 5-3. IFFP development architecture.

Figure 4. *IFFP page* data flow. The numbers represent the steps within the process described in Section 5.2.5.2.

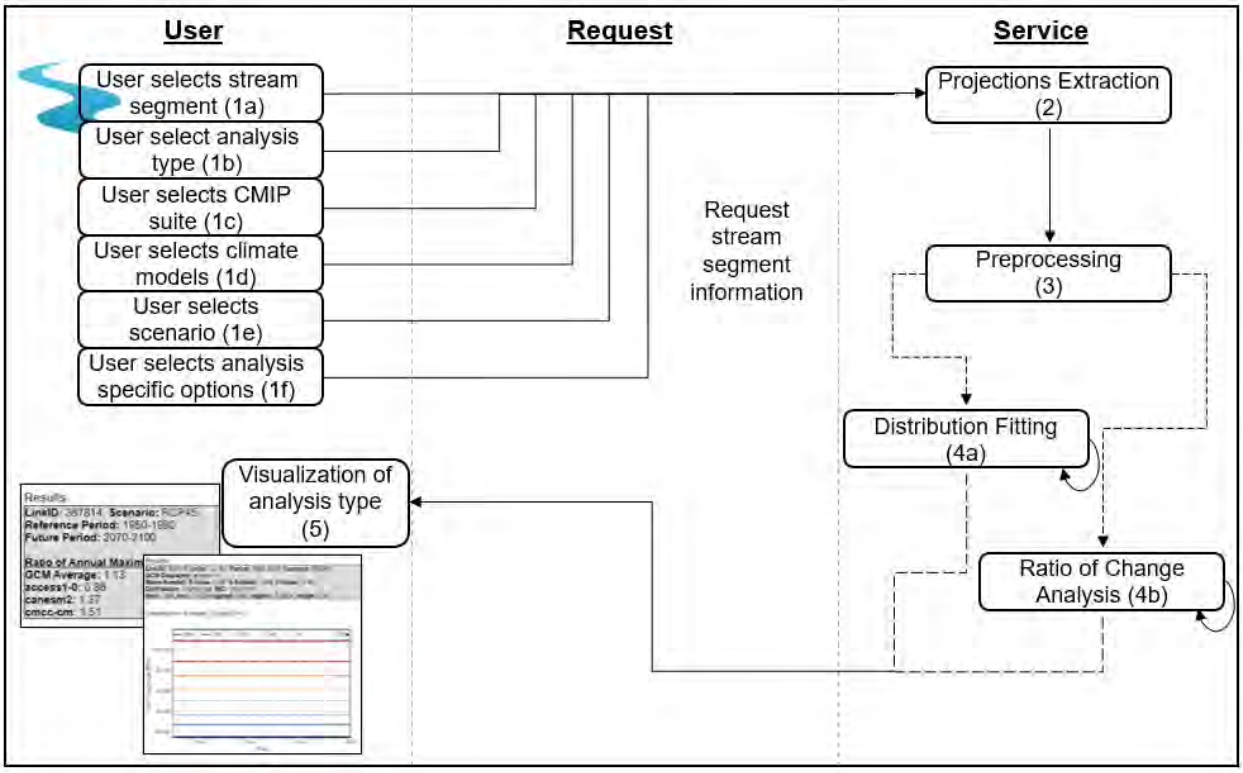

Figure 5-5. *For Research* page data flow. The numbers represent the steps within the process described in Section 5.2.5.3.

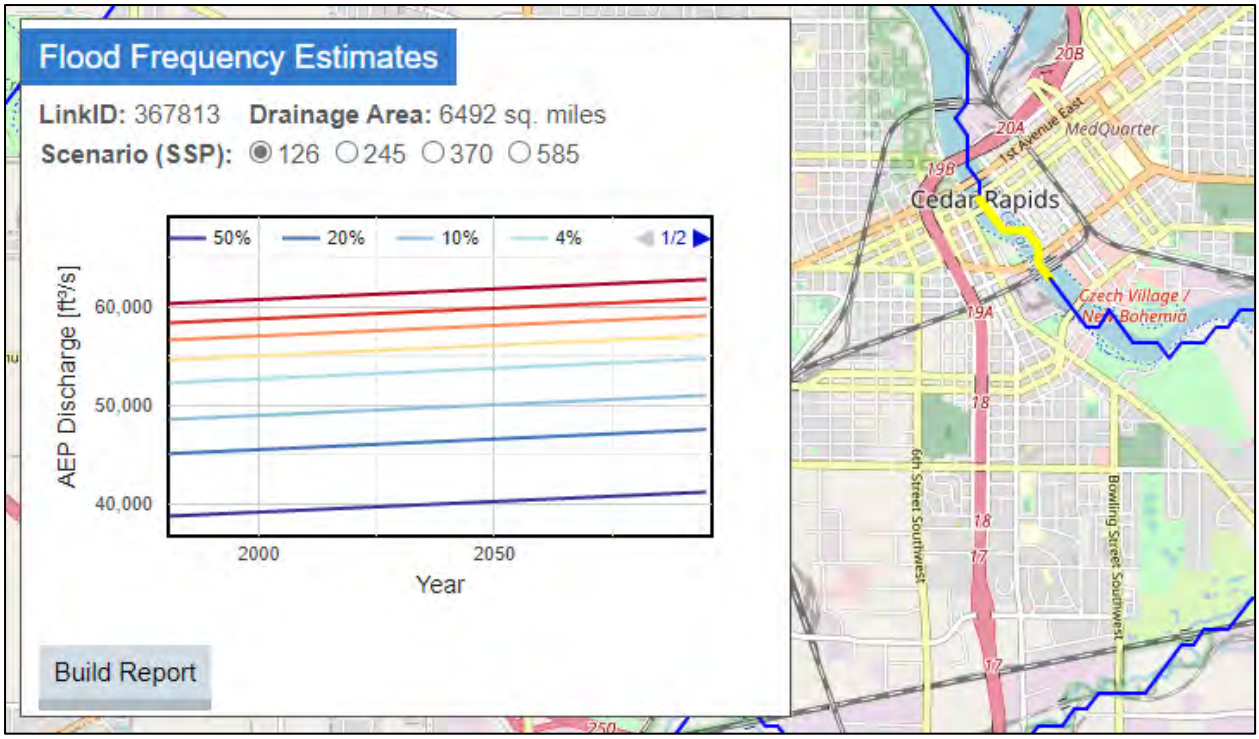

Figure 5-6. Screenshot showing test case when selecting a link (yellow) at Cedar Rapids, Iowa, and the returned AEP discharges under emission scenario SSP1-2.6.

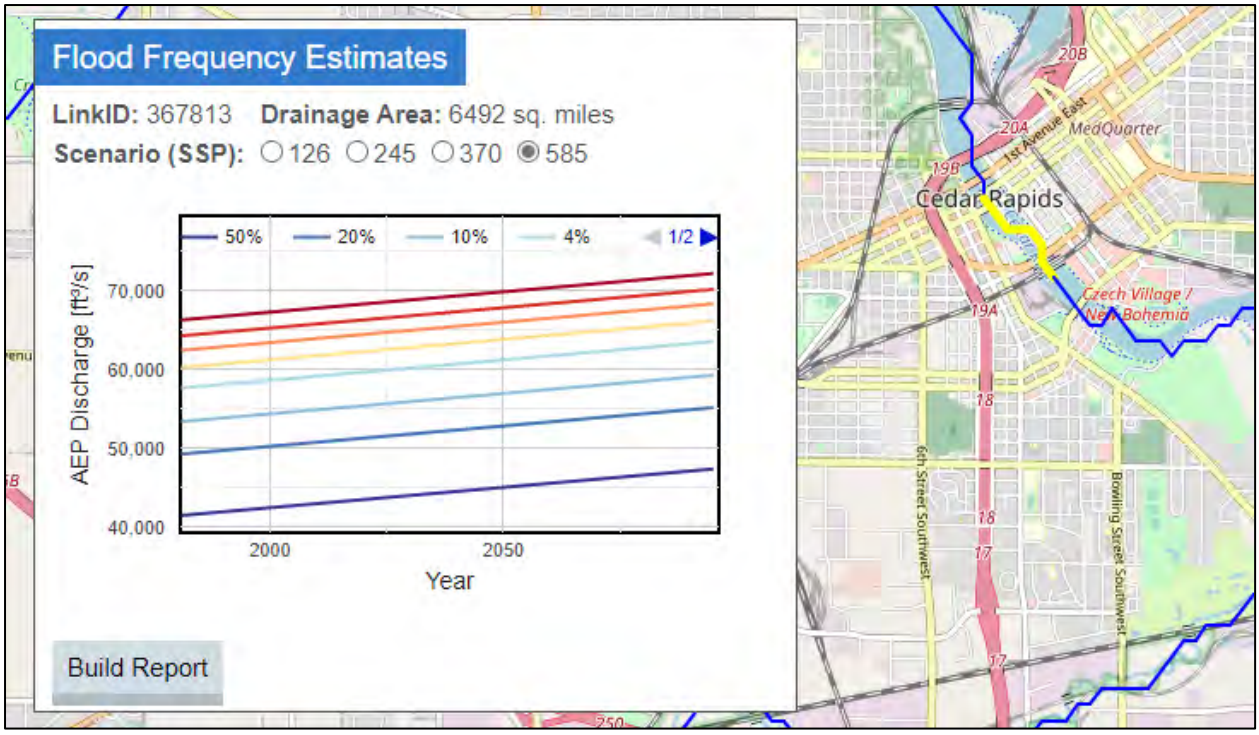

Figure 5-7. Screenshot showing test case when selecting a link (yellow) at Cedar Rapids, Iowa, and the returned AEP discharges under emission scenario SSP5-8.5.

IFFP Report (SSP585)

LinkID: 367813
Longitude/Latitude: -91.67, 41.97
Upstream Area: 6492 sq. miles
Time: Wed Feb 22 2023 12:21:14 GMT-0600 (Central Standard Time)

IFFP Estimates [ft³/s]:

Year	50%	20%	10%	4%	2%	1%	0.5%	0.2%
1981	41506	49293	53412	57672	60269	62452	64297	66310
1982	41558	49345	53464	57724	60320	62504	64349	66362
1983	41610	49397	53515	57775	60372	62556	64401	66413
1984	41661	49448	53567	57827	60424	62607	64452	66465
1985	41713	49500	53619	57879	60475	62659	64504	66517
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2090	47138	54925	59044	63304	65900	68084	69929	71942
2091	47190	54977	59095	63355	65952	68136	69981	71993
2092	47241	55028	59147	63407	66004	68187	70032	72045
2093	47293	55080	59199	63459	66055	68239	70084	72097
2094	47345	55132	59250	63510	66107	68291	70136	72148
2095	47396	55183	59302	63562	66159	68342	70187	72200

Print Report | Download Simulation Data | Download Selected AEP

Figure 5-8. The top and bottom of the report bult when selection the “Build Report” button in Fig. 5-2. It provides the AEP projections in tabular format along with a map of the area selected, and buttons to download the data.

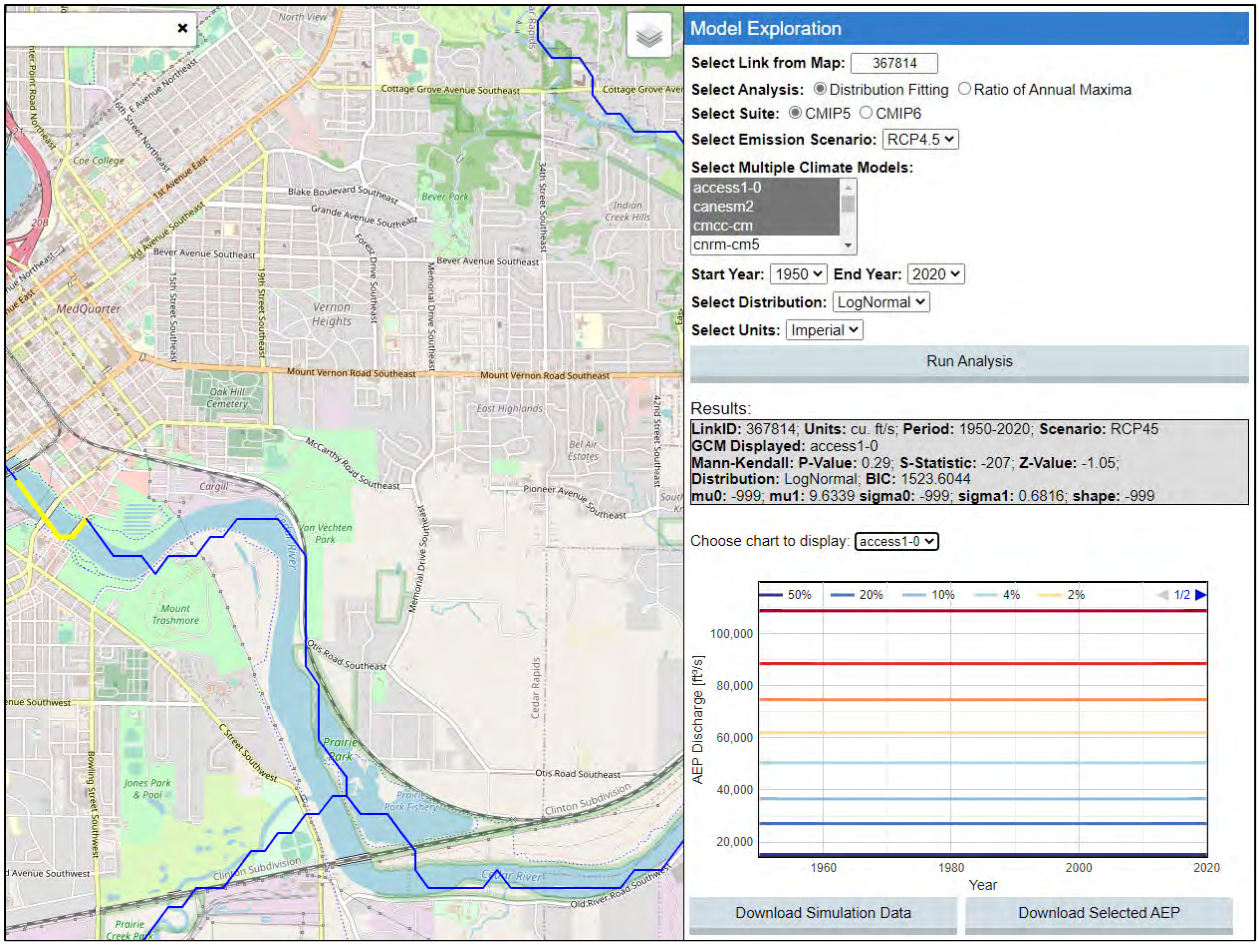

Figure 5-9. Screenshot of distribution fitting of the Lognormal distribution from three selected CMIP5 climate models from 1950 to 2020 under RCP4.5 at Cedar Rapids, Iowa (yellow highlight) The "Results" box shows the distribution parameters with the Mann-Kendall statistic and an AEP discharge plot over time for the average of the models.

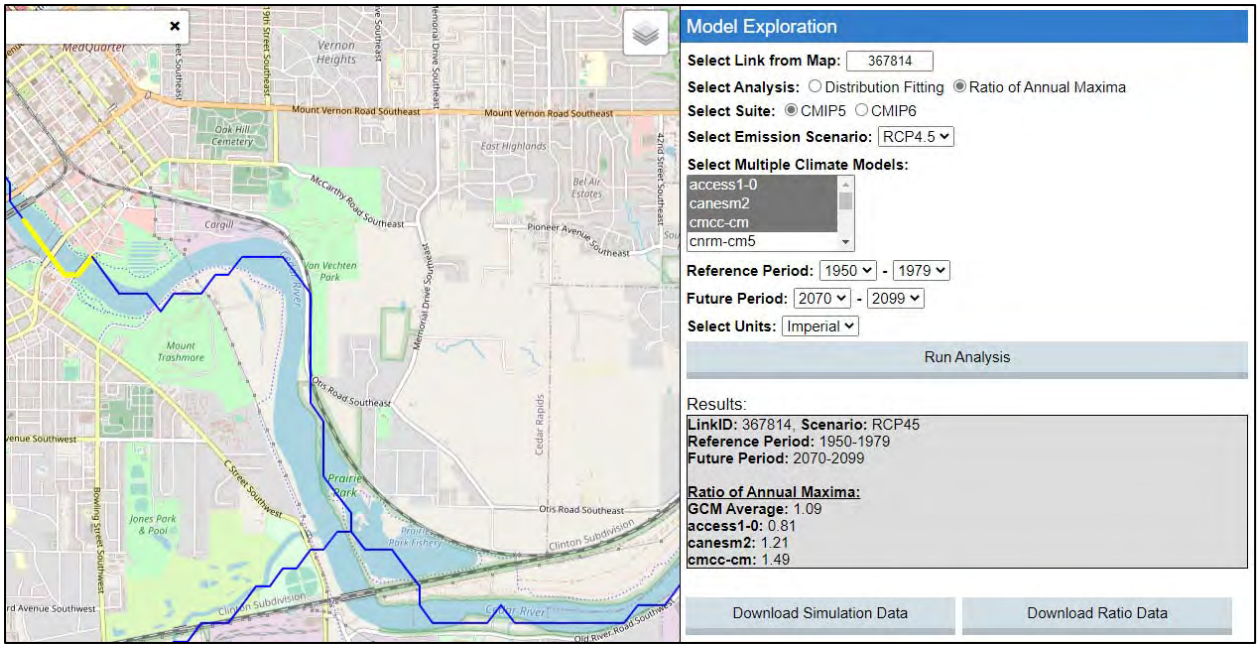

Figure 5-10. Screenshot of ratio of median annual maxima between periods of 1950-1979 and 2070-2099 for three selected CMIP5 climate models at Cedar Rapids, Iowa (yellow highlight) The "Results" box provides a summary of the user's selections as well as the ratio of median annual maxima between periods for the models.

6. Summary and Conclusions

We have focused on the examination of the projected changes in flooding across Iowa using the hydrologic model developed by the IFC using outputs from CMIP5 and CMIP6 global climate models and different scenarios. The main findings of our effort can be summarized as follows:

- The use of a physically-based model like HLM in flood frequency analysis provides a consistent method for modeling exceedance discharges necessary for engineering design as well as provides a framework to reassess current regional flood quantile estimation practices. This is a departure from the current standard approach for USGS regional regression equations for Iowa, which defines flood regions where no true physical or climate boundary exists.
- We investigated the response of annual maximum discharge over time under climate change scenarios (RCP4.5 and RCP8.5) at the community level across Iowa. A significant positive trend is detected for 70% of sites under RCP8.5, compared to only 2% of the sites under RCP4.5 for the selected GCMs, pointing to the increase in flood hazard under higher emission scenarios.
- We evaluated the potential for the CMIP6 HighResMIP suite to be utilized for hydrologic simulations, and found these models to be fit for purpose. We found that the median in flood peaks from most communities is projected to increase by ~20% on average across the state. The river communities across Iowa which should be focused on by stakeholders are those along the Des Moines and Cedar River, with Des Moines and Cedar Rapids that are two major cities in Iowa.
- Our study was one of the first examining the viability of CMIP6 HighResMIP forcings for hydrologic application. We recommend that hydrologic modeling with HighResMIP projections utilizes precipitation forcings instead of runoff outputs. Furthermore, bias-correction and downscaling of precipitation led to a good hydrologic performance in reproducing the distribution of annual maximum discharges across Iowa's river communities.
- Projected changes in the magnitude of annual maximum discharge based on HighResMIP were not detected for the immediate future across the state, but significant changes in interannual variability for the future were, with results differing between precipitation and runoff simulations. Investigation of the discrepancies between approaches should be the focus of future endeavors to quantify uncertainties in projections of flooding.
- We have developed the IFFP web-based tool, which provides flood frequency projections for Iowa and allows the exploration of future changes in flood peaks. It fills a current software gap for stakeholders as current tools do not provide flood projections for end-users. IFFP provides users the ability to download the simulation data, AEP data, as well as a pdf report for a specific river link anywhere in the state. Furthermore, it easily allows users to search a location with a simple geographical interface. The research page expands on the ability of the main tool by allowing users to explore how climate model selection, time period, emission scenario, and distribution type impact flood frequency projections. Finally, it provides users the ability to see the median change between two periods based on the selected scenario.

