Elsevier Editorial System(tm) for Journal of Hydrology or its open access mirror
Manuscript Draft

Manuscript Number: HYDROL32937R1

Title: Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

Article Type: Research paper

Keywords: frequency of flood events; projections; statistical modeling; CORDEX; LOCA; CMIP5

Corresponding Author: Dr. Andrea Neri,

Corresponding Author's Institution:

First Author: Andrea Neri

Order of Authors: Andrea Neri; Gabriele Villarini; Francesco Napolitano

Abstract: There is growing empirical evidence that many river basins across the U.S. Midwest have been experiencing an increase in the frequency of flood events over the most recent decades. Albeit these detected changes are important to understand what happened in our recent past, they cannot be directly extrapolated to obtain information about possible future changes in the frequency of flood events. Building on recent statistically-based attribution studies, we project seasonal changes in the frequency of flood events at 286 U.S. Geological Survey gauging stations across the U.S. Midwest using projections of precipitation, antecedent wetness conditions and temperature as drivers. The projections of the covariates are obtained from two datasets obtained by downscaling global circulation models from the Fifth Coupled Model Intercomparison Project (CMIP5). We focus on the representative concentration pathway (RCP) 8.5 and on four different flood thresholds (i.e., from more common to less frequent flood events). We find that the frequency of flood events during the 21st century increases during spring at most of the analyzed gauging stations, with larger changes in the Northern Great Plains and regardless of the flood threshold value. Our findings also point to a projected increasing number of flood events during the winter, especially in the stations in the southern and western part of the domain (Iowa, Missouri, Illinois, Ohio, Indiana and Michigan). A marked change in the frequency of flood events is not projected for the summer and fall.

Suggested Reviewers: Stacey Archfield
sarch@usgs.gov

Kenneth Porter
kwpotter@facstaff.wisc.edu

Robert Hirsch
rhirsch@usgs.gov

Karen Ryberg
kryberg@usgs.gov

David Hannah
d.m.hanna@bham.ac.uk

***Declaration of Interest Statement**

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

by

ANDREA NERI, GABRIELE VILLARINI, AND FRANCESCO NAPOLITANO

- We use statistical models to project changes in the frequency of flood events
- We use precipitation, temperature and antecedent wetness conditions as drivers
- The drivers are derived from centennial projections by global circulation models
- The biggest projected changes in the frequency of flooding are in spring and winter

Abstract
Click here to download Abstract: Abstract_6.docx

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

ANDREA NERI[1], GABRIELE VILLARINI[2], AND FRANCESCO NAPOLITANO[1]

[1] Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Rome, Italy.

[2] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, Iowa, USA.

Manuscript submitted to

Journal of Hydrology

29 March 2019

Corresponding author address:

Andrea Neri, Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Via Eudossiana 18, 00184, Rome, Italy. E-mail: andrea.neri@uniroma1.it. Tel.: +390644585063

ABSTRACT

There is growing empirical evidence that many river basins across the U.S. Midwest have been experiencing an increase in the frequency of flood events over the most recent decades. Albeit these detected changes are important to understand what happened in our recent past, they cannot be directly extrapolated to obtain information about possible future changes in the frequency of flood events. Building on recent statistically-based attribution studies, we project seasonal changes in the frequency of flood events at 286 U.S. Geological Survey gauging stations across the U.S. Midwest using projections of precipitation, antecedent wetness conditions and temperature as drivers. The projections of the covariates are obtained from two datasets obtained by downscaling global circulation models from the Fifth Coupled Model Intercomparison Project (CMIP5). We focus on the representative concentration pathway (RCP) 8.5 and on four different flood thresholds (i.e., from more common to less frequent flood events). We find that the frequency of flood events during the 21st century increases during spring at most of the analyzed gauging stations, with larger changes in the Northern Great Plains and regardless of the flood threshold value. Our findings also point to a projected increasing number of flood events during the winter, especially in the stations in the southern and western part of the domain (Iowa, Missouri, Illinois, Ohio, Indiana and Michigan). A marked change in the frequency of flood events is not projected for the summer and fall.

*Revised Manuscript with no changes marked
Click here to view linked References

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

ANDREA NERI[1], GABRIELE VILLARINI[2], AND FRANCESCO NAPOLITANO[1]

[1] Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Rome, Italy.

[2] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, Iowa, USA.