The primary limitations of our work are related to the hydrologic and climate models. First, we assumed that the streams in Iowa were under naturalized flow conditions, but a few streams are regulated. This should be investigated in studies when trying to apply future streamflow simulations to engineering design as the regulatory actions are important to consider. Next, we selected the climate models based on how well they reproduced historical trends, magnitude, and distribution. However, uncertainty in

CMIP5/CMIP6 projections due to original model spatial coarseness and bias-correction and temporal disaggregation processes exist and can contribute to the underestimation of extreme events. We explored the impact of model selection by rerunning our analyses using all available climates models, finding differences in the results over our primary analysis. We suggest that when providing flood projections for communities that a sensitivity analysis of model selection should be performed.

References

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, Journal of Climate, 25(3), 939-957, 2012.

An, B., Yu, Y., Bao, Q., He, B., Li, J., Luan, Y., et al., CAS FGOALS-f3-H dataset for the High-Resolution Model Intercomparison Project (HighResMIP) Tier 2. Advances in Atmospheric Sciences, https://doi.org/10.1007/s00376-022-2030-5, 2022.

Archfield, S. A., R. M. Hirsch, A. Viglione and G. Bloschl, Fragmented patterns of flood change across the United States, Geophysical Research Letters, 43(19), 10232-10239, 2016.

Ayalew, T. B., and W. F. Krajewski, Effect of river network geometry on flood frequency: A tale of two watersheds in Iowa, Journal of Hydrologic Engineering, 22 (8), 2017.

Ayalew, T. B., W. F. Krajewski, R. Mantilla, and S. J. Small, Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge, Advances in Water Resources, 64, 9-20. 2014.

Baeck, M. L., J. A. Smith, G. Villarini, D. B. Wright, and W. Krajewski, Extreme flood response: The June 2008 flooding in Iowa, Journal of Hydrometeorology, 14 (6), 1810-1825, 2013.

Balaguru, K., Leung, L. R., Van Roekel, L. P., Golaz, J. C., Ullrich, P. A., Caldwell, P. M., et al., Characterizing tropical cyclones in the Energy Exascale Earth System Model Version 1, Journal of Advances in Modeling Earth Systems, 12(8), 2020.

Ball, J. E., Modelling accuracy for urban design flood estimation, Urban Water Journal, 19 (1), 87-96. 2022.

Bates, P. D., N. Quinn, C. Sampson, A. Smith, O. Wing, J. Sosa, et al., Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates, Water Resources Research, 57(2), 2021.

Bellucci, A., Athanasiadis, P. J., Scoccimarro, E., Ruggieri, P., Gualdi, S., Fedele, G., et al., Air-Sea interaction over the Gulf Stream in an ensemble of HighResMIP present climate simulations, Climate Dynamics, 56(7-8), 2093-2111, 2021.

Benjamini, Y., & Yekutieli, D., The control of the false discovery rate in multiple testing under dependency, The annals of statistics, 29(4), 1165-1188, 2001.

Berne, A., G. Delrieu, J.-D. Creutin, and C. Obled, Temporal and spatial resolution of rainfall measurements required for urban hydrology, Journal of Hydrology, 299 (3-4), 166-179. 2004.

Blazkova, S., and K. Beven, Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty)." Water Resources Research, 38 (8), 14-11-14-14. 2002.

Blazkova, S., and K. Beven, A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka Catchment, Czech Republic, Water Resources Research, 45(12), 2009.

Bobée, B., G. Cavadias, F. Ashkar, J. Bernier, and P. Rasmussen, Towards a systematic approach to comparing distributions used in flood frequency analysis, Journal of Hydrology, 142 (1), 121-136. 1993.

Bonferroni, C., Teoria statistica delle classi e calcolo delle probabilita, Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8, 3-62, 1936.

Boughton, W., and O. Droop, Continuous simulation for design flood estimation—a review, Environmental Modeling and Software, 18 (4), 309-318, 2003.

Bradley, A. A., and K. W. Potter, Flood frequency analysis of simulated flows, Water Resources Research, 28 (9), 2375-2385, 1992.

Byun, K., C. M. Chiu and A. F. Hamlet, Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US, Science of the Total Environment, 650(Pt 1), 1261-1277, 2019.

Cameron, D. S., K. J. Beven, J. Tawn, S. Blazkova, and P. Naden, Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty), Journal of Hydrology, 219 (3), 169-187, 1999.

Cannon, A. J., Sobie, S. R., & Murdock, T. Q., Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? Journal of Climate, 28(17), 6938-6959, 2015.

Centre for Research on the Epidemiology of Disaster, The human cost of natural disasters: A global perspective, 2015.

Chang, P., Zhang, S., Danabasoglu, G., Yeager, S. G., Fu, H., Wang, H., et al., An unprecedented set of high-resolution earth system simulations for understanding multiscale interactions in climate variability and change, Journal of Advances in Modeling Earth Systems, 12(12), 2020.

Charts, G., Google Visualization API Reference. Retrieved from https://developers.google.com/chart/, 2023.

Cohn, T. A., J. F. England, C. E. Berenbrock, R. R. Mason, J. R. Stedinger, and J. R. Lamontagne, A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series, Water Resources Research, 49 (8), 5047-5058, 2013.

Cunha, L. K., P. V. Mandapaka, W. F. Krajewski, R. Mantilla, and A. A. Bradley, Impact of radar-rainfall error structure on estimated flood magnitude across scales: An investigation based on a parsimonious distributed hydrological model, Water Resources Research, 48 (10), 2012.

Cunnane, C., A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, Journal of Hydrology, 18 (3), 257-271. 1973.

Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris, A knowledge-based approach to the statistical mapping of climate, Climate Research, 22 (2), 99-113. 2002.

Dawdy, D. R., V. W. Griffis, and V. K. Gupta, Regional flood-frequency analysis: How we got here and where we are going, Journal of Hydrologic Engineering, 17 (9), 953-959. 2012.

Demaria, E. M. C., R. N. Palmer and J. K. Roundy, Regional climate change projections of streamflow characteristics in the Northeast and Midwest U.S, Journal of Hydrology: Regional Studies, 5309-323, 2016a.

Demaria, E. M. C., J. K. Roundy, S. Wi and R. N. Palmer, The effects of climate change on seasonal snowpack and the hydrology of the Northeastern and Upper Midwest United States, Journal of Climate, 29(18), 6527-6541, 2016b.

Dornes, P. F., B. A. Tolson, B. Davison, A. Pietroniro, J. W. Pomeroy, and P. Marsh, Regionalisation of land surface hydrological model parameters in subarctic and arctic environments, Physics and Chemistry of the Earth, Parts A/B/C, 33 (17-18), 1081-1089. 2008.

Dottori, F., W. Szewczyk, J.-C. Ciscar,F. Zhao, L. Alfieri, Y. Hirabayashi, et al., Increased human and economic losses from river flooding with anthropogenic warming, Nature Climate Change, 8(9), 781-786, 2018.

Douglas, E. M., R. M. Vogel and C. N. Kroll, Trends in floods and low flows in the United States: impact of spatial correlation, Journal of Hydrology, 240(1), 90-105, 2000.

Eagleson, P. S., Dynamics of flood frequency, Water Resources Research, 8 (4), 878-898, 1972.

Eash, D. A., K. K. Barnes, and A. G. Veilleux,Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010, U.S. Geological Survey Scientific Investigations Report, 63 p. with appendix, 2013.

ElSaadani, M., and W. F. Krajewski, A time-based framework for evaluating hydrologic routing methodologies using wavelet transform, Journal of Water Resource and Protection, 09 (07), 723-744, 2017.

England Jr, J. F., T. A. Cohn, B. A. Faber, J. R. Stedinger, W. O. Thomas Jr, A. G. Veilleux, J. E. Kiang, and J. R. R. Mason, Guidelines for determining flood flow frequency—Bulletin 17C, Reston, VA, 168, 2019.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937-1958, 2016.

Fang, X., J. W. Pomeroy, C. R. Ellis, M. K. MacDonald, C. M. DeBeer, and T. Brown, Multi-variable evaluation of hydrological model predictions for a headwater basin in the Canadian Rocky Mountains, Hydrology and Earth System Science, 17 (4), 1635-1659, 2013.

Fang, X., J. W. Pomeroy, C. J. Westbrook, X. Guo, A. G. Minke, and T. Brown, Prediction of snowmelt derived streamflow in a wetland dominated prairie basin, Hydrology and Earth System Science, 14 (6), 991-1006, 2010.

Feng, Z., Houze, R. A., Leung, L. R., Song, F., Hardin, J. C., Wang, J., et al., Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the Rocky Mountains, Journal of Climate, 32(21), 7303-7328, 2019.

Ficchì, A., C. Perrin, and V. Andréassian, Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events, Journal of Hydrology, 538, 454-470, 2016.

Flanagan, P. X., Mahmood, R., Umphlett, N. A., Haacker, E., Ray, C., Sorensen, W., et al., A hydrometeorological assessment of the historic 2019 flood of Nebraska, Iowa, and South Dakota, Bulletin of the American Meteorological Society, 101(6), E817-E829, 2020.

Gabellani, S., G. Boni, L. Ferraris, J. von Hardenberg, and A. Provenzale, Propagation of uncertainty from rainfall to runoff: A case study with a stochastic rainfall generator, Advances in Water Resources, 30 (10), 2061-2071, 2007

Gao J.; Minobe, S., Roberts, M., Haarsma, R., Putrasahan, D., Roberts, C., Scoccimarro, E.; Terray, L., Vannière, B., Vidale, P. L., Influence of model resolution on bomb cyclones revealed by HighResMIP-PRIMAVERA simulations. Environmental Research Letters, 2020.

Ghimire, G. R., W. F. Krajewski, T. B. Ayalew, and R. Goska, Hydrologic investigations of radar-rainfall error propagation to rainfall-runoff model hydrographs, Advances in Water Resources, 161, 2022.

Giuntoli, I., G. Villarini, C. Prudhomme and D. M. Hannah, Uncertainties in projected runoff over the conterminous United States, Climatic Change, 150(3-4), 149-162, 2018.

Grist, J. P., Josey, S. A., Sinha, B., Catto, J. L., Roberts, M. J., & Coward, A. C., Future evolution of an eddy rich ocean associated with enhanced East Atlantic storminess in a coupled model projection, Geophysical Research Letters, 48(7), 2021.

Gupta, V. K., B. M. Troutman, and D. R. Dawdy, Towards a nonlinear geophysical theory of floods in river networks: An overview of 20 years of progress, In Nonlinear Dynamics in Geosciences, Anastasios, T. A., Elsner, J.B., Eds.; Springer, 121–151, 2007.

Gutjahr, O., & Heinemann, G., Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM, Theoretical and Applied Climatology, 114(3-4), 511-529, 2013.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., et al., Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), Geoscientific Model Development, 12(7), 3241-3281, 2019.

Haarsma, R., Acosta, M., Bakhshi, R., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., et al., HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR – description, model computational performance and basic validation. Geoscientific Model Development, 13(8), 3507-3527, 2020.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al., High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geoscientific Model Development, 9(11), 4185-4208, 2016.

Hariadi, M. H., Schrier, G., Steeneveld, G. J., Sopaheluwakan, A., Tank, A. K., Roberts, M. J., et al., Evaluation of onset, cessation and seasonal precipitation of the Southeast Asia rainy season in CMIP5

regional climate models and HighResMIP global climate models, International Journal of Climatology, 42(5), 3007-3024, 2021.

Hay, L. E., S. L. Markstrom and C. Ward-Garrison, Watershed-Scale Response to Climate Change through the Twenty-First Century for Selected Basins across the United States, Earth Interactions, 15(17), 1-37, 2011.

Hicks, M. J., and M. L. Burton, Preliminary flood damage estimates for Iowa: Great flood of 2008. In: Muncie, IN: Ball State University, 2008.

Hodgkins, G. A. and R. W. Dudley, Changes in the timing of winter–spring streamflows in eastern North America, 1913–2002, Geophysical Research Letters, 33(6), 2006.

Hrachowitz, M., H. H. G. Savenije, G. Blöschl, J. J. McDonnell, M. Sivapalan, J. W. Pomeroy, B. Arheimer, T. Blume, M. P. Clark, U. Ehret, F. Fenicia, J. E. Freer, A. Gelfan, H. V. Gupta, D. A. Hughes, R. W. Hut, A. Montanari, S. Pande, D. Tetzlaff, P. A. Troch, S. Uhlenbrook, T. Wagener, H. C. Winsemius, R. A. Woods, E. Zehe, and C. Cudennec, A decade of predictions in ungauged basins (PUB)—A review, Hydrological Sciences Journal, 58 (6), 1198-1255, 2013.

Huang, H., Patricola, C. M., & Collins, W. D., The influence of ocean coupling on simulated and projected tropical cyclone precipitation in the HighResMIP–PRIMAVERA simulations, Geophysical Research Letters, 48(20), 2021.

IPCC, Climate Change 2021: The Physical Science Basis, 2021.

IPCC,. Climate Change 2022: Impacts, Adaptation and Vulnerability, 2022.

Judt, F., Klocke, D., Rios-Berrios, R., Vanniere, B., Ziemen, F., Auger, L., et al., Tropical cyclones in global storm-resolving models, Journal of the Meteorological Society of Japan. Ser. II, 99(3), 579-602, 2021.

Kandel, D. D., A. W. Western, and R. B. Grayson, Scaling from process timescales to daily time steps: A distribution function approach, Water Resources Research, 41 (2), 2005.

Kendall, M., Rank correlation methods., Charles Griffin, London, 1975.

Klemes, V.,Some problems in pure and applied stochastic hydrology, In Proceedings of the symposium on statistical hydrology, 2-15, 1974.

Krajewski, W. F., D. Ceynar, I. Demir, R. Goska, A. Kruger, C. Langel, R. Mantilla, J. Niemeier, F. Quintero, B.-C. Seo, S. J. Small, L. J. Weber, and N. C. Young, Real-time flood forecasting and information system for the state of Iowa, Bulletin of the American Meteorological Society, 98 (3), 539-554, 2017.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S. C. Jain, Monte Carlo study of rainfall sampling effect on a distributed catchment model, Water Resources Research, 27 (1), 119-128, 1991.

Kreussler, P., Caron, L. P., Wild, S., Loosveldt Tomas, S., Chauvin, F., Moine, M. P., et al., Tropical cyclone integrated kinetic energy in an ensemble of HighResMIP simulations, Geophysical Research Letters, 48(5), 2021.

Leaflet-js, A JavaScript library for interactive maps. Retrieved from https://leafletjs.com/, 2023.

Leander, R., & Buishand, T. A., Resampling of regional climate model output for the simulation of extreme river flows. Journal of Hydrology, 332(3-4), 487-496, 2007.

Lehner, F., A. W. Wood, J. A. Vano, D. M. Lawrence, M. P., Clark, and J. S. Mankin, The potential to reduce uncertainty in regional runoff projections from climate models. Nature Climate Change, 9(12), 926-933, 2019.

Lettenmaier, D. P., E. F. Wood and J. R. Wallis, Hydro-climatological trends in the continental United States, 1948-88, Journal of Climate, 7(4), 586-607, 1994.

Li, J., Bao, Q., Liu, Y., Wang, L., Yang, J., Wu, G., et al., Effect of horizontal resolution on the simulation of tropical cyclones in the Chinese Academy of Sciences FGOALS-f3 climate system model, Geoscientific Model Development, https://doi.org/10.5194/gmd-2021-19, 2021.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research: Atmospheres, 99(D7), 14415-14428, 1994.

Liang, X., E. F. Wood, and D. P., Lettenmaier, Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, Global and Planetary Change, 13(1), 195-206, 1996.

Lin, Y., and K. E. Mitchell, The NCEP Stage II/IV hourly precipitation analyses: Development and applications, In Proceedings of the 19th Conference Hydrology, American Meteorological Society, San Diego, CA, USA, 2005.

Lins, H. F. and J. R. Slack, Streamflow trends in the United States, Geophysical Research Letters, 26(2), 227-230, 1999.

Lins, H. F. and J. R. Slack, Seasonal and Regional Characteristics of U.S. Streamflow Trends in the United States from 1940 to 1999, Physical Geography, 26(6), 489-501, 2005.

Lohmann, K., Putrasahan, D. A., Storch, J. S., Gutjahr, O., Jungclaus, J. H., & Haak, H., Response of northern North Atlantic and Atlantic Meridional overturning circulation to reduced and enhanced wind stress forcing, Journal of Geophysical Research: Oceans, 126(11), 2021.

Lopez-Parages, J., & Terray, L., Tropical North Atlantic response to ENSO: Sensitivity to model spatial resolution, Journal of Climate, 35(1), 3-16, 2022.

Mallakpour, I. and G. Villarini, The changing nature of flooding across the central United States, Nature Climate Change, 5(3), 250-254, 2015.

Mann, H. B., Nonparametric tests against trend, Econometrica, 13 (3), 245-259, 1945.

Mann, H. B., & Whitney, D. R., On a test of whether one of two random variables is stochastically larger than the other, The annals of mathematical statistics, 50-60, 1947.

Mantilla, R., Gupta, V. K., & J. Mesa, O., Role of coupled flow dynamics and real network structures on Hortonian scaling of peak flows, Journal of Hydrology, 322(1-4), 155-167, 2006.

Mantilla, R., W. F. Krajewski, N. Velasquez, S. J. Small, T. B. Ayalew, F. Quintero, N. Jadidoleslam, and M. Fonley, The hydrological hillslope-link model for space-time prediction of streamflow: Insights and applications at the Iowa Flood Center, In Extreme Weather Forecasting, Astitha, M., Nikolopoulos, E. I., Eds.; Elsevier, 2022.

Martinec, J., & Rango, A., Areal distribution of snow water equivalent evaluated by snow cover monitoring, Water Resources Research, 17(5), 1480-1488, 1981.

Massey, F. J., The Kolmogorov-Smirnov test for goodness of fit, Journal of the American Statistical Association, 46 (253), 68-78, 1951.

McCabe, G. J. and D. M. Wolock, A step increase in streamflow in the conterminous United States, Geophysical Research Letters, 29(24), 38-31-38-34, 2002.

McSweeney, C. F., R.G. Jones, R. W. Lee, and D. P. Rowell, Selecting CMIP5 GCMs for downscaling over multiple regions, Climate Dynamics, 44(11-12), 3237-3260, 2014.

Merz, B., G. Blöschl, S. Vorogushyn, F. Dottori, J. C. J. H. Aerts, P. Bates, M. Bertola, M. Kemter, H. Kreibich, U. Lall and E. Macdonald, Causes, impacts and patterns of disastrous river floods, Nature Reviews Earth & Environment, 2(9), 592-609, 2021.

Michalek, A., F. Quintero, G. Villarini, and W. F. Krajewski, Advantages of Physically-based Flood Frequency Analysis with Long Term Simulations for Iowa, ASCE Journal of Hydrologic Engineering. 27(12), 1-11, 2022.

Michalek, A., Quintero, F., Villarini, G., & Krajewski, W. F., Advantages of physically based flood frequency analysis with long-term simulations for Iowa. Journal of Hydrologic Engineering, 27(12), 2022.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J.,. Stationarity Is Dead: Whither Water Management? Science, 319(5863), 573-574, 2008.

Mitchell, K. E.,The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple gcip products and partners in a continental distributed hydrological modeling system, Journal of Geophysical Research, 109 (D7), 2004.

Moreno-Chamarro, E., Caron, L.-P., Loosveldt Tomas, S., Vegas-Regidor, J., Gutjahr, O., Moine, M.-P., et al., Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models, Geoscientific Model Development, 15(1), 269-289, 2022.

Moreno-Chamarro, E., Caron, L. P., Ortega, P., Loosveldt Tomas, S., & Roberts, M. J., Can we trust CMIP5/6 future projections of European winter precipitation? Environmental Research Letters, 16(5), 2021.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, et al., The next generation of scenarios for climate change research and assessment, Nature, 463(7282), 747-756, 2010.

Murphy, A. H., and R. L. Winkler, Diagnostic verification of probability forecasts, International Journal of Forecasting, 7 (4), 435-455, 1992.

National Research Council, Disaster Resilience: A National Imperative, Washington, DC: The National Academies Press, 2012.

Naz, B. S., S.-C. Kao, M. Ashfaq, D. Rastogi, R. Mei and L. C. Bowling, Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations, Global and Planetary Change, 143100-117, 2016.

Neri, A., G. Villarini, L. J. Slater and F. Napolitano, On the statistical attribution of the frequency of flood events across the U.S. Midwest, Advances in Water Resources, 127225-236, 2019.

Nikolopoulos, E. I., E. N. Anagnostou, M. Borga, E. R. Vivoni, and A. Papadopoulos, Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability, Journal of Hydrology, 402, 165–178, 2011.

O'Brien, N. L., and D. H. Burn, A nonstationary index-flood technique for estimating extreme quantiles for annual maximum streamflow, Journal of Hydrology, 519, 2040-2048, 2014.

O'Connell, P., On the relation of the freshwater floods of rivers to the areas and physical features of their basins and on a method of classifying rivers and streams with reference to the magnitude of their floods, In Minutes of the Proceedings of the Institution of Civil Engineers, Thomas Telford-ICE Virtual Library, 204-217, 1868.

Pathiraja, S., S. Westra, and A. Sharma, Why continuous simulation? The role of antecedent moisture in design flood estimation, Water Resources Research, 48 (6), 2012.

Perez, G., R. Mantilla, and W. F. Krajewski, The influence of spatial variability of width functions on regional peak flow regressions, Water Resources Research, 54 (10), 7651-7669, 2018.

Perez, G., R. Mantilla, W. F. Krajewski, and F. Quintero, Examining observed rainfall, soil moisture, and river network variabilities on peak flow scaling of rainfall-runoff events with implications on regionalization of peak flow quantiles, Water Resources Research, 55 (12), 10707-10726, 2019.

Perica, S., D. Martin, S. Pavlovic, I. Roy, M. S. Laurent, C. Trypaluk, D. Unruh, M. Yekta, and G. Bonnin, Noaa Atlas 14 volume 8, precipitation-frequency atlas of the United States, Midwestern States, NOAA, National Weather Service, Silver Spring, MD, 2013.

Piani, C., Haerter, J. O., & Coppola, E., Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99(1-2), 187-192, 2009.

PostgreSQL, Learn PostgreSQL from Scratch. Retrieved from https://www.postgresqltutorial.com/, 2023.

Praskievicz, S., and P. Bartlein, Hydrologic modeling using elevationally adjusted NARR and NARCCAP regional climate-model simulations: Tucannon River, Washington, Journal of Hydrology, 517, 803-814, 2014.

Priestley, M. D. K., & Catto, J. L., Improved representation of extratropical cyclone structure in HighResMIP models, Geophysical Research Letters, 49(5), 2022.

Prior, J. C., Landforms of Iowa, University of Iowa Press, 1991.

Quintero, F., and N. Velasquez, Implementation of TETIS hydrologic model into the Hillslope Link Model framework, Water, 14(17), 2610, 2022.

Quintero, F., W. F. Krajewski, R. Mantilla, S. Small, and B.-C. Seo, A spatial–dynamical framework for evaluation of satellite rainfall products for flood prediction, Journal of Hydrometeorology, 17 (8), 2137-2154, 2016.

Quintero, F., W. F. Krajewski, B.-C. Seo, and R. Mantilla, Improvement and evaluation of the Iowa Flood center Hillslope Link Model (HLM) by calibration-free approach, Journal of Hydrology, 584, 2020.

Quintero, F., R. Mantilla, C. Anderson, D. Claman and W. Krajewski, Assessment of Changes in Flood Frequency Due to the Effects of Climate Change: Implications for Engineering Design, Hydrology, 5(1), 2018.

Rango, A., The Snowmelt Runoff Model (SRM). Computer models of watershed hydrology., 477-520, 1995.

Rankinen, K., T. Karvonen, and D. Butterfield, A simple model for preidcting soil temperature in snow-covered and seasonally frozen soil: model description and testing, Hydrology and Earth System Sciences, 8(4), 706-716, 2004.

Rigby, R. A., and Stasinopoulos, D. M., Generalized additive models for location, scale and shape, Journal of the Royal Statistical Society: Series C (Applied Statistics), 54(3), 507-554, 2005.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., et al., Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments, Geoscientific Model Development, 12(12), 4999-5028, 2019.

Santander Meteorology Group, downscaleR: Climate data manipulation and statistical downscaling (Version R package version 0.6-0), 2015.

Santhi, C., N. Kannan, J. Arnold, and M. Di Luzio, Spatial calibration and temporal validation of flow for regional scale hydrologic modeling, Journal of the American Water Resources Association, 44 (4), 829-846, 2008.

Schaake, J., A. Henkel, and S. Cong, Application of PRISM climatologies for hydrologic modeling and forcasting in the western US, In Proc. 18th Conf. on Hydrology, Seattle, WA, American Meteorological Society, 2004.

Schmidli, J., Frei, C., & Vidale, P. L., Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, International Journal of Climatology, 26(5), 679-689, 2006.

Scoccimarro, E., Peano, D., Gualdi, S., Bellucci, A., Lovato, T., Fogli, P. G., & Navarra, A., Extreme events representation in CMCC-CM2 standard and high-resolution general circulation models, Geoscientific Model Development, 15(4), 1841-1854, 2022.

Seager, R., M. Ting, C. Li, N. Naik, B. Cook, J. Nakamura, and H. Liu, Projections of declining surface-water availability for the southwestern United States, Nature Climate Change, 3(5), 482-486, 2012

Shuster, W. D., J. Bonta, H. Thurston, E. Warnemuende, and D. R. Smith, Impacts of impervious surface on watershed hydrology: A review, Urban Water Journal, 2 (4), 263-275, 2005.

Sivapalan, M., Prediction in ungauged basins: A grand challenge for theoretical hydrology, Hydrological Processes, 17 (15), 3163-3170, 2003.

Slater, L., G. Villarini, S. Archfield, D. Faulkner, R. Lamb, A. Khouakhi and J. Yin, Global Changes in 20-Year, 50-Year, and 100-Year River Floods, Geophysical Research Letters, 48(6), 2021.

Slater, L. J. and G. Villarini, Recent trends in US flood risk, Geophysical Research Letters, 43(24), 12428-12436, 2016.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., et al. (2021), Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management, Hydrology and Earth System Sciences, 25(7), 3897-3935, 2021.

Small, D., S. Islam and R. M. Vogel, Trends in precipitation and streamflow in the eastern U.S.: Paradox or perception?, Geophysical Research Letters, 33(3), 2006.

Smith, A. B., and R. W. Katz, US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases, Natural hazards, 67(2), 387-410, 2013.

Smith, J. A., Baeck, M. L., Villarini, G., Wright, D. B., & Krajewski, W., Extreme flood response: The June 2008 flooding in Iowa. Journal of Hydrometeorology, 14(6), 1810-1825, 2013.

Squintu, A. A., van der Schrier, G., van den Besselaar, E., van der Linden, E., Putrasahan, D., Roberts, C., et al., Evaluation of trends in extreme temperatures simulated by HighResMIP models across Europe. Climate Dynamics, 56(7-8), 2389-2412, 2021.

Tan, X., and T. Y. Gan, Nonstationary Analysis of Annual Maximum Streamflow of Canada, Journal of Climate, 28(5), 1788-1805, 2015.

Tate, E., A. Strong, T. Kraus, and H. Xiong, Flood recovery and property acquisition in Cedar Rapids, Iowa, Natural hazards, 80(3), 2055-2079, 2016.

Taylor, K. E., R. J. Stouffer and G. A. Meehl, An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93(4), 485-498, 2012.

Tohver, I. M., A. F. Hamlet, and S.-Y Lee, Impacts of 21st-Century Climate Change on Hydrologic Extremes in the Pacific Northwest Region of North America, Journal of the American Water Resources Association, 50(6), 1461-1476, 2014.

Tsartsali, E. E., Haarsma, R. J., Athanasiadis, P. J., Bellucci, A., de Vries, H., Drijfhout, S., et al., Impact of resolution on the atmosphere–ocean coupling along the Gulf Stream in global high resolution models, Climate Dynamics, 58(11-12), 3317-3333, 2022.

USGS, WDN (USGS Water Data for the Nation). Washington, DC: US Dept. of the Interior. Retrieved from: https://waterdata.usgs.gov/nwis/rt, 2022.

USGS, Streamstats: Streamflow statistics and spatial analysis tools for water-resources applications, Accessed October 20, 2021. https://streamstats.usgs.gov/ss/, 2021.

Vannière B. , M. Roberts, P.L. Vidale, K. Hodges, M-E Demory, L-P Caron, E. Scoccimarro , L. Terray, and R. Senan, The moisture budget of tropical cyclones in HighResMIP models: Large-scale environmental balance and sensitivity to horizontal resolution, Journal of Climate, 2020.

Ventura, V., Paciorek, C. J., & Risbey, J. S., Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data, Journal of Climate, 17(22), 4343-4356, 2004.

Vidale, P. L., Hodges, K., Vannière, B., Davini, P., Roberts, M. J., Strommen, K., et al., Impact of stochastic physics and model resolution on the simulation of tropical cyclones in climate gcms, Journal of Climate, 34(11), 4315-4341, 2021.

Villarini, G., and C. Wasko, Humans, climate and streamflow, Nature Climate Change, 11(9), 725-726, 2021.

Villarini, G., F. Serinaldi, J. A. Smith and W. F. Krajewski, On the stationarity of annual flood peaks in the continental United States during the 20th century, Water Resources Research, 45(8), 2009.

Villarini, G., J. A. Smith, M. L. Baeck and W. F. Krajewski, Examining Flood Frequency Distributions in the Midwest U.S.1, JAWRA Journal of the American Water Resources Association, 47(3), 447-463, 2011a.

Villarini, G., J. A. Smith, M. L. Baeck, R. Vitolo, D. B. Stephenson, and W. F. Krajewski, On the frequency of heavy rainfall for the Midwest of the United States, Journal of Hydrology, 400 (1-2), 103-120, 2011b.

Villarini, G., J. A. Smith and G. A. Vecchi, Changing Frequency of Heavy Rainfall over the Central United States, Journal of Climate, 26(1), 351-357, 2013.

Villarini, G. and L. Slater. Climatology of flooding in the United States, Oxford research encyclopedia of natural hazard science. 2017.

Villarini, G. and W. Zhang, Projected changes in flooding: a continental US perspective, Annals of the New York Academy of Sciences, 1472(1), 95-103, 2020.

Villarini, G., W. Zhang, F. Quintero, W. F. Krajewski and G. A. Vecchi, Attribution of the impacts of the 2008 flooding in Cedar Rapids (Iowa) to anthropogenic forcing, Environmental Research Letters, 15(11), 2020.

Villarini, G. and C. Wasko, Humans, climate and streamflow, Nature Climate Change, 11(9), 725-726, 2021.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al., Evaluation of CMIP6 DECK experiments With CNRM-CM6-1, Journal of Advances in Modeling Earth Systems, 11(7), 2177-2213, 2019.

Wasko, C., R. Nathan, L. Stein and D. O'Shea, Evidence of shorter more extreme rainfalls and increased flood variability under climate change, Journal of Hydrology, 6032021.

Wasko, C., S. Westra, R. Nathan, H. G. Orr, G. Villarini, R. Villalobos Herrera, and H. J. Fowler, Incorporating climate change in flood estimation guidance, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379(2195), 20190548, 2021.

Wehner, M., Lee, J., Risser, M., Ullrich, P., Gleckler, P., & Collins, W. D., Evaluation of extreme sub-daily precipitation in high-resolution global climate model simulations, Philos Trans A Math Phys Eng Sci, 379(2195), 20190545, 2021.

Wright, D. B., G. Yu, and J. F. England, Six decades of rainfall and flood frequency analysis using stochastic storm transposition: Review, progress, and prospects, Journal of Hydrology, 585, 2020.

Wu, T., Yu, R., Lu, Y., Jie, W., Fang, Y., Zhang, J., et al., BCC-CSM2-HR: a high-resolution version of the Beijing Climate Center Climate System Model, Geoscientific Model Development, 14(5), 2977-3006, 2021.

Yamada, Y., Kodama, C., Satoh, M., Sugi, M., Roberts, M. J., Mizuta, R., et al., Evaluation of the contribution of tropical cyclone seeds to changes in tropical cyclone frequency due to global warming in high-resolution multi-model ensemble simulations, Progress in Earth and Planetary Science, 8(1). https://doi.org/10.1186/s40645-020-00397-1, 2021.

Zhang, W., Villarini, G., Scoccimarro, E., Roberts, M., Vidale, P. L., Vanniere, B., et al., Tropical cyclone precipitation in the HighResMIP atmosphere-only experiments of the PRIMAVERA Project. Climate Dynamics, 57(1-2), 253-273, 2021.

Zhao, M., Simulations of Atmospheric Rivers, Their Variability, and Response to Global Warming Using GFDL's New High-Resolution General Circulation Model, Journal of Climate, 33(23), 10287-10303, 2020.

Zhao, M., A study of AR-, TS-, and MCS-associated precipitation and extreme precipitation in present and warmer climates, Journal of Climate, 35(2), 479-497, 2022.

Zhu, Z., D. B. Wright, and G. Yu., The impact of rainfall space-time structure in flood frequency analysis, Water Resources Research, 54 (11), 8983-8998, 2018.

Zoccatelli, D., M. Borga, G.B. Chirico, and E.I. Nikolopoulos, The relative role of hillslope and river network routing in the hydrologic response to spatially variable rainfall fields, Journal of Hydrology, 531, 349–359, 2015.

Appendix

Table S4-1. USGS gauges utilized for model validation in this study. The table below provides the gauge ID, drainage area (sq. miles), and description as provided on the USGS web page for each location.

ID Number	Description	Area
05387490	Dry Run Creek near Decorah (05387490)	21
05387500	Upper Iowa River at Decorah (DEHI4-05387500)	491
05388250	Upper Iowa River near Dorchester (DCHI4-05388250)	769
05388310	Waterloo Creek near Dorchester (05388310)	43
05412500	Turkey River at Garber (GRBI4-05412500)	1557
05414400	Middle Fk Little Maquoketa R near Rickardsville (05414400)	31
05418500	Maquoketa River near Maquoketa (MAQI4-05418500)	1550
05420850	Little Wapsipinicon River near Oran (05420850)	95
05421000	Wapsipinicon River at Independence (IDPI4-05421000)	1052
05421890	Silver Creek at Welton (05421890)	10
05422000	Wapsipinicon River near De Witt (DEWI4-05422000)	2287
05448400	West Main Drainage Ditch 1 & 2 near Britt (05448400)	21
05449500	Iowa River near Rowan (ROWI4-05449500)	426
05451500	Iowa River at Marshalltown (MIWI4-05451500)	1534
05451700	Timber Creek near Marshalltown (05451700)	119
05451900	Richland Creek near Haven (05451900)	57
05451955	Stein Creek near Clutier (05451955)	24
05452000	Salt Creek near Elberon (05452000)	201
05452200	Walnut Creek near Hartwick (05452200)	71
05453000	Big Bear Creek at Ladora (05453000)	187
05453100	Iowa River at Marengo (MROI4-05453100)	2795
05453200	Price Creek at Amana (05453200)	29
05454000	Rapid Creek near Iowa City (05454000)	26
05454300	Clear Creek near Coralville (05454300)	98
05454500	Iowa River at Iowa City (IOWI4-05454500)	3267
05455100	Old Mans Creek near Iowa City (05455100)	201
05455230	Deep River at Deep River (05455230)	30
05455500	English River at Kalona (KALI4-05455500)	576
05455700	Iowa River Near Lone Tree (Tri-County Bridge) (LNTI4-05455700)	4291
05457440	Deer Creek near Carpenter (05457440)	96
05457700	Cedar River at Charles City (CCYI4-05457700)	1077
05458000	Little Cedar River near Ionia (05458000)	298
05458500	Cedar River at Janesville (JANI4-05458500)	1676
05458900	West Fork Cedar River at Finchford (FNHI4-05458900)	852
05459490	Spring Creek near Mason City (05459490)	31
05459500	Winnebago River at Mason City (MCWI4-05459500)	484
05460100	Willow Creek near Mason City (05460100)	80
05462000	Shell Rock River at Shell Rock (SHRI4-05462000)	1711
05463000	Beaver Creek at New Hartford (NHRI4-05463000)	349

05463500	Black Hawk Creek at Hudson (05463500)	300
05464000	Cedar River at Waterloo (ALOI4-05464000)	5133
05464500	Cedar River at Cedar Rapids (CIDI4-05464500)	6492
05465000	Cedar River near Conesville (CNEI4-05465000)	7753
05465500	Iowa River at Wapello (WAPI4-05465500)	12483
05469860	Mud Lake Drainage Ditch 71 at Jewell (05469860)	66
05469990	Keigley Branch near Story City (05469990)	32
05470000	South Skunk River near Ames (AMEI4-05470000)	314
05470500	Squaw Creek at Ames (AMWI4-05470500)	213
05471000	South Skunk River below Squaw Creek near Ames (AESI4-05471000)	563
05471050	South Skunk River at Colfax (CFXI4-05471050)	813
05471200	Indian Creek near Mingo (05471200)	288
05471500	South Skunk River near Oskaloosa (OOAI4-05471500)	1657
05472390	Middle Creek near Lacey (05472390)	23
05472500	North Skunk River near Sigourney (SIGI4-05472500)	735
05473400	Cedar Creek near Oakland Mills (05473400)	533
05474000	Skunk River at Augusta (AGSI4-05474000)	4333
05476750	Des Moines River at Humboldt (HBTI4-05476750)	2286
05479000	East Fork Des Moines River at Dakota City (DAKI4-05479000)	1304
05480500	Des Moines River at Fort Dodge (FODI4-05480500)	4244
05480930	White Fox Creek at Clarion (05480930)	6
05481000	Boone River near Webster City (WBCI4-05481000)	840
05481300	Des Moines River near Stratford (STRI4-05481300)	5500
05481510	Bluff Creek at Pilot Mound (05481510)	20
05481650	Des Moines River near Saylorville (05481650)	5892
05481950	Beaver Creek near Grimes (GRMI4-05481950)	371
05482300	North Raccoon River near Sac City (05482300)	726
05482500	North Raccoon River near Jefferson (EFWI4-05482500)	1658
05483318	Brushy Fork Creek near Templeton (05483318)	45
05483349	Middle Raccoon River Tributary at Carroll (05483349)	6.58
05483450	Middle Raccoon River near Bayard (05483450)	367
05483600	Middle Raccoon River at Panora (05483600)	411
05484000	South Raccoon River at Redfield (REDI4-05484000)	972
05484500	Raccoon River at Van Meter (VNMI4-05484500)	3436
05484800	Walnut Creek at Des Moines (05484800)	78
05484900	Raccoon River at Fleur Drive at Des Moines (DEMI4-05484900)	3612
05485500	Des Moines River blw Raccoon R. at Des Moines (DESI4-05485500)	9926
05485640	Fourmile Creek at Des Moines (05485640)	91
05486000	North River near Norwalk (NRWI4-05486000)	350
05486490	Middle River near Indianola (IDNI4-05486490)	499
05487470	South River near Ackworth (AKWI4-05487470)	461
05487980	White Breast Creek near Dallas (05487980)	332
05488200	English Creek near Knoxville (05488200)	91
05488500	Des Moines River near Tracy (TRCI4-05488500)	12526
05489000	Cedar Creek near Bussey (BSSI4-05489000)	374
05489490	Bear Creek at Ottumwa (05489490)	23