Manuscript submitted to

Journal of Hydrology

18 August 2019

Corresponding author address:

Andrea Neri, Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Via Eudossiana 18, 00184, Rome, Italy. E-mail: andrea.neri@uniroma1.it.

Tel.: +390644585063

Abstract

There is growing empirical evidence that many river basins across the U.S. Midwest have been experiencing an increase in the frequency of flood events over the most recent decades. Albeit these detected changes are important to understand what happened in our recent past, they cannot be directly extrapolated to obtain information about possible future changes in the frequency of flood events. Building on recent statistically-based attribution studies, we project seasonal changes in the frequency of flood events at 286 U.S. Geological Survey gauging stations across the U.S. Midwest using projections of precipitation, antecedent wetness conditions and temperature as drivers. The projections of the covariates are obtained from two datasets obtained by downscaling global circulation models from the Fifth Coupled Model Intercomparison Project (CMIP5). We focus on the representative concentration pathway (RCP) 8.5 and on four different flood thresholds (i.e., from more common to less frequent flood events). We find that the frequency of flood events during the 21st century increases during spring at most of the analyzed gauging stations, with larger changes in the Northern Great Plains and regardless of the flood threshold value. Our findings also point to a projected increasing number of flood events during the winter, especially in the stations in the southern and western part of the domain (Iowa, Missouri, Illinois, Ohio, Indiana and Michigan). A marked change in the frequency of flood events is not projected for the summer and fall.

Keywords: frequency of flood events; projections; statistical modeling; CORDEX; LOCA; CMIP5

1 Introduction

The hydrologic impacts of climate change have been the topic of a growing body of research, and have attracted significant interest from decision makers and stakeholders in terms of their projected changes and related societal and economic impacts. This awareness, together with the evidence of other natural disasters attributed to climate change (e.g., Seneviratne et al., 2012), has recently spread across many countries and pushed governments to react in terms of both adaptation to and management of extreme flood events (e.g., Lavell and Oppenheimer, 2012). For instance, the Government of Canada developed the Federal Floodplain Mapping Framework, which is a document intended to describe the entire process to define reliable flood risk maps and the effect that climate alterations have on them (Natural Resources Canada, 2017); the Australian Disaster Resilience Handbook 7 (AIDR, 2017) describes the main practices and activities for a proper floodplain management, focusing on the land use and development and on how climate change affects flood modeling; in the United States, four federal agencies (U.S. Geological Survey (USGS), U.S. Army Corps of Engineers, Bureau of Reclamation and National Oceanic and Atmospheric Administration) prepared a report (Brekke et al., 2009) which proposes better practices and activities for water resources management by considering the effects of global warming.

Although there is still low confidence about the changes of the frequency and/or magnitude of flood events at the global scale (Seneviratne et al., 2012), these recent actions and strategies for flood risk management have been encouraged by several studies that detect statistically significant trends in flooding at the regional level. For instance, focusing on the continental United States, Mallakpour and Villarini (2015) analyzed daily streamflow records at 774 USGS stream gauge stations across central United States covering the period from 1962 to

2011, and detected statistically significant increases in the frequency of flood events for the majority of the stations (see also Neri et al. (2019b)). Slater and Villarini (2016) showed that the frequency of the water level exceeding the National Weather Service's four flood level categories in 2042 water basins across the United States was subject to significant changes, with different parts of the country exhibiting spatially coherent signals of increasing or decreasing trends. Archfield et al. (2016) analyzed the frequency, duration, magnitude and volume of floods at 345 streamgages across the United States, showing that certain regions present significant changes in these flood properties. For a recent review of the detected changes in flooding across the continental United States, see Villarini and Slater (2017).