05489500	Des Moines River at Ottumwa (OTMI4-05489500)	13421
05490500	Des Moines River at Keosauqua (KEQI4-05490500)	14083
06482610	Big Sioux River near Hawarden (06483950)	482
06483500	Rock River near Rock Valley (RKVI4-06483500)	1592
06599950	Perry Creek near Hinton (PCHI4-06599950)	33
06600100	Floyd River at Alton (ALTI4-06600100)	268
06600500	Floyd River at James (JMEI4-06600500)	889
06602020	West Fork Ditch at Hornick (HOKI4-06602020)	524
06602190	Elliott Creek at Lawton (06602190)	35
06602400	Monona-Harrison Ditch near Turin (06602400)	764
06605000	Ocheyedan River near Spencer (SPOI4-06605000)	424
06605750	Willow Creek near Cornell (06605750)	81
06605850	Little Sioux River at Linn Grove (LNNI4-06605850)	1541
06606600	Little Sioux River at Correctionville (CRRI4-06606600)	2479
06607200	Maple River at Mapleton (MAPI4-06607200)	671
06607500	Little Sioux River near Turin (TURI4-06607500)	764
06608500	Soldier River at Pisgah (PSGI4-06608500)	402
06609500	Boyer River at Logan (LGNI4-06609500)	859
06609560	Willow Creek near Soldier (06609560)	29
06807410	West Nishnabotna River at Hancock (HNKI4-06807410)	611
06808500	West Nishnabotna River at Randolph (RDPI4-06808500)	1324
06809210	East Nishnabotna River near Atlantic (ATCI4-06809210)	440
06809500	East Nishnabotna River at Red Oak (RDOI4-06809500)	897
06810000	Nishnabotna River above Hamburg (HMBI4-06810000)	2822
06811760	Tarkio River near Elliott (06811760)	10
06811800	East Tarkio Creek near Stanton (06811800)	5
06816290	West Nodaway River at Massena (06816290)	23
06817000	Nodaway River at Clarinda (ICLI4-06817000)	763
06819185	East Fork 102 River at Bedford (BDFI4-06819185)	86
06898000	Thompson River at Davis City (DVSI4-06898000)	698
06903400	Chariton River near Chariton (CHTI4-06903400)	185
06903700	South Fork Chariton River near Promise City (PRMI4-06903700)	169
06903900	Chariton River near Rathbun (06903900)	542
06904010	Chariton River near Moulton (MOLI4-06904010)	741

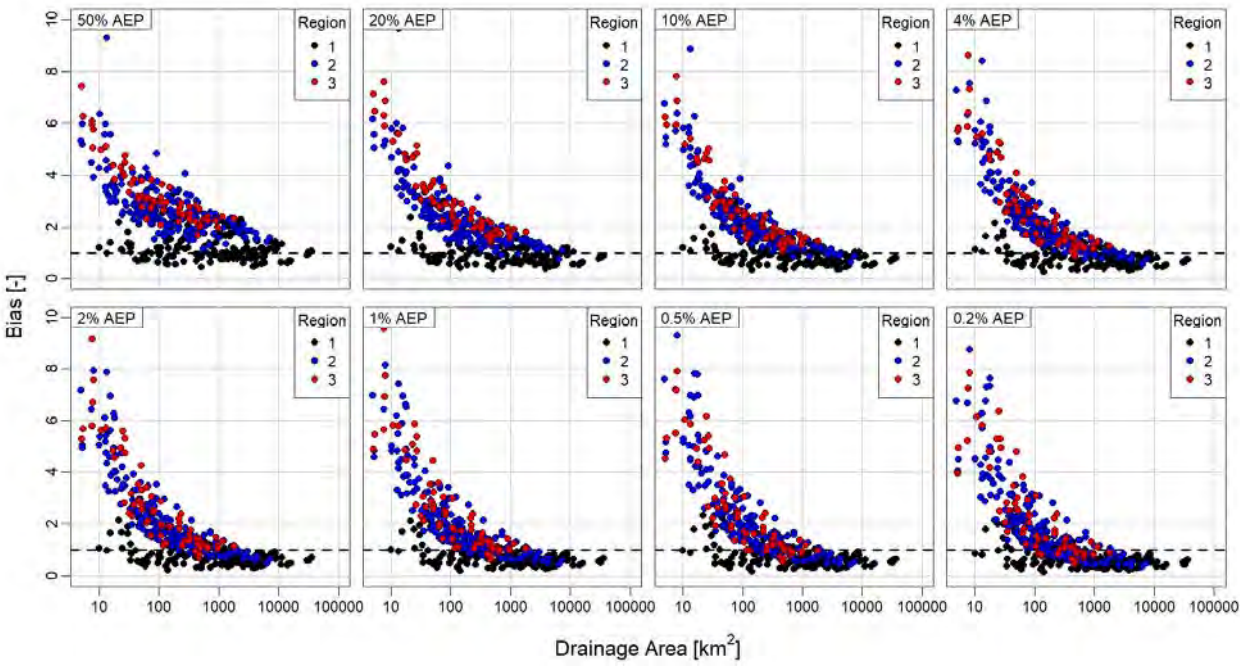

Fig. S2-1. Relation between scale of the watershed and bias for annual exceedance discharge at probabilities of 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2%. The dashed line represents a value of 1 in which the Streamstats derived discharge is equal to the HLM derived discharge. Bias values above (below) 1 indicate Streamstats overestimates (underestimate) AEP discharge compared HLM.

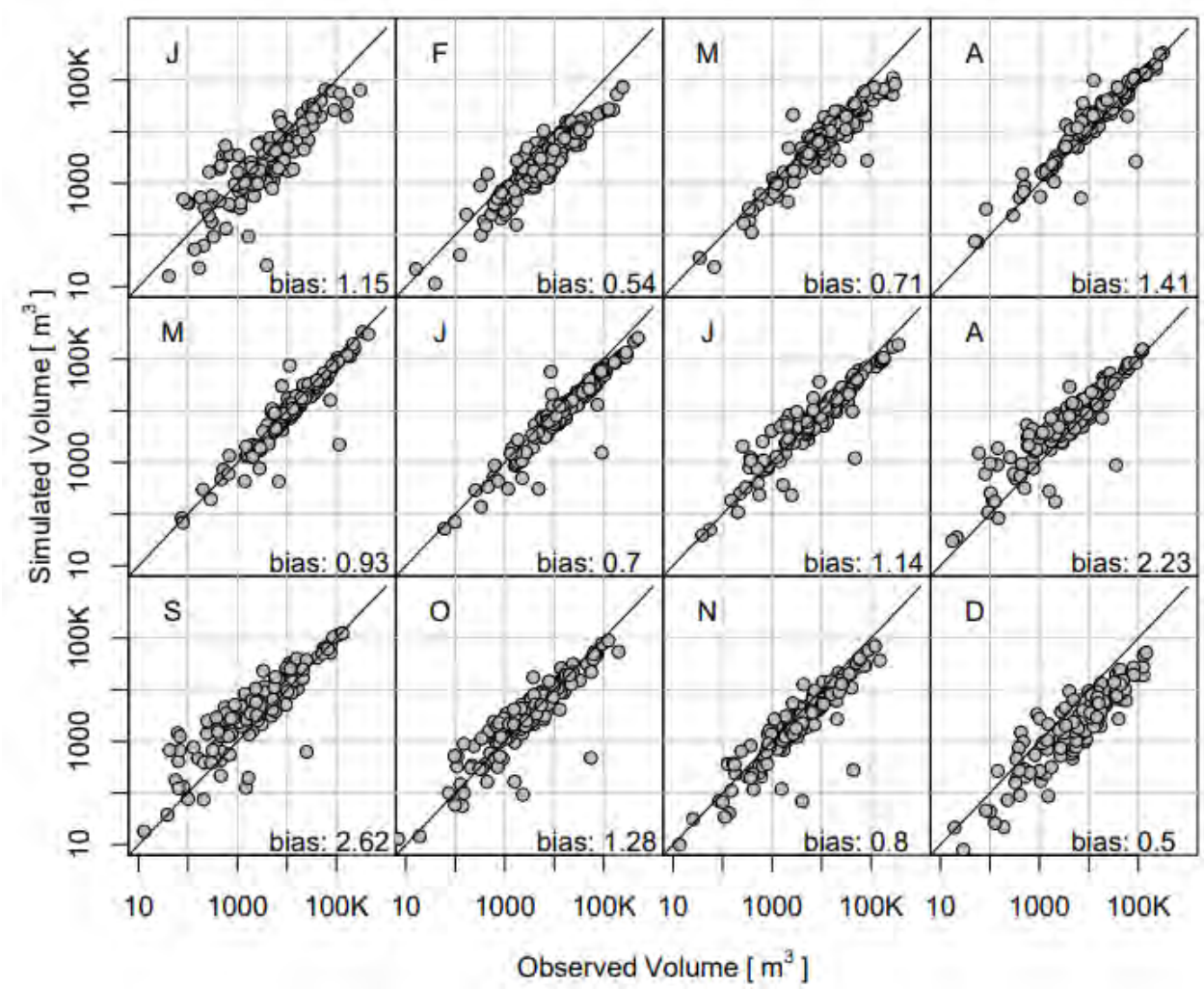

Figure S3-1. Ability of the HLM using daily precipitation, daily temperature, monthly ET derived from PRISM input to reproduce the observed average monthly volume at 150 USGS gauges in Iowa from 1981 to 2020 where instantaneous gauge data are available. The bias reported is the average ratio of simulation volume to USGS observation volume for all 150 gauges. A bias above (below) 1 indicates simulations overestimate (underestimate) volume compared observations. Months are indicated by the letters in the upper left corner.

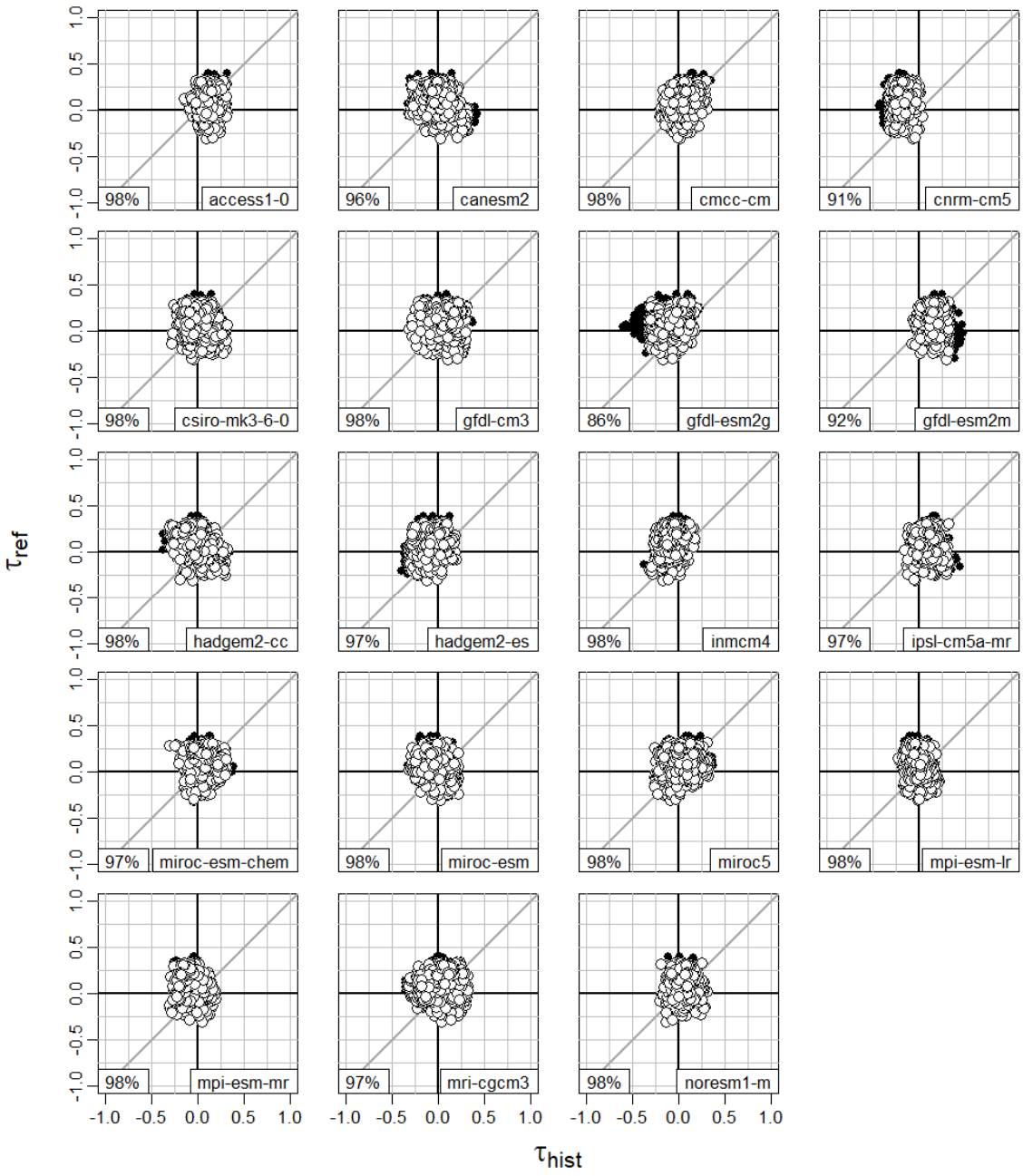

Figure S3-2. Comparison of Kendall's estimates for the historical period (1981-2005) from HLM simulations forced with PRISM rainfall (τ_{ref}) and GCM historical experiments (τ_{hist}). The white circles indicate communities where the test statistics match in sign and are statistically significant (5% level) or where both test statistics are not statistically significant. The percentage indicates the number of "correct detections" with respect to the total sample. Global climate models with percentages greater than 90% were deemed to match historical trends and selected for the primary subset of models.

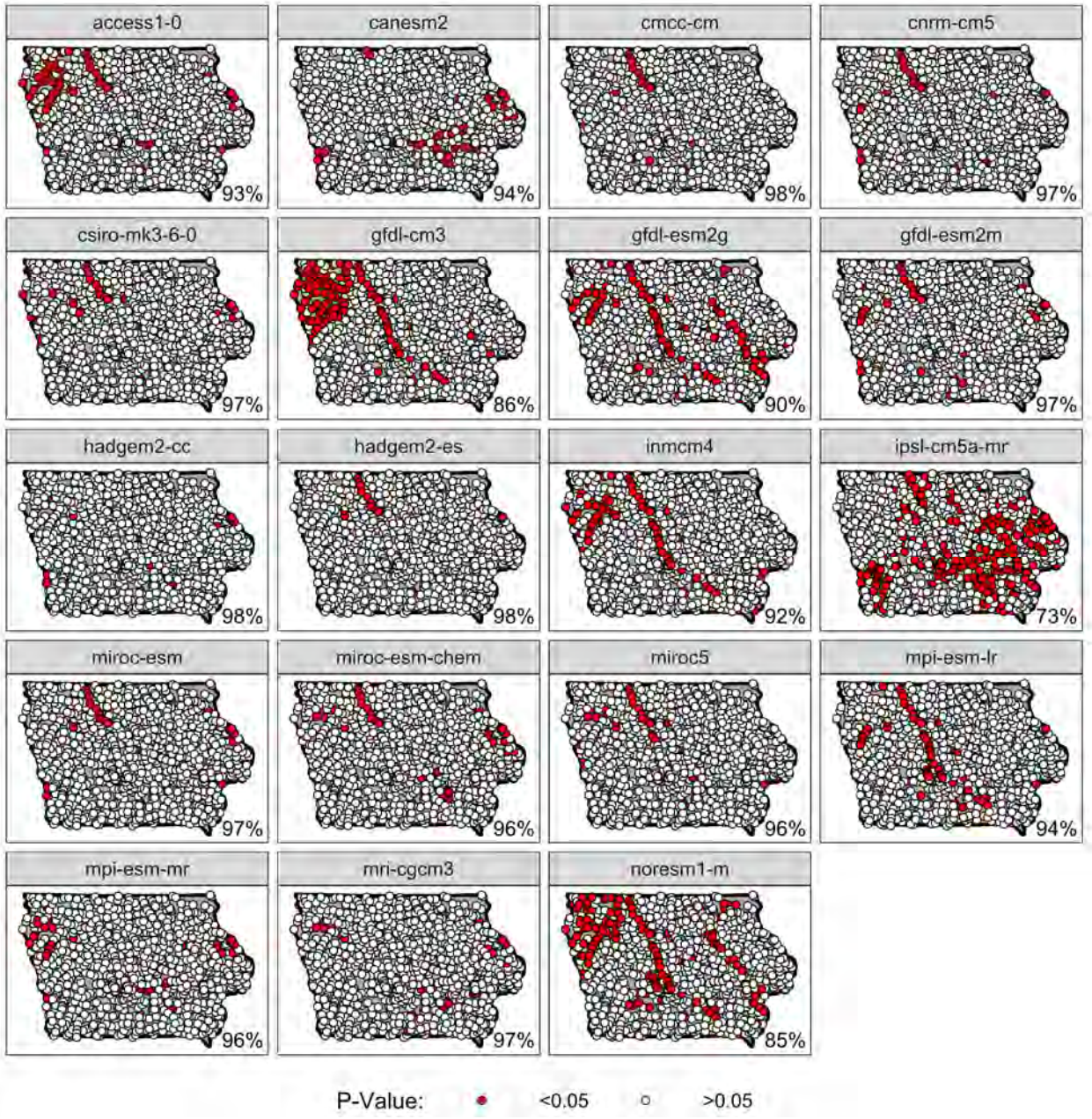

Figure S3-3. Mann-Whitney U test results for the historical period (1981-2005) between HLM simulations forced with PRISM rainfall and GCM historical experiments. The white circles indicate communities where the test statistics are not statistically significant at the 5% level. The percentage indicates the number of non-significant points with respect to the total sample.

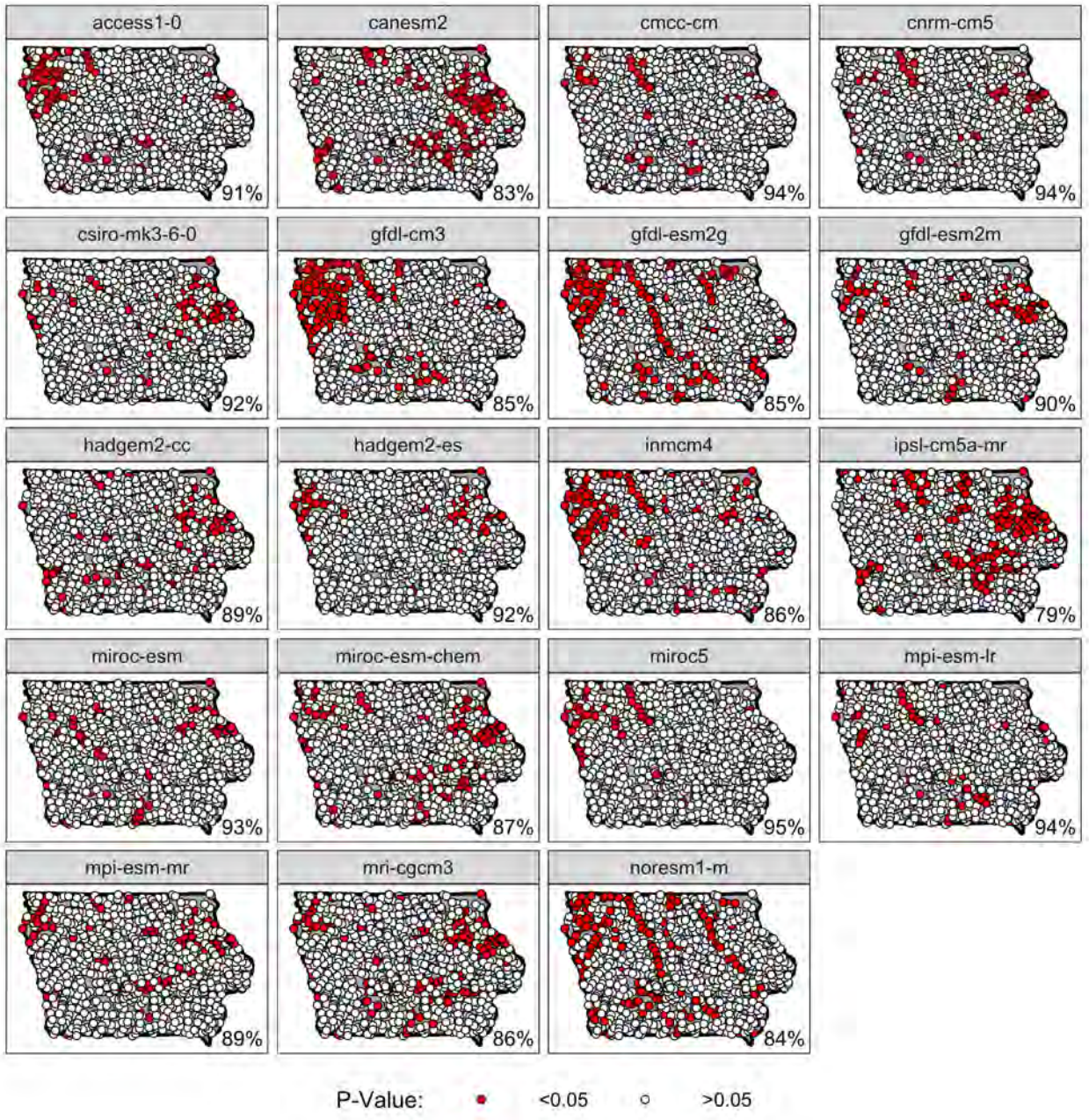

Figure S3-4. Kolmogorov-Smirnov test results for the historical period (1981-2005) between HLM simulations forced with PRISM rainfall and GCM historical experiments. The white circles indicate communities where the test statistics are not statistically significant at the 5% level. The percentage indicates the number of non-significant points with respect to the total sample.

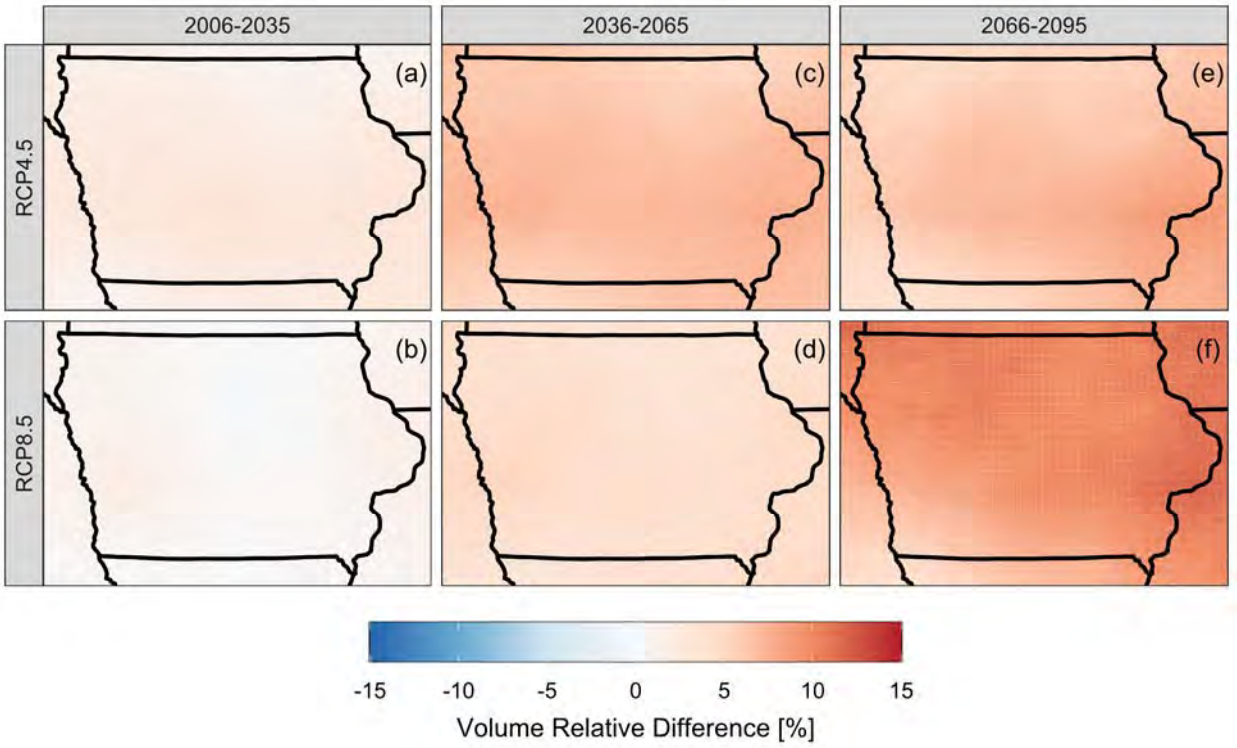

Figure S3-5. Relative differences in total precipitation volume between historical (1976-2005) and three future periods (2006-2035 in the left column; 2036-2065 in the middle column; 2066-2095 in the right column) for the two emission scenarios (one per row).

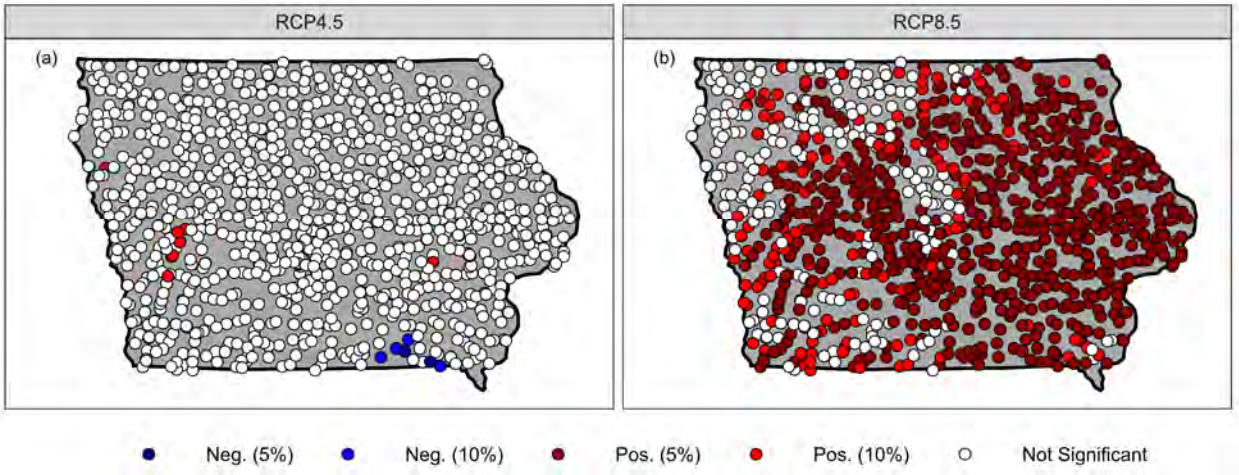

Figure S3-6. Annual maximum discharge trend from 2006 to 2099 determined by the Mann-Kendall trend test. The left (right) panel shows the results from simulations under the RCP4.5 (RCP8.5) emission scenario. Blue (red) circles indicate a negative (positive) trend in annual maxima. White circles are locations where the trend is not statistically significant. These results are based on all available 19 GCM as opposed to the subset of models in Fig. 3.

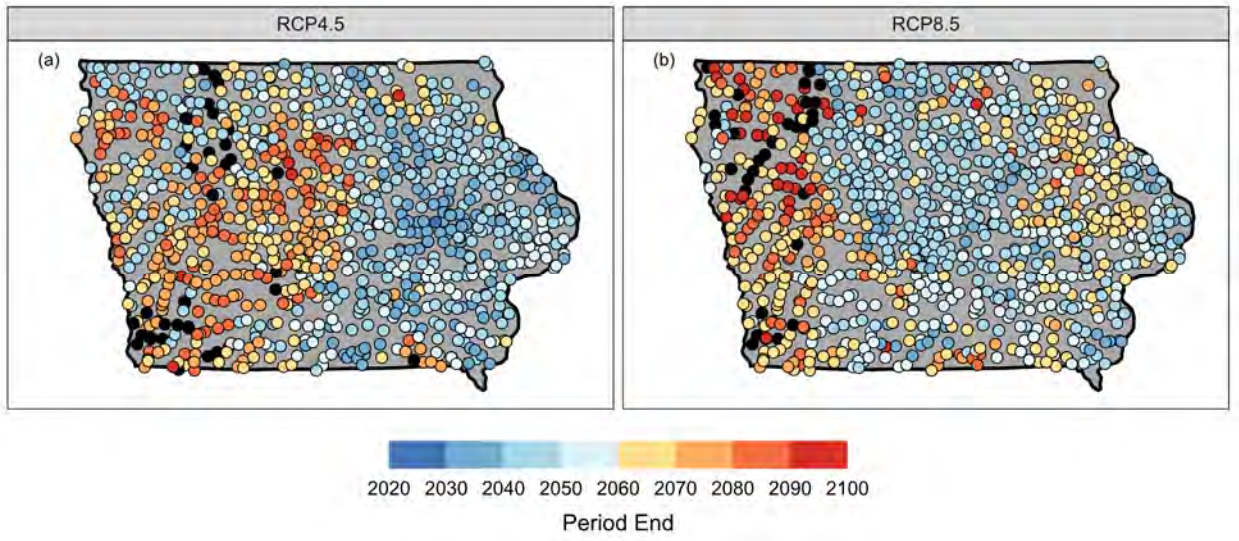

Figure S3-7. Time of emergence where the 50-year distribution of annual maxima discharge is significantly different ($p<0.05$) compared to the reference distribution from 1956 to 2005. The years associated with the color bars indicate the end of the comparison period (i.e., 2060 represents 2011 to 2060). Black points indicate sites where the p-value never went below 0.05 for the entire study period. For this figure, the analysis was performed using all available 19 GCMs as opposed to the subset shown in Fig. 5.

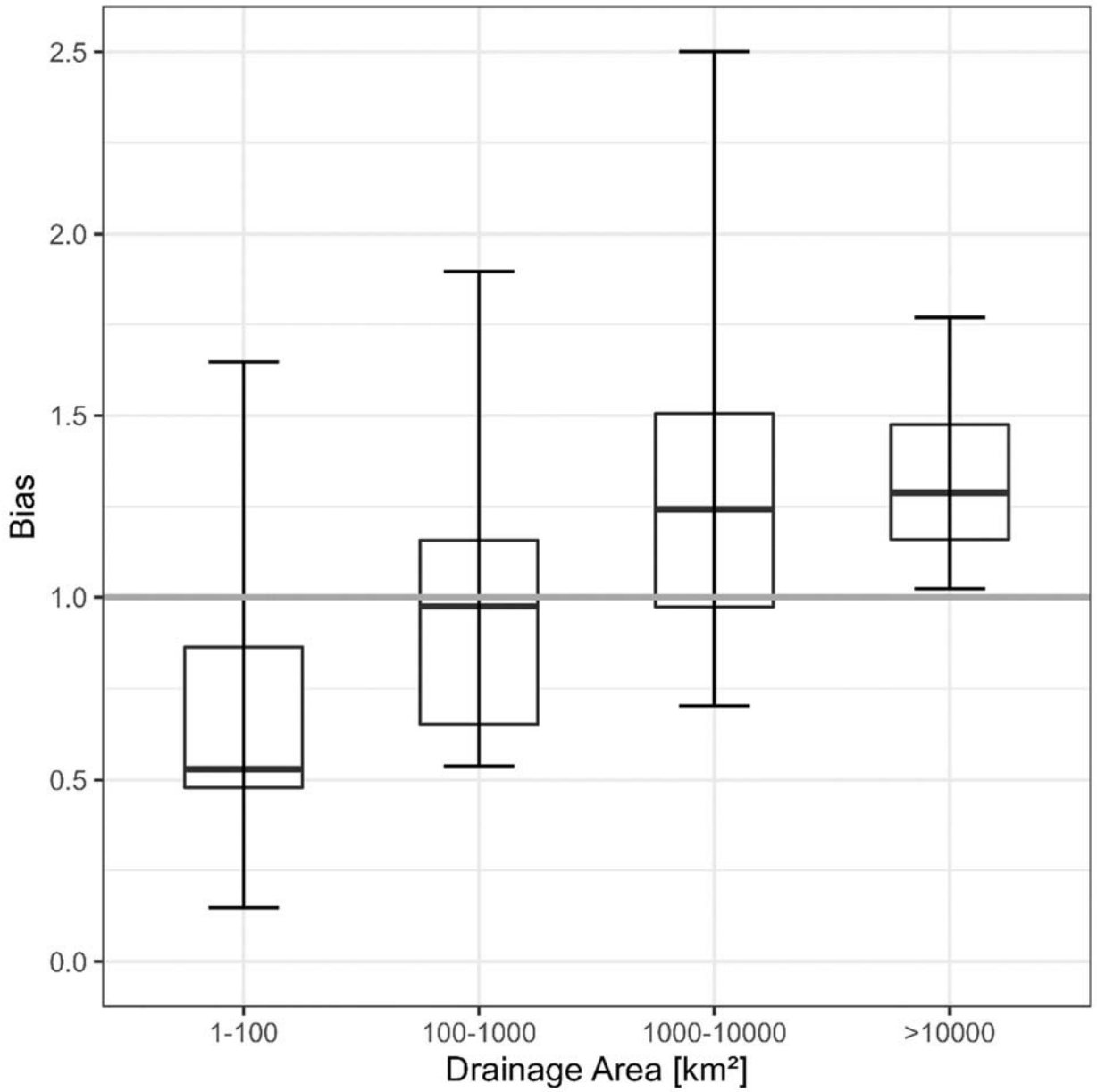

Figure S4-1. Bias between the PRISM-driven reference simulations and observations at USGS locations with over 30-years of annual maximum discharge data stratified by drainage area. A value greater (smaller) than 1 indicates an over-(under-) estimation of annual maximum discharge by the model compared to observations The dark grey line represents a bias of 1. The limits of the box represent the 25th and 75th percentiles, while the line in the middle the median; the whiskers represent the 5th and 95th percentiles. The breakdown of number of sites in each category is: 1-100 km^2 (22 sites), 100-1000 km^2 (30 sites), 1000-10000 km^2 (53 sites), >10000 km^2 (14 sites).

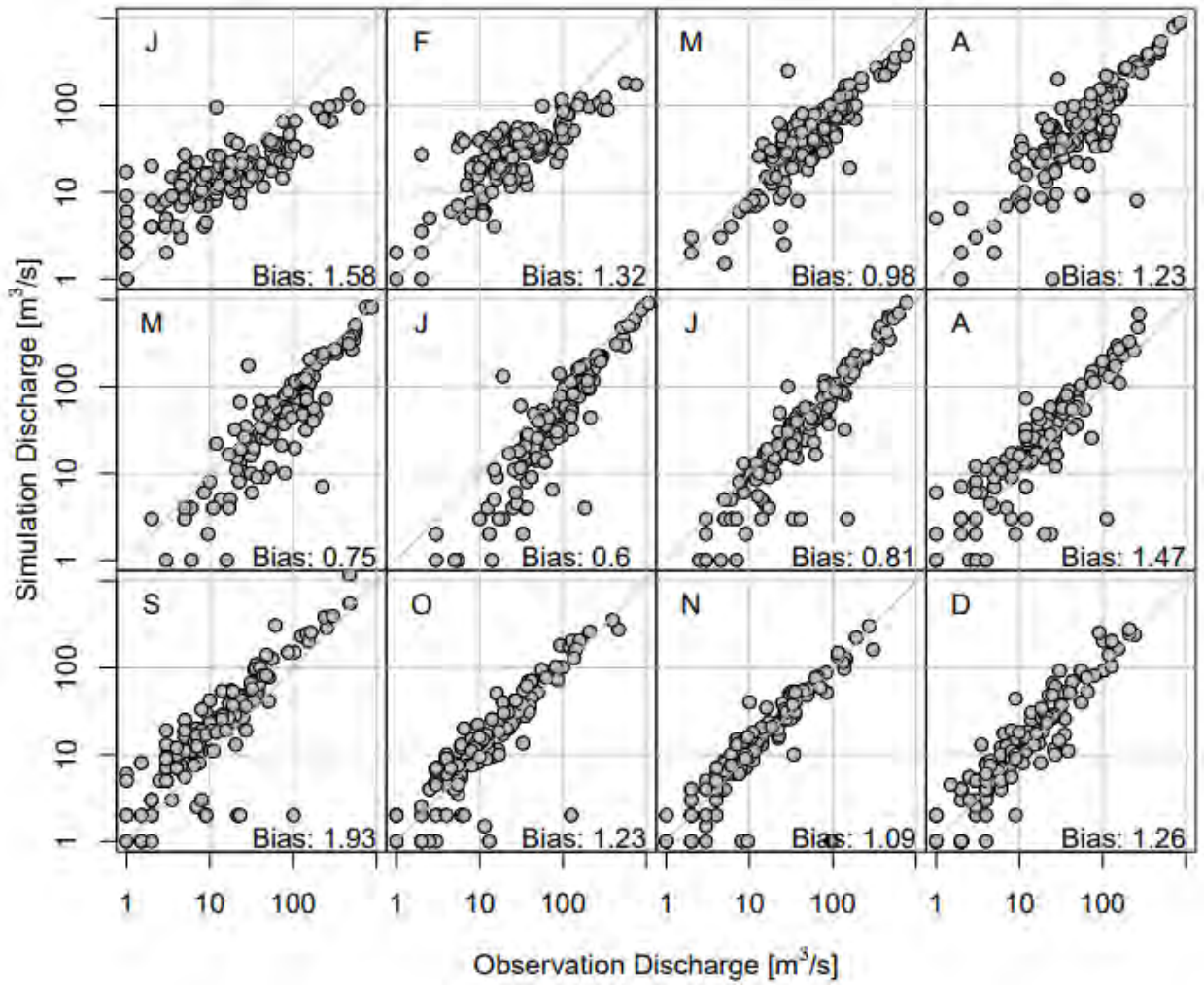

Figure S4-2. Scatterplot of monthly peak discharge between PRISM reference simulations and USGS observations from 1981 to 2021. The bias in the corner represents the average bias across all sites for each month. A bias greater (less) than zero indicates an overestimation (under) of monthly peak discharge.

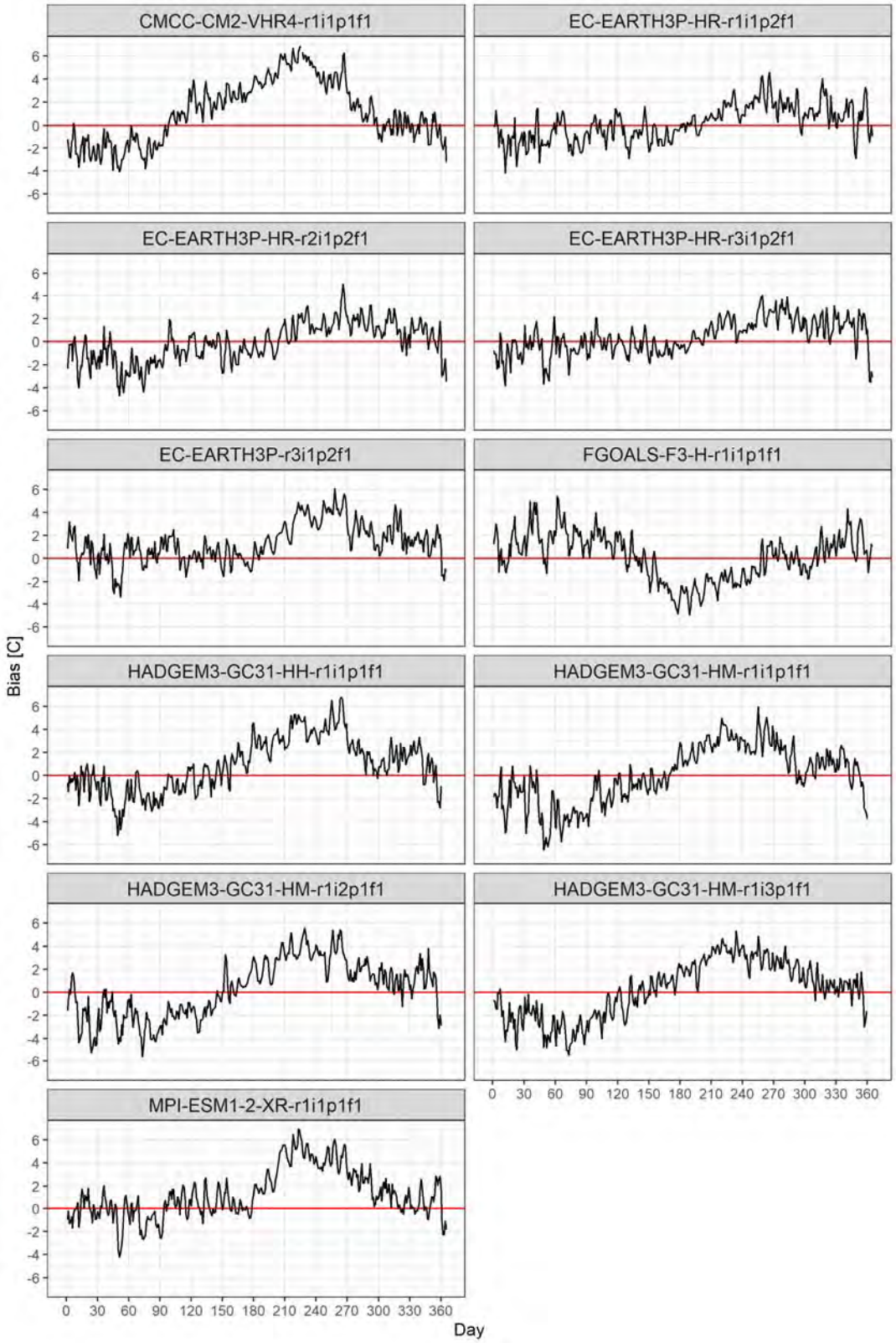

Figure S4-3. Average daily bias in temperature compared to PRISM temperature from 1981 to 2014. For the HLM, daily temperature is spatially averaged across the state. A bias greater (less) than zero (red line) Celsius indicates an overestimation (under) of temperature.