Even though these findings represent a key step towards our improved characterization of the changes in flood events over the past several decades, they do not provide useful information about projected changes in flood-related quantities. To address this knowledge gap, different methods have been proposed and developed, which can be classified into two broad classes: hydrological and statistical models (e.g., Chang and Chen, 2018; Eldho and Kulkarni, 2017; François et al., 2019; Giuntoli et al., 2018; Villarini et al., 2015). The former aims to predict future flood conditions by using projections of the hydrologic forcings as input to physical or conceptual equations that describe the main processes regulating the transformation of precipitation into runoff, as for instance through the use of global impact models (e.g., Dankers et al., 2014). For instance, Arnell and Gosling (2016) analyzed the effects of climate change on global flood risk by combining projections of population and of climate variables of 21 global climate models (GCMs) to force the Mac-PDM.09 hydrological model (Gosling and Arnell, 2011) to obtain flood frequency curves at 0.5° resolutions. Alfieri et al. (2017) assessed global projections of the frequency and magnitude of river floods using dynamical downscaled projections of climate variables by seven GCMs as input to the LISFLOOD hydrological model (van der Knijff et al., 2010). Lehner et al. (2006) used temperature, precipitation and water use future projections as input to the WaterGAP (Water – Global Assessment and Prognosis; Alcamo et al. 2003, Döll et al. 2003) hydrological model to analyzes changes in the magnitude and frequencies of floods and droughts in Europe at 0.5° spatial resolution. The second approach focuses on establishing statistical relationships between the hydrological drivers and discharge through the analysis of historical data, and use the projections of the most relevant covariates to project changes in discharge (e.g., Neri et al., 2019a). For instance, Villarini et al. (2015) focused on one watershed in the central United States and considered historical observations of precipitation and agricultural intensity to estimate the parameters of a statistical model; they then used the projections of these predictors to obtain the projected changes in discharge magnitude. Regardless of the selected approach, the overarching methodology involves: 1) establishing relationships (either statistical in nature or based on hydrologic models) between discharge and its drivers; 2) using projected changes in the identified drivers to obtain the projected changes in the discharge-related quantities of interest (Seneviratne et al., 2012; Villarini and Slater, 2017).

The projection studies based on hydrologic models (see also Li et al. 2015; Lu et al. 2018; Zheng et al. 2018) are of great importance for the assessment of the impacts that climate change and land use / land cover (LULC) can have on flood characteristics and they give a global perspective of their most relevant macro-scale patterns. However, because the outputs of these hydrological models are distributed over grid cells and not necessarily related to the specific gauge station (Giuntoli et al., 2015), their applicability loses strength and reliability for the analyses at the catchment scale (Gudmundsson et al., 2012), and it represents a limitation when specific and localized flood mitigation plans are needed. Focusing on the U.S. Midwest, there are different studies analyzing the future characteristics of floods at specific catchments, but they either consider a single catchment (Ahiablame et al., 2017; Choi et al., 2017; Jha et al., 2004; Sunde et al., 2017) or they analyze flood magnitude (Byun et al., 2019; Chien et al., 2013; Kollat et al., 2012; Schlef et al., 2018; Wobus et al., 2017). A regional study about the projected frequency of flood events at the catchment scale is still lacking: this is a particularly important research topic given the increasing trends detected in the frequency of flood events across the U.S. Midwest (e.g., Mallakpour and Villarini, 2015; Slater and Villarini 2016; Neri et al. 2019b). Moreover, given the increasing availability and length of the observed time series of streamflow and climate variables, data-driven statistical attribution (see, e.g., Neri et al., 2019b; Slater and Villarini 2017) represents a more straightforward approach to analyze projected changes in the characteristics of flood events at the catchment scale compared to the physically-based hydrological models, because these models are faster to implement, less time-consuming and less affected by model parameters uncertainties (Duethmann et al., 2015).

Here we adopt a statistical framework similar to the one described in Neri et al. (2019a, 2019b) by using Poisson regression to attribute changes in the frequency of flood events (i.e., the predictand) to changes in precipitation, temperature and antecedent wetness conditions (i.e., the predictors) at the seasonal scale; we then use centennial precipitation and temperature projections from two different ensembles of downscaled and bias-corrected GCMs as input to these statistical models to investigate the projected changes in the frequency of flood events. The research questions we want to answer are: how is the frequency of flood events projected to change across the U.S. Midwest during the 21st century? Are the changes uniform over the region and across seasons, or are there “hotspots” that exhibit a stronger signal of change? Are the changes sensitive to the threshold values used to identify the events?