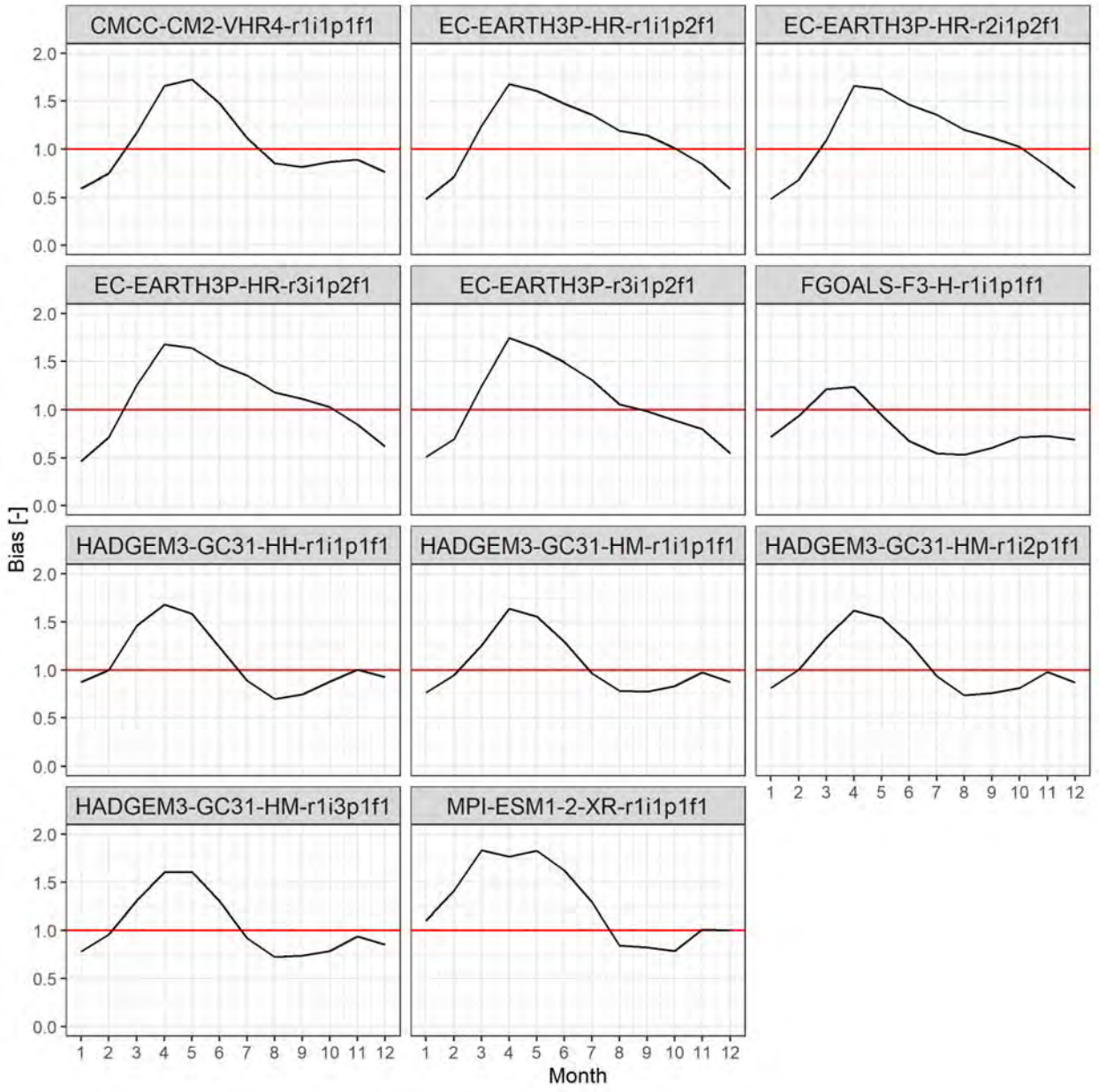

Figure S4-4. Average monthly bias in potential evapotranspiration (PET) compared to PRISM derived evapotranspiration from 1981 to 2014. For the HLM, monthly evapotranspiration is spatially averaged across the state. A bias greater (less) than one (red line) indicates an overestimation (under) of evapotranspiration.

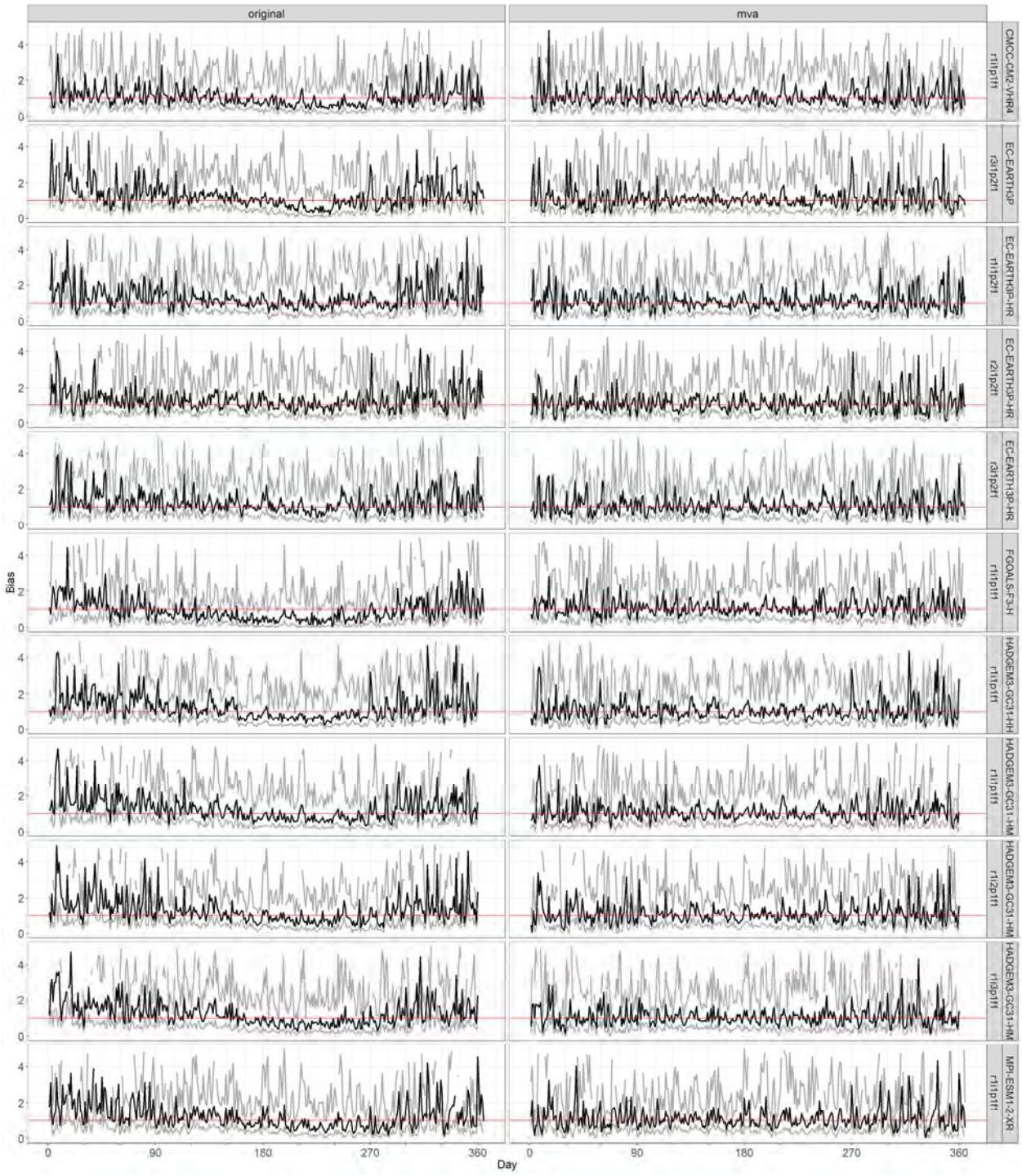

Figure S4-5. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and mva corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

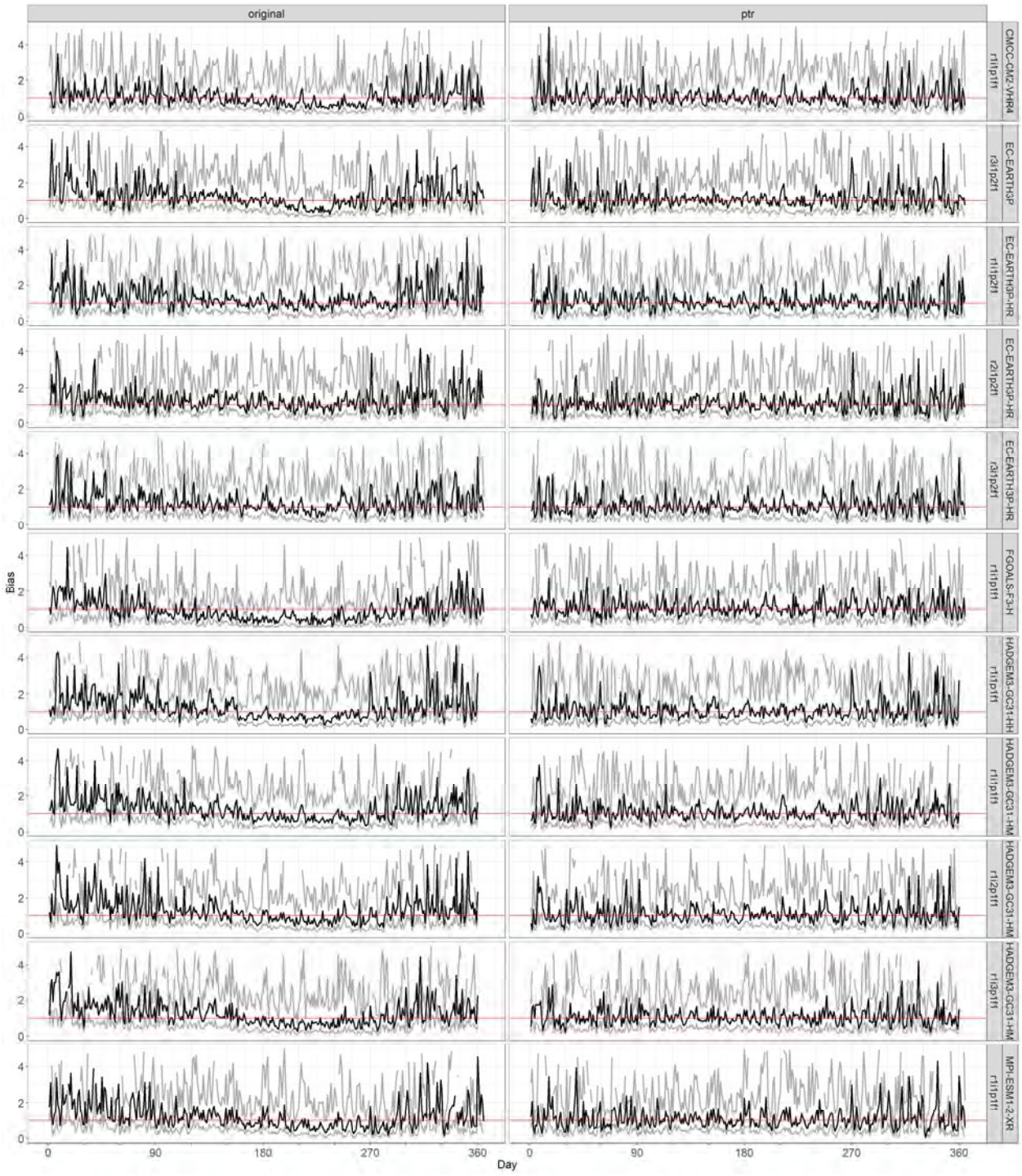

Figure S4-6. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and ptr corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

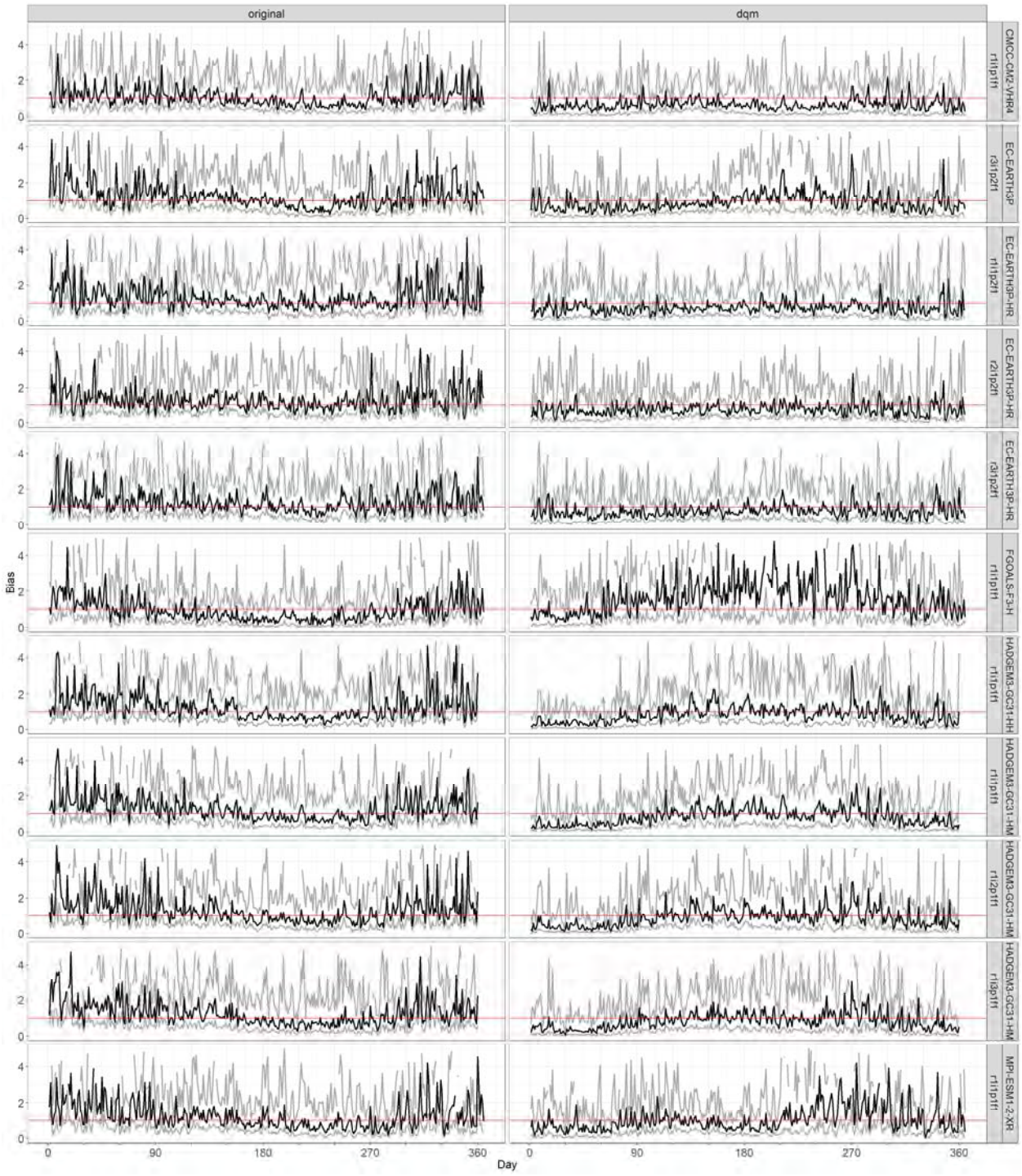

Figure S4-7. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and dqm corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

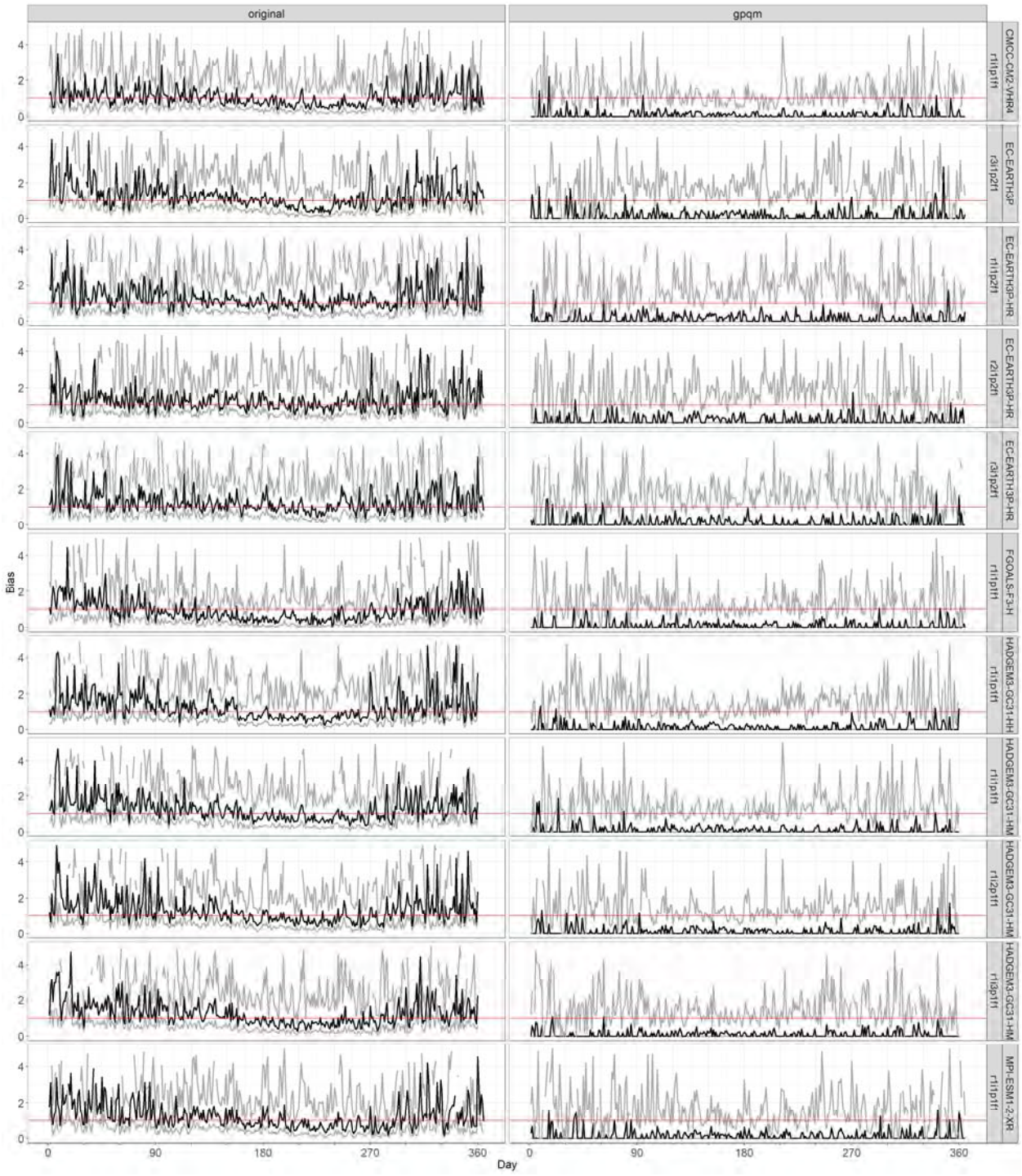

Figure S4-8. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and gpqm corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