2. Data

We focus on 286 USGS gauging stations located across the U.S. Midwest (the area includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota and Wisconsin) (Figure 1). The streamflow time series of each station has at least 50 consecutive complete years (we consider a year complete if it has more than 330 daily observations) of data, and is not affected by any regulation (i.e., not classified with code “5” or “6” according to the USGS notation). We create the time series of the frequency of flood events using a peak-over-threshold (POT) approach, counting the number of events with a discharge value greater than a selected threshold during each season (winter: December-February; spring: March-May; summer: June-August; fall: September-November) of every year. The flood threshold value is site-specific and selected to give 1, 2, 3, or 4 events per year on average (see also Neri et al. (2019b); Mallakpour and Villarini (2015)). For instance, if at a given site we focus on two events per year on average over the 1940-2016 period (i.e., 77 years), we set a threshold so that we select the top 144 events, making sure that each event is separate by 5 days plus the logarithm of the drainage area (in square miles) (Lang et al., 1999); this threshold varies from site to site, and decreases as we move from 1 to 4 events per year on average. The time series obtained using these thresholds represent the predictand for our statistical models. It is worth mentioning that the daily values are smaller than the instantaneous peak values, especially at small basins; however, given that we work with events exceeding a threshold, as long as the ratio between daily averages and instantaneous peaks is constant for a given basin, the selected flood events would be the same.

Observed precipitation and temperature records are derived from the PRISM dataset (PRISM, 2017), which provides monthly values across the entire United States at a resolution of ~4km. For each gauging station, we compute the basin-averaged value of both variables for each month, and then we aggregate it at the seasonal time scale to obtain the basin-averaged seasonal precipitation and temperature time series. We focus on the period starting from 1940 to 2016.

Centennial projections of precipitation and temperature are derived from two different datasets: the North America Coordinated Regional Downscaling Experiment (NA-CORDEX) (Mearns et al., 2017) and the Localized Constructed Analogs (LOCA) (Pierce et al., 2014). NA-CORDEX provides outputs of regional climate models (RCMs) using boundary conditions from GCMs from the Coupled Model Intercomparison Project phase five (CMIP5) (Taylor et al., 2012) archive, covering most of North America at a resolution of 0.22° and monthly. The LOCA dataset provides daily time series of climate variables across North America at a resolution of $1/16^{th}$ of a degree obtained by means of statistically downscaling the CMIP5 GCMs. Here we focus on the historical simulations of precipitation and temperature covering the 1950-2005 period, and the representative concentration pathway (RCP) 8.5 for the projections from 2006 to 2100. We consider ten members of the NA-CORDEX obtained by using five GCMs providing initial and boundary conditions to five RCMs (not all the RCMs are used for each GCM). LOCA has 32 members obtained by downscaling 32 GCMs. Similar to the observations, we use the GCM outputs to compute the basin-averaged time series of seasonal precipitation and temperature. The third predictor, i.e., the antecedent wetness conditions, is defined as the accumulated precipitation during the three months prior to the analyzed season (e.g., Kam and Sheffield, 2016; Neri et al., 2019a, 2019b; L. Slater and Villarini, 2017).

We then estimate the ensemble mean for each basin-averaged driver, with each GCM member having the same weight. To correct for the biases in the ensemble mean of the LOCA and NA-CORDEX, we use the delta-change bias-correction approach (Maraun, 2016) with a modification that adjusts the variability of the historical and projected time series according to the observations. The correction of the mean is simply accomplished by shifting the time series by the difference between the average of the simulated and observed variable over 1950-2005 (i.e., the historical period). The correction of the variance is performed in two steps. First we compute the difference between the shifted time series and a moving average, which allows us to estimate the variability of the time series locally; then we multiply this difference by a factor which is estimated in such a way that the standard deviation of the GCM outputs over the historical period matches the one from the observation (over the same period). Figure 2 shows an example (USGS station 07014500; Meramec River near Sullivan, Missouri) of the type of time series we create for each site and for precipitation and temperature based on observations and bias-corrected GCM outputs.

3. Methodology

Our methodology builds on the approach described by Neri et al. (2019b) and here we provide just a brief overview. Neri et al. (2019b) used Poisson regression to relate the occurrence of flood events to six different predictors: precipitation, antecedent wetness conditions, temperature, population density (as a proxy for urbanization) and agricultural intensity (i.e., combined harvested corn and soybean acreage). They found that precipitation (x_P), wetness conditions (x_M) and temperature (x_T) are the most important drivers across the study region, and this is why we only consider these predictors in this study. We combine these three variables to build four different statistical models relating the parameter of the Poisson distribution to these covariates as described in Table 1. Model P only considers precipitation (x_P) as covariate; model $P.T$ considers precipitation (x_P) and temperature (x_T); model $P.M$ considers precipitation (x_P) and wetness conditions (x_M); model $Mixed$ considers all the three drivers. In this last model, which is not used for the winter season, the value of temperature changes according to the analyzed season: during spring, the temperature is the average temperature for March and April ($X_{T_{Mar-Apr}}$), as a simple way to account for the generation of flood peaks caused by snowmelt and/or rain-on-snow processes; during summer and fall, it considers the average temperature over the summer months, as a proxy for the effects of evapotranspiration during summer and drying soils during fall.

Similar to Neri et al. (2019a), we fit the four models over the observational period from 1940 to 2016 (pending data availability) to each station, season and flood threshold value, and perform the model selection using the Bayesian Information Criterion (Schwarz, 1978). We estimate the α, β, γ and δ parameters (Table 1) for each of the best models over the 1940-2005 period and we evaluate their skill in reproducing the observed time series by computing the correlation coefficient between the observed and simulated flood count time series. We then use the NA-CORDEX/LOCA outputs as inputs to these Poisson regression models to describe the projected changes in flood counts, in a similar way as Neri et al. (2019a) used decadal predictions as input to the models to investigate the future conditions of the frequency of flood events with a lead time up to ten years.

To quantify the temporal changes in the frequency of flood events during the historical and projection period, we use Poisson regression with time as predictor, and focus on the sign and significance (i.e., 5% and 10% level) of the slope coefficient. We show these values only if the time series has at least five years of non-zero flood counts. Furthermore, we compute the difference between the average number of flood events during three consecutive periods of the 21st century (i.e., 2005-2035, 2036-2069, 2070-2099) and the historical period (i.e., 1976-2005) to quantify the magnitude of these changes.

4. Results and Discussion

In this section we focus on the results based on a threshold that gives two event per year on average because the results for the other threshold values are similar (see Supplemental Material).

Figure 3 shows the best models that were selected at every gauging station and the corresponding correlation coefficient between observed and predicted flood counts time series for each season and for a flood threshold value that gives two peaks per year on average. The same results for all the flood threshold values are shown in Figure S1 and Figure S2 of the supplemental material. The P and $P.M$ models are selected at most of the stations, suggesting that precipitation and antecedent wetness conditions are the two most important drivers of the frequency of flood events. The $Mixed$ model is selected only in the northern stations during spring, where temperature and antecedent wetness conditions are crucial for snow-related flood-generating processes. Lastly, the $P.T$ model is the best model only in few stations, with no significant consistency in space or season. These results are consistent with previous similar analyses (Neri et al., 2019a, 2019b; Slater and Villarini, 2016, 2017), which show that the frequency of flood events during spring at stations in the Northern Great Plains is driven by a combination of temperature and antecedent wetness conditions and that precipitation is an important driver, particularly during summer. Because across much of the study area the only drivers responsible for the frequency of flood events are precipitation and antecedent wetness conditions, it is clear that changes in this flood hazard during the 21st century are mostly driven by projected changes in precipitation rather than temperature. The skill of the Poisson regression statistical models in reproducing the observed flood counts is overall good, with an average correlation coefficient among the different seasons of 0.56 (Figure 3) (consult Neri et al. (2019b) for a more detailed evaluation of the model performance).

The trends of the seasonal frequency of flood events during the historical period for the flood threshold value that gives two peaks per year on average and according to the observations and to the median of the Poisson regression model when using observations, LOCA and NA-CORDEX datasets are shown in Figure 4 (Figure S3 of the supplemental material shows the same results also for the other threshold values). In general, the Poisson regression models using the observed precipitation and temperature as predictors are able to well reproduce the trends in the observed number of flood events (compare the first and second columns of Figure 4). These trends in the frequency of flood events are consistent with those obtained in Mallakpour and Villarini (2015) and Neri et al. (2019b) with respect to a comparable historical period. Moreover, these findings further support what mentioned in the introduction, i.e., that it is the frequency of flood events, rather than its magnitude (Villarini et al., 2011; Mallakpour and Villarini, 2015), which presents significant trends. The statistical models forced with the NA-CORDEX well reproduce the observed positive and negative trends during all seasons except for winter, where many trends in central Indiana, northern Illinois and southern Michigan are different. The LOCA dataset also performs comparatively well, even though it presents some trends which are discordant with the observations. The two datasets behave similarly with respect to the spring season, where most of the gauging stations present positive trends in agreement with the observations. These findings are similar also for the trends obtained using a flood threshold value that gives 3 and 4 peaks per year on average (Figure S3 of the supplemental material). The acceptable skill of the statistical models in reproducing the observed frequency of flood events, when forced with climate observations and the ensemble of the historical runs by the GCMs, enables us to use the same models to project future changes in the frequency of flood events up to the end of the 21st century.