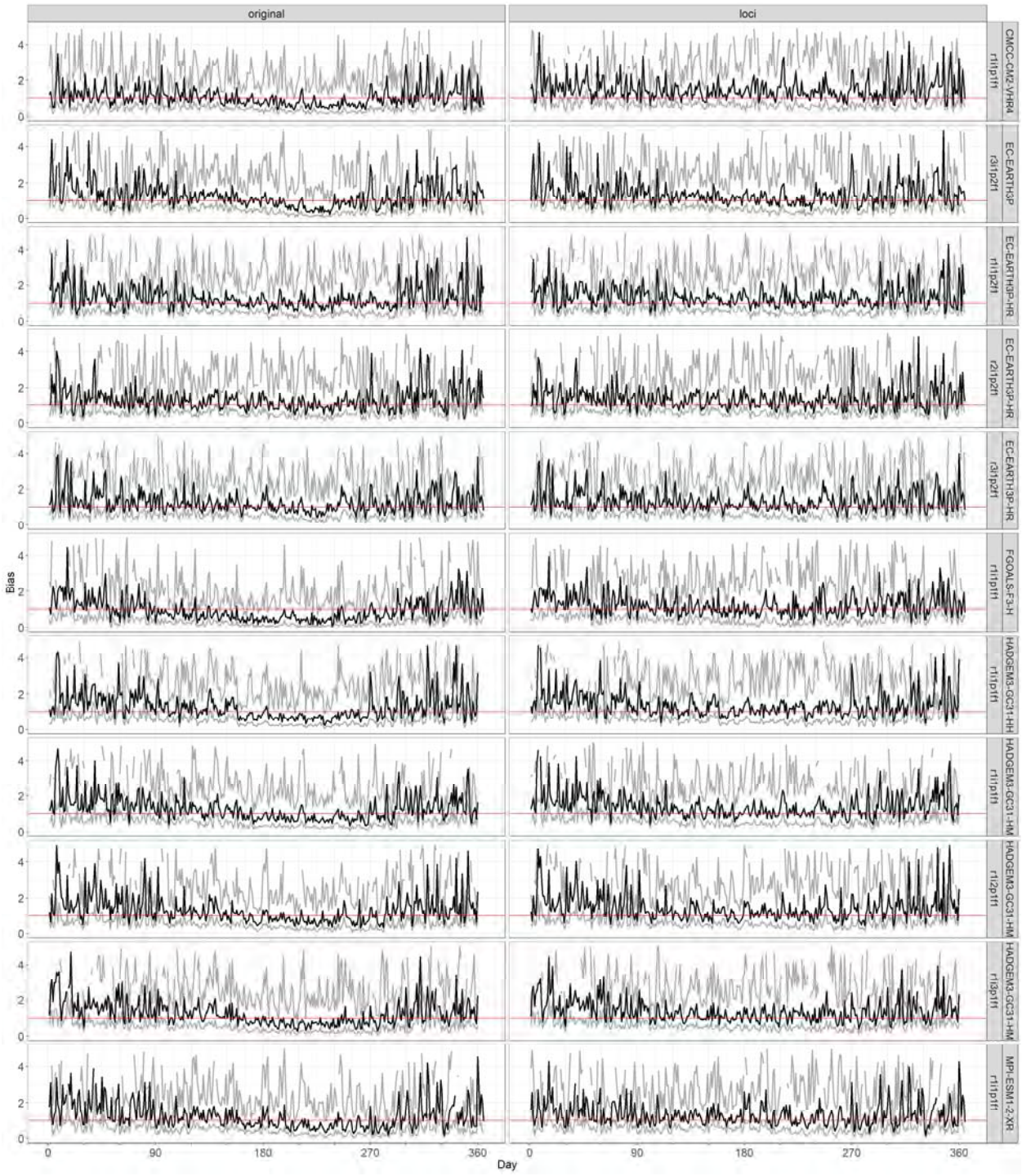

Figure S4-9 Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and loci corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

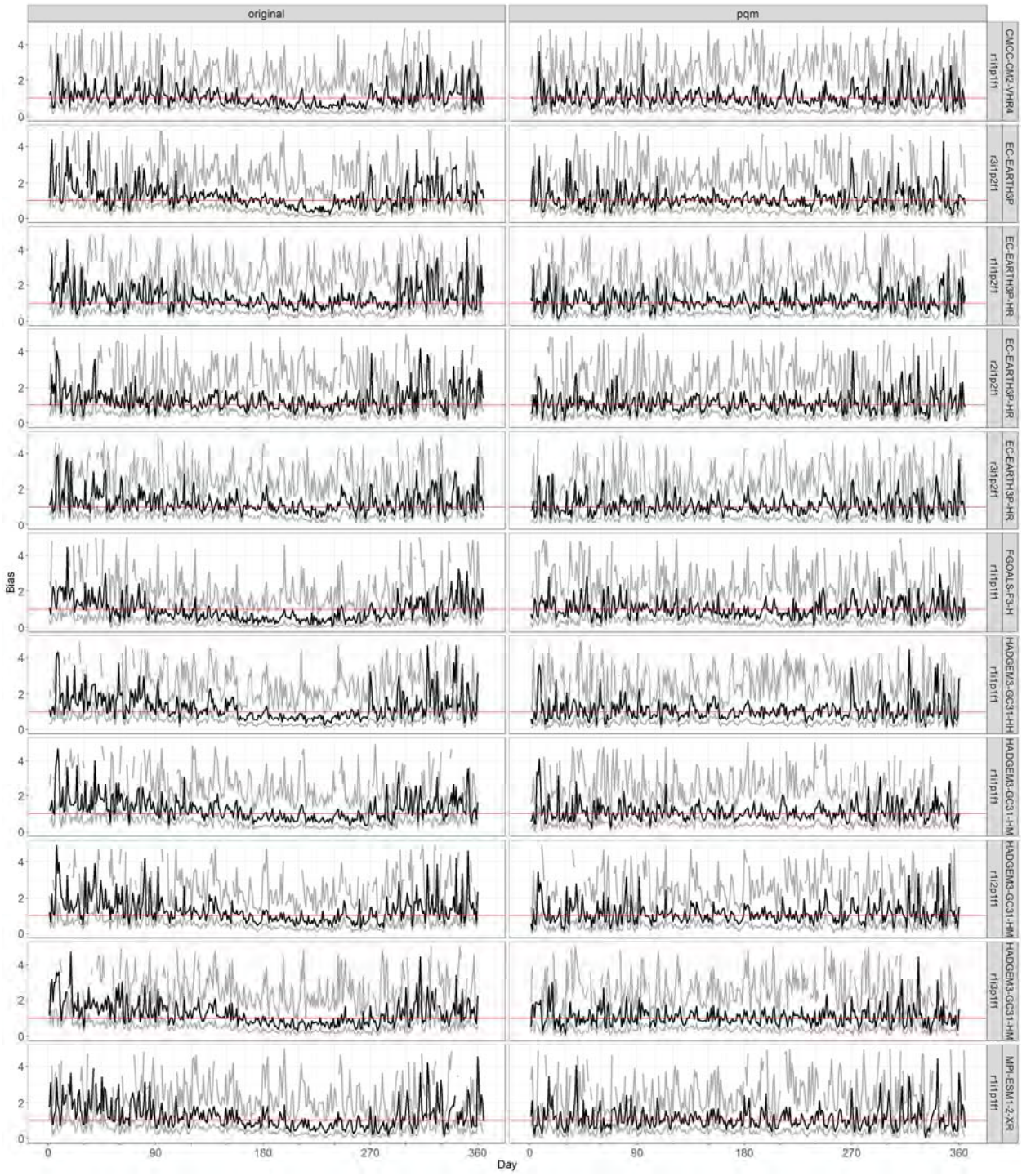

Figure S4-10. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and pqm corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

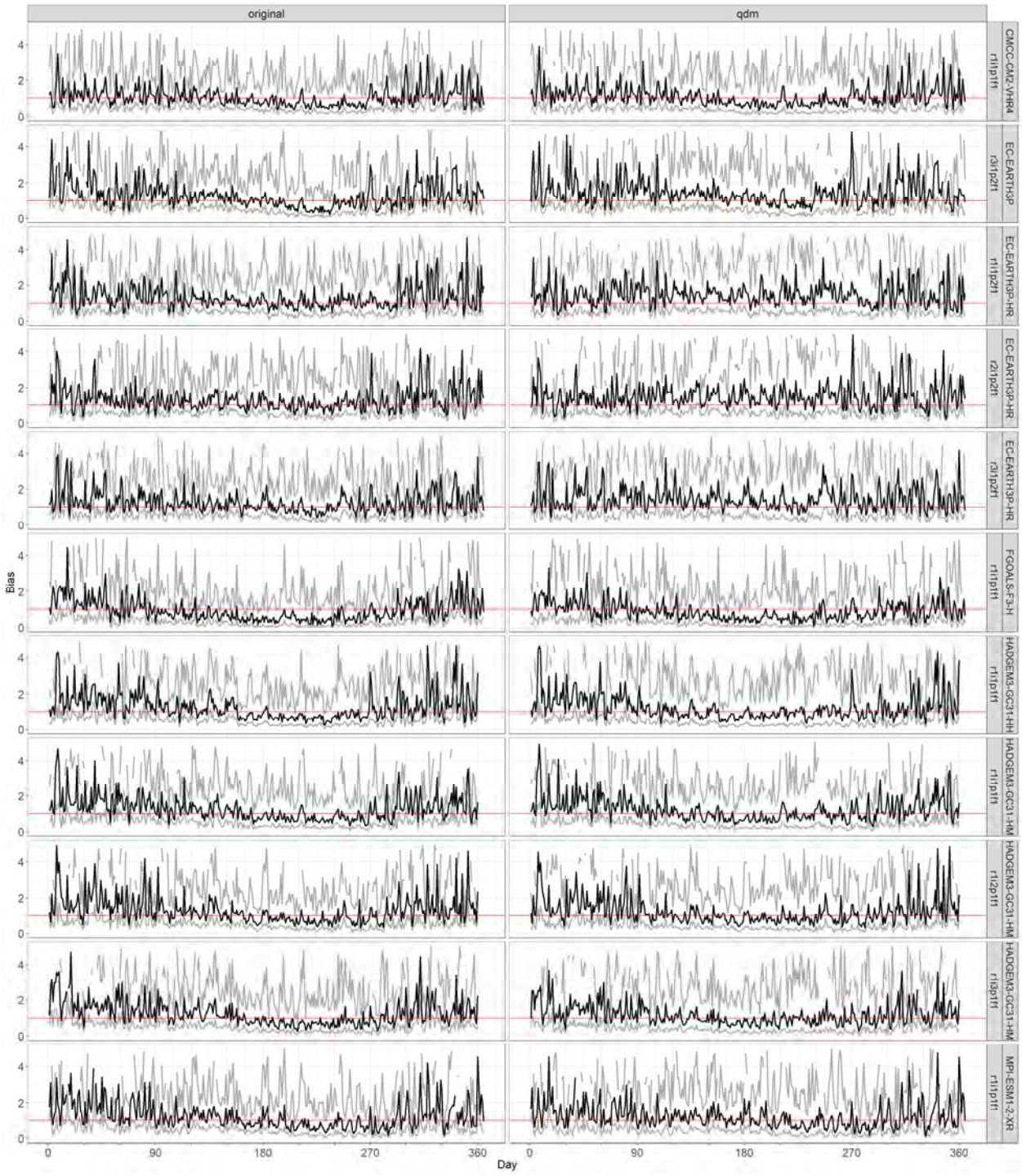

Figure S4-11. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and qdm corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

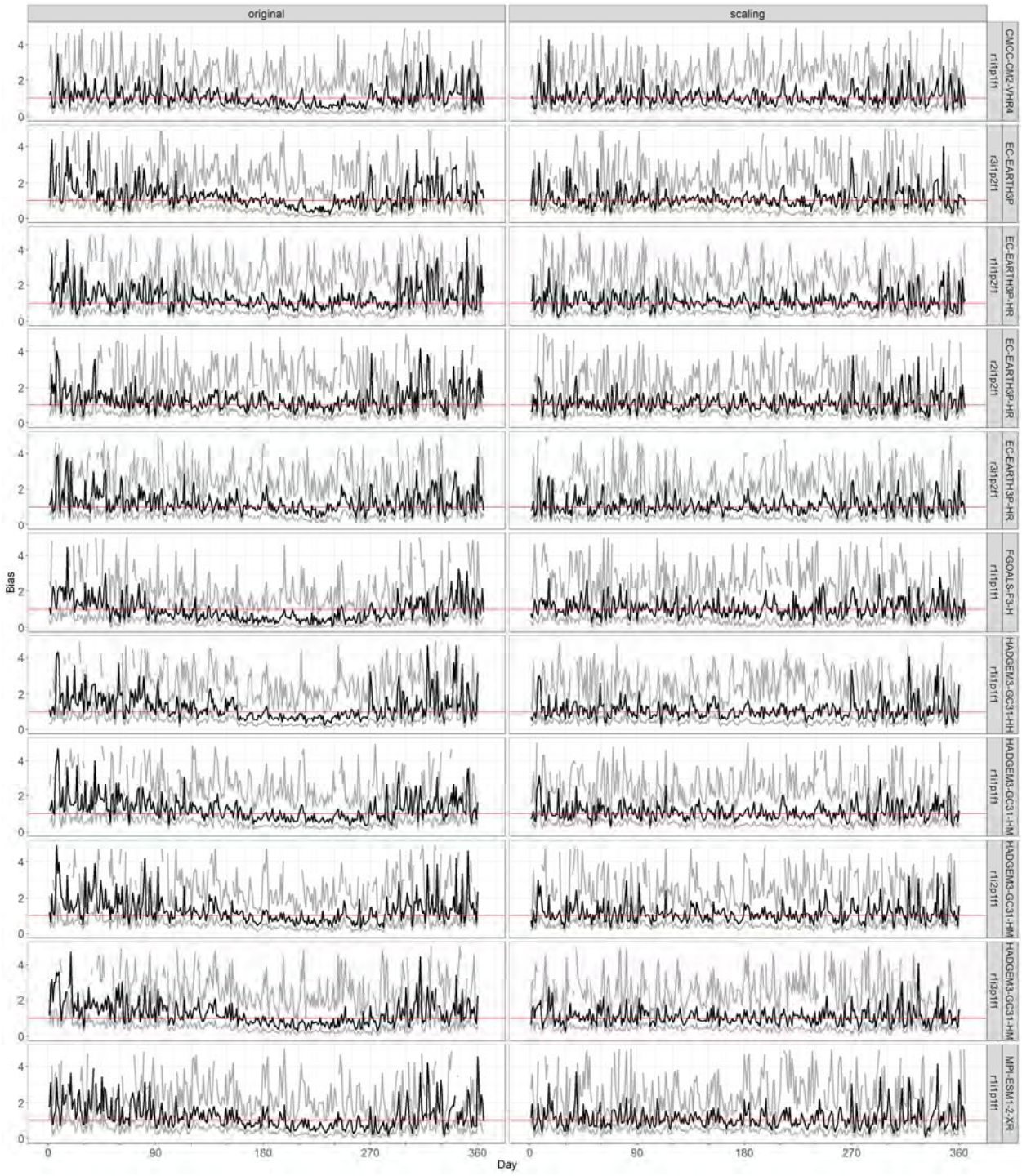

Figure S4-12. Average daily bias in precipitation based on leave-one-out cross validation (looCV). The columns are the results for original (before bias-correction and downscaling) and scaling corrected HighResMIP precipitation, respectively. The rows represent each of the climate models. The black line represents the median, and the gray lines represent the 5th and 95th quantiles of daily bias. The red line represents a bias of one or perfect match with respect to PRISM precipitation.

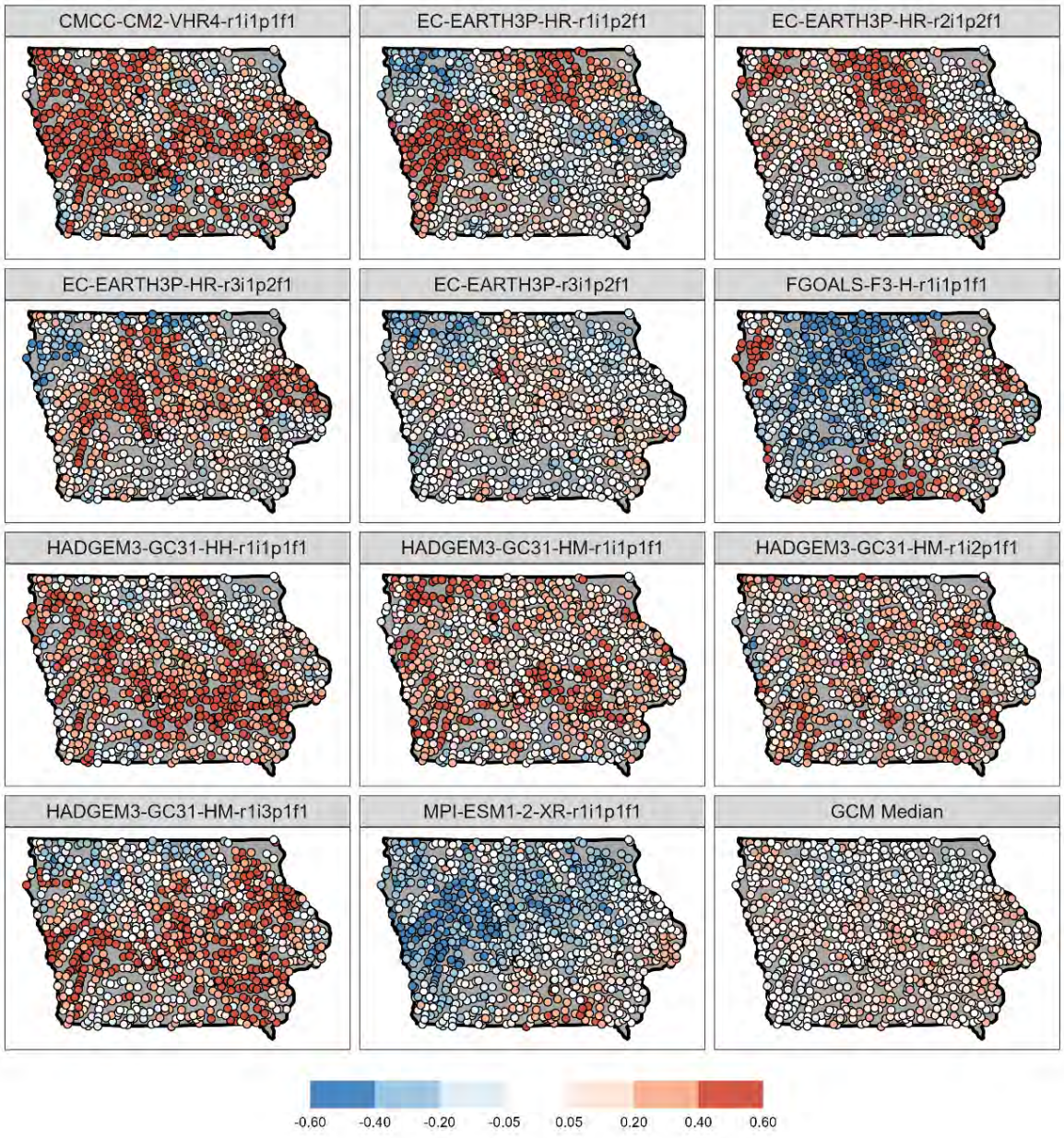

Figure S4-13. Relative difference (Eq. 4-2) in median annual maximum discharges between historical (1981-2014) and future (2017-2050) experiments from mva corrected simulations. The GCM Median represents the relative difference based on the median of the first member (r1) of the three GCMs.