Figure 5 shows the trends in the frequency of flood events over the 2006-2100 period based on the LOCA and NA-CORDEX dataset. These results suggest that flood events are projected to become more frequent during the 21st century across much of the U.S. Midwest during winter and spring. The fall season presents, instead, spatially consistent negative trends. With respect to the summer season, no reliable conclusions can be drawn because the two datasets provide discordant results. To quantify the magnitude of these changes, Figure 6 shows the difference between the average flood counts during three different future periods (i.e., 2006-2035, 2036-2069 and 2070-2099) compared to the last 30 years of the historical period (1976-2005) according to the LOCA and NA-CORDEX ensemble. At a very general level, we project a considerable increase in the frequency of flood events during winter and spring, with larger changes as we move towards the end of the 21st century. In particular, the largest increases in the frequency of flood events occur in the stations located in the northern Great Plains during spring, suggesting that projected precipitation during the wintertime and temperature play an important role in driving the future changes of the frequency of flood events in the context of snowmelt and potential changes in the seasonality of precipitation. The winter season is subject to a considerable increase in flood events, especially at stations located in the south-eastern part of the domain that experience flooding associated with atmospheric rivers and extratropical cyclones (e.g., Lavers and Villarini, 2013; Nakamura et al., 2013; Nayak and Villarini, 2017). With regards to summer and fall, the results obtained using LOCA and NA-CORDEX (Figure 6) suggest that there is not a strong change (at least compared to spring) in terms of flood counts during the 21st century. These last findings appear to be in contrast with many studies showing that precipitation is projected to slightly decrease during summer and fall in the U.S. Midwest (i.e. Byun and Hamlet, 2018; Swain and Hayhoe, 2015), which should lead to a reduction in the number of flood events. One way to reconcile these discrepancies is by considering that the projected decrease in precipitation is small during these seasons (see also Winkler et al. (2012)); therefore, at the stations where the model with precipitation as the only predictor is selected, there are minor or no changes in the frequency of flood events, leading to a muted response. Moreover, some of the positive trends can be due to the possible increase in precipitation towards the end of the 21st century which leads to an increase in the frequency of these events, because those are the years that exert a significant leverage in terms of detected trends.

It is worth pointing out that these results are based on the assumption that the regression coefficients of the drivers of the best models estimated on the 1940-2005 period are the same also for the 2006-2100 projection period. To gain insights with respect to the validity of this assumption, we use a splitting-sample validation approach: we calibrate the statistical models on the 1940-1977 period and then we estimate the median of the Poisson distribution on the 1978-2016 period (i.e., the validation period). Figure S7 of the supplemental material shows the correlation coefficient between the observed and modeled flood counts over the validation period. The models present good skill in reproducing the interannual variability of flood counts, suggesting that the parameters of the best drivers obtained during the calibration period are also representative of the rainfall-runoff processes of the following period. Even though this test provides encouraging results regarding the reliability and performance of our statistical models over the observational period, it does not ensure the same robustness with respect to the projection period, because we do not know how the future hydrological system is going to behave. The uncertainties associated with the above-mentioned assumption represent therefore a limitation of our approach, which is however "an attribute of information and therefore does not mean lack of knowledge" (Blöschl and Montanari, 2010). We took these uncertainties into consideration in our results given that we developed probabilistic models rather than deterministic outputs.

5. Conclusions

In this study we used a statistical approach to investigate the projected changes in the seasonal frequency of flood events during the 21st century at 286 USGS station across the U.S. Midwest. The results are based on downscaled and bias corrected GCM outputs and the RCP 8.5. The selection of the flood events is carried out through a peak-over-threshold approach and considering four different flood threshold values. Compared to previous studies, here we provide a regional perspective of the projected changes in the frequency of flood events. The main findings of this study can be summarized as follows:

- The trends over the historical period (1950-2005) based on the NA-CORDEX reproduce reasonably well those from the observations, especially during spring. The LOCA dataset also performs well, with the exception of the summer season where most of the trends have opposite signs with respect to the observations.
- Our findings suggest that the spring season is projected to experience a substantial increase in the frequency of flood events during the 21st century across much of the study

region, and in particular across the Northern Great Plains. The average number of flood events is also projected to increase in the winter, especially in the south-eastern part of the domain which is within the storm track of the extratropical cyclones. Despite summer and fall present statistically significant trends, the change of the average number of flood counts is negligible for most of the gauging stations.

- It is worth reminding that these results are based on the extrapolation of the modeling results for the historical period to the future; this means that we assumed that the relationship between the response variable and the predictor(s) is expected to remain constant. Moreover, we also assumed that the performance of the GCMs for the historical period is a reflection of their performance in the future.
- This framework provides a simple and rapid methodology to assess projected changes in flood events, which can be further updated and improved with new and higher resolution GCMs (e.g., Haarsma et al., 2016).

Acknowledgements:

This study was supported in part by the Broad Agency Announcement (BAA) Program and the Engineer Research and Development Center (ERDC)–Cold Regions Research and Engineering Laboratory (CRREL) under Contract No. W913E5-16-C-0002, USACE Institute for Water Resources, and by the National Science Foundation under CAREER Grant AGS-1349827. The comments and suggestions by two anonymous reviewers are gratefully acknowledged.

Table 1 – List of the four statistical models used to relate the seasonal occurrence of flood events to the three drivers: precipitation (x_P), wetness conditions (x_M) and temperature (x_T).

Model Name	Dependence
$\boldsymbol{P}$	$log\lambda_1 = \alpha_1 + \beta_1 x_P$
$\boldsymbol{P.T}$	$log\lambda_2 = \alpha_2 + \beta_2 x_P + \gamma_2 x_T$
$\boldsymbol{P.M}$	$log\lambda_3 = \alpha_3 + \beta_3 x_P + \gamma_3 x_M$
$\boldsymbol{Mixed}$	$log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Mar-Apr}}$ $log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Summer}}$