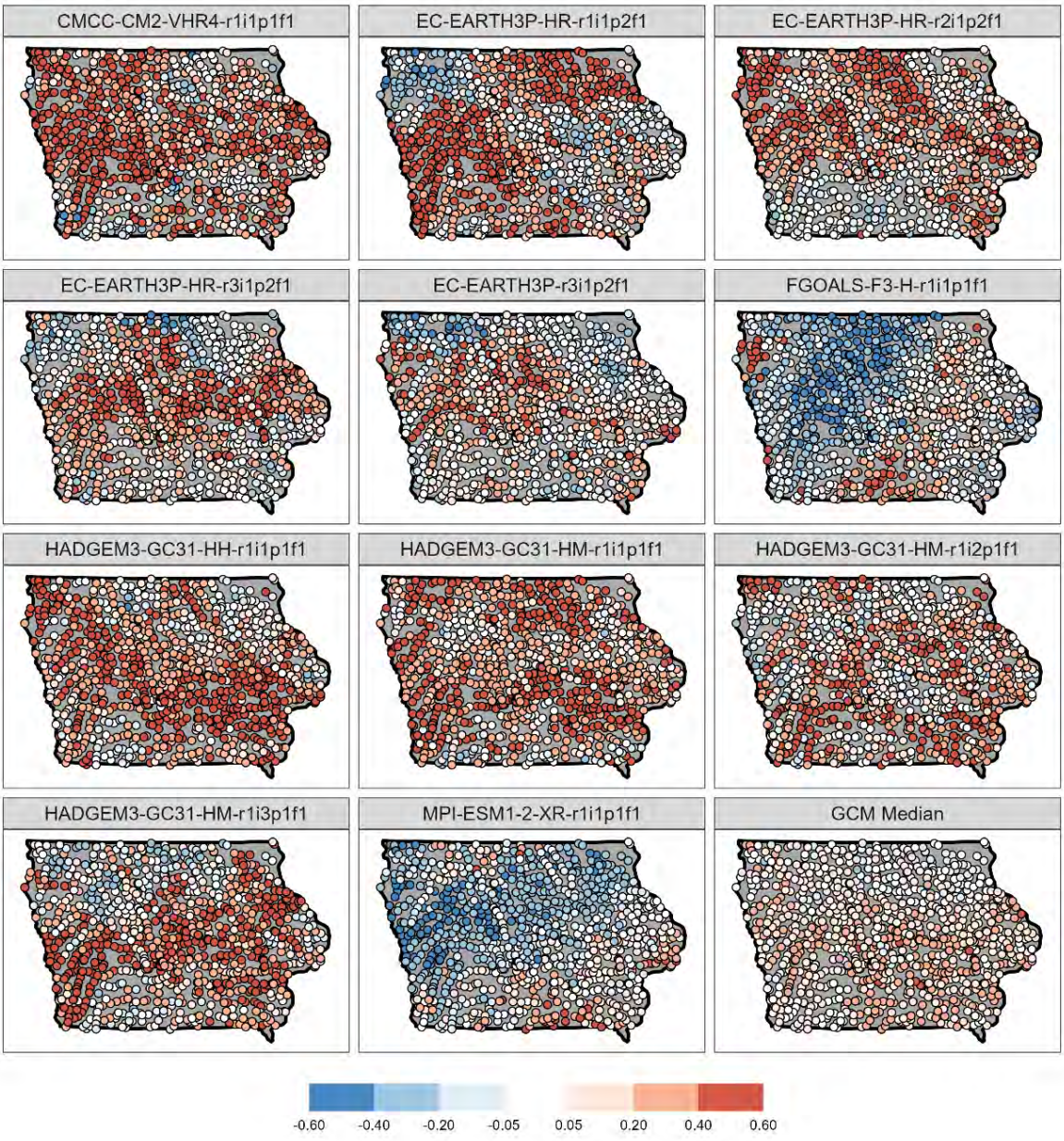

Figure S4-14. Relative difference (Eq. 4-2) in median annual maximum discharges between historical (1981-2014) and future (2017-2050) experiments from ptr corrected simulations. The GCM Median represents the relative difference based on the median of the first member (r1) of the three GCMs.

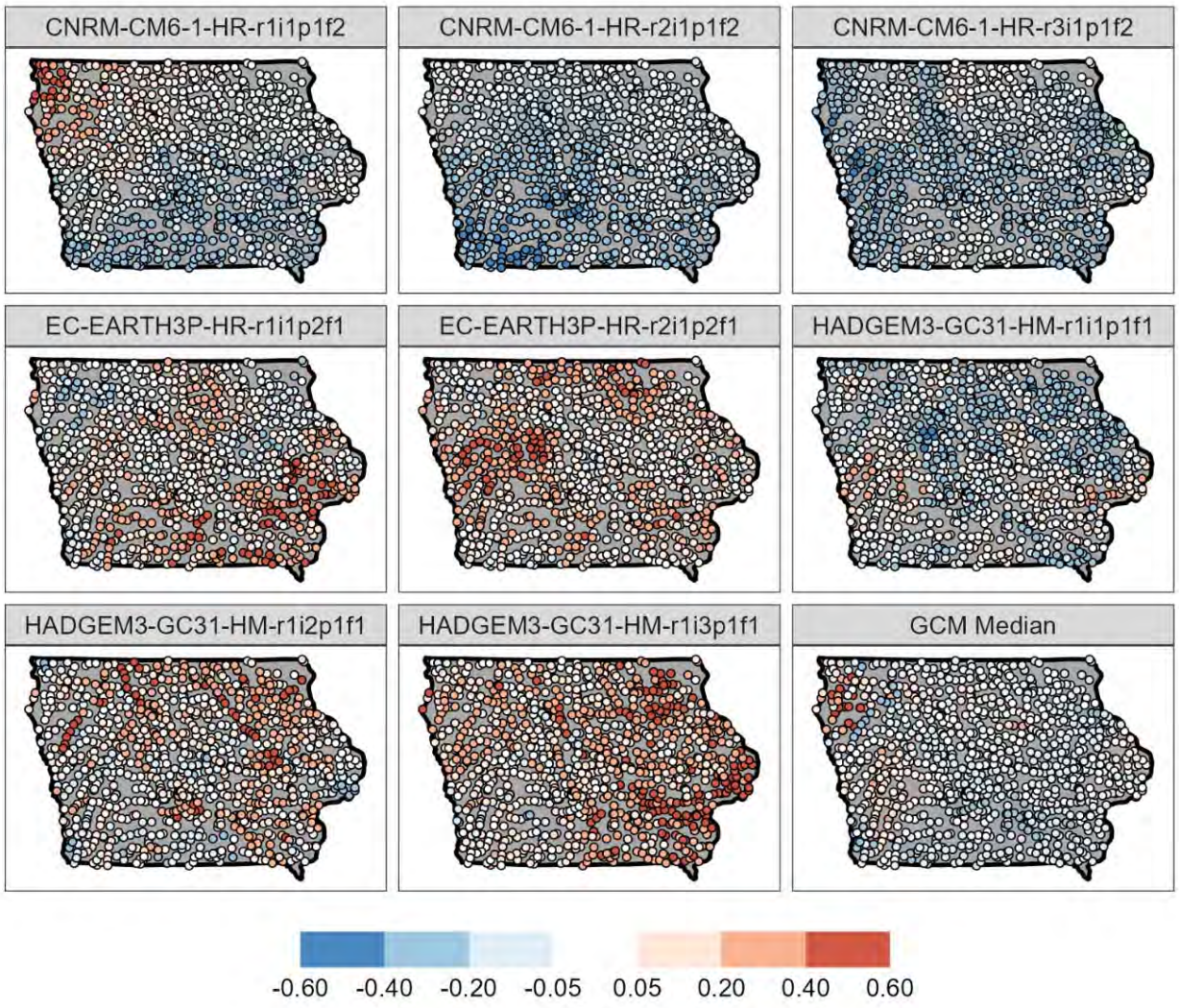

Figure S4-15. Relative difference (Eq. 4-2) in median annual maximum discharges between historical (1981-2014) and future (2017-2050) experiments from routed runoff simulations. The GCM Median represents the relative difference based on the median of the first member (r1) of the three GCMs.

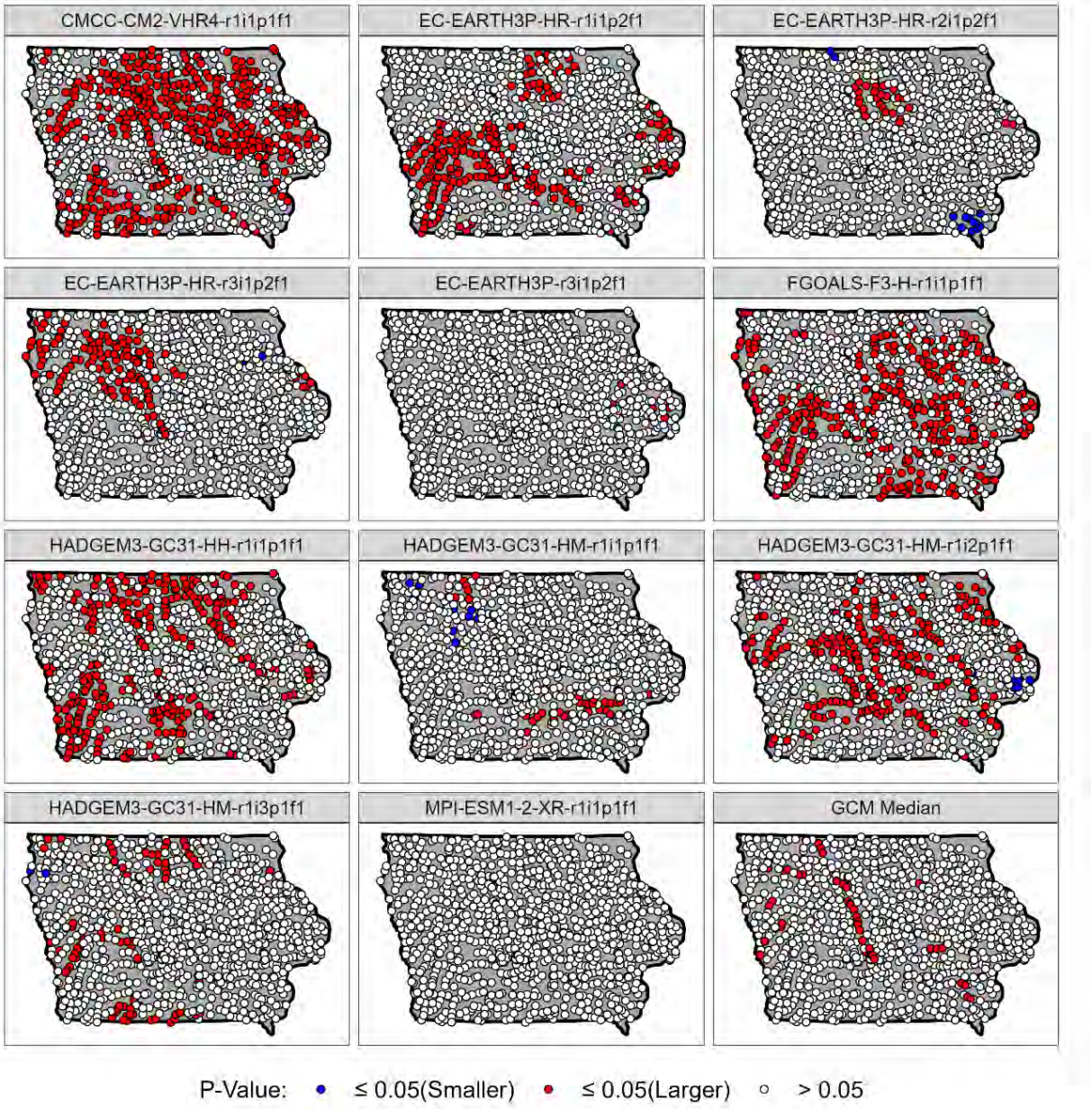

Figure S4-16. F-test p-values for annual maximum discharge between historical (1981-2014) and future (2017-2050) experiments for mva-based simulations. The red (blue) represent a larger (smaller) variance in future simulations compared to the historical period. The test results account for false discovery rate.

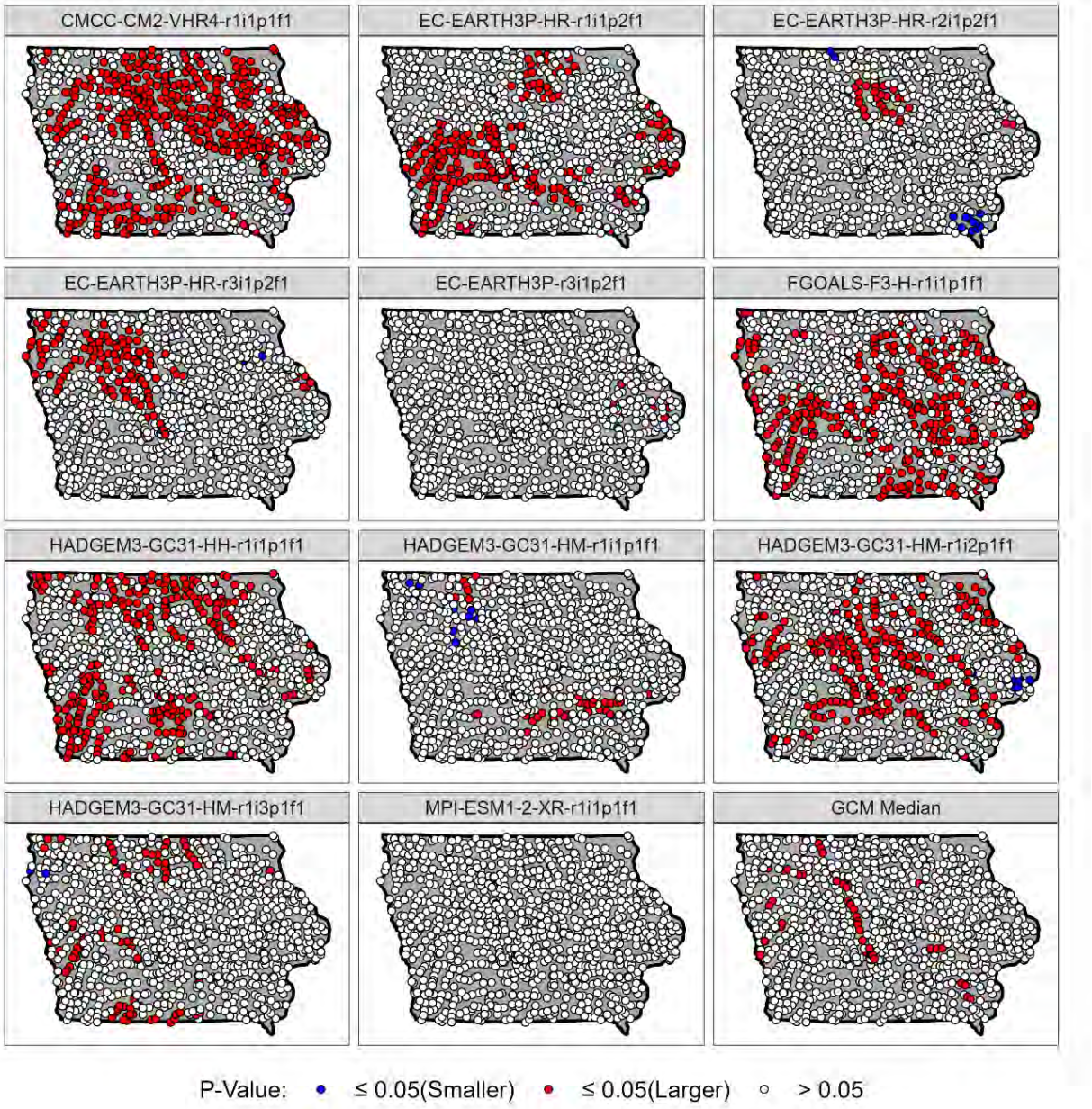

Figure S4-17. F-test p-values for annual maximum discharge between historical (1981-2014) and future (2017-2050) experiments for ptr-based simulations. The red (blue) represent a larger (smaller) variance in future simulations compared to the historical period. The test results account for false discovery rate.

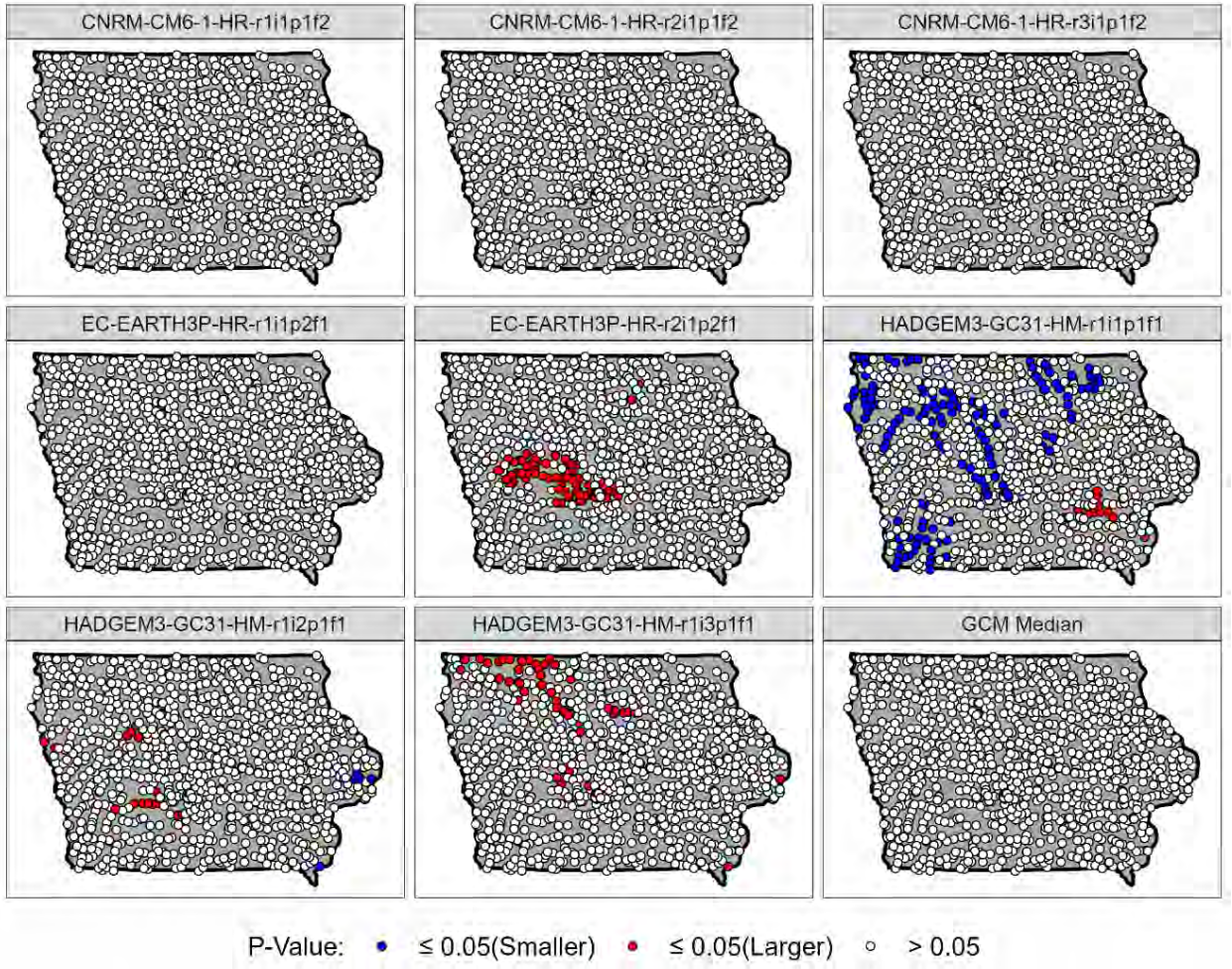

Figure S4-18. F-test p-values for annual maximum discharge between historical (1981-2014) and future (2017-2050) experiments for runoff-based simulations. The red (blue) represent a larger (smaller) variance in future simulations compared to the historical period. The test results account for false discovery rate.