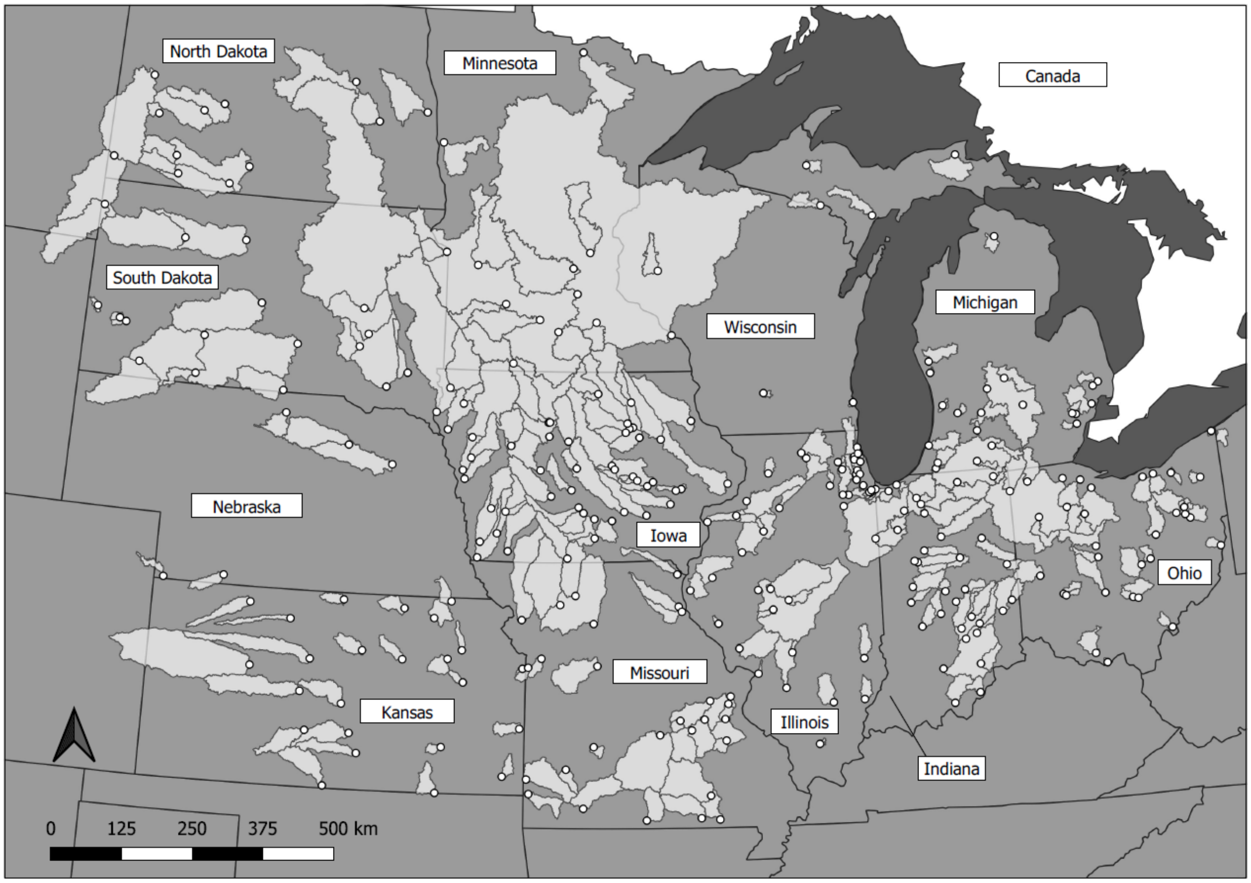


Figure 1 - Location of the 286 USGS gauging stations (white circles) and the relative upstream drainage area (light gray polygons).

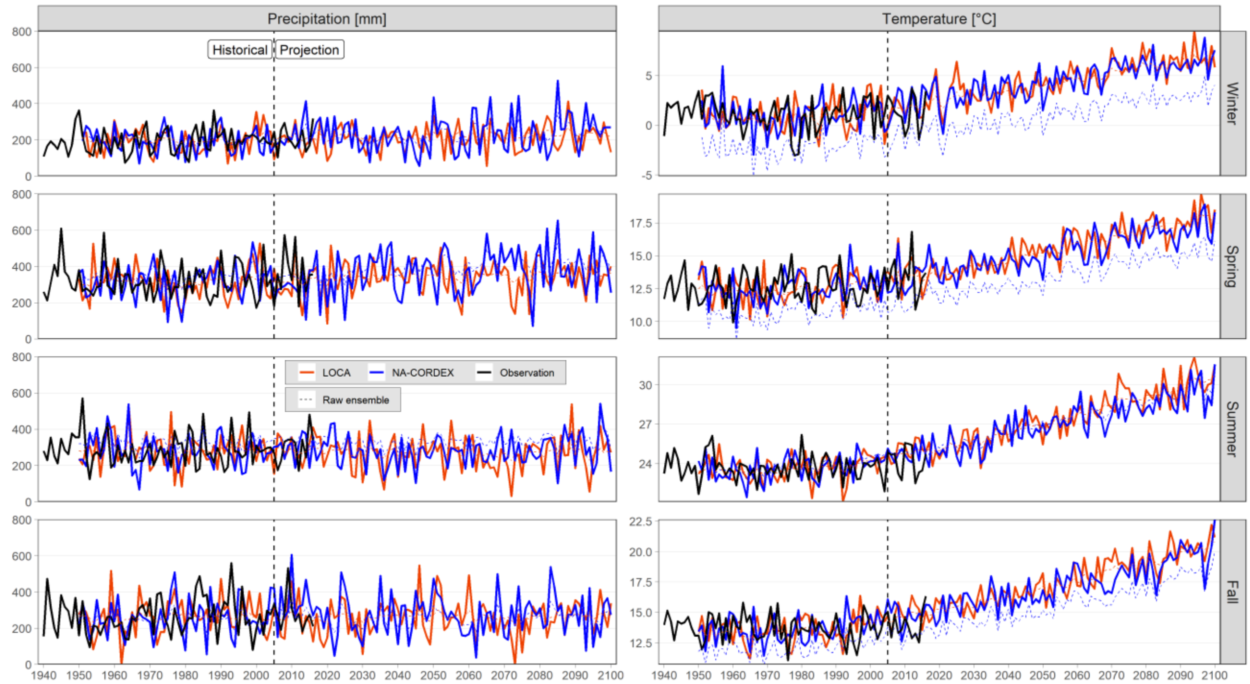


Figure 2 – Observed, historical and projected basin-averaged bias-corrected precipitation (left panels) and temperature (right panels) seasonal time series for USGS station 07014500 (Meramec River near Sullivan, Missouri). The black line represents the observed values, while the red and blue solid (dotted) lines the values based on bias corrected (raw) LOCA and NA-CORDEX, respectively.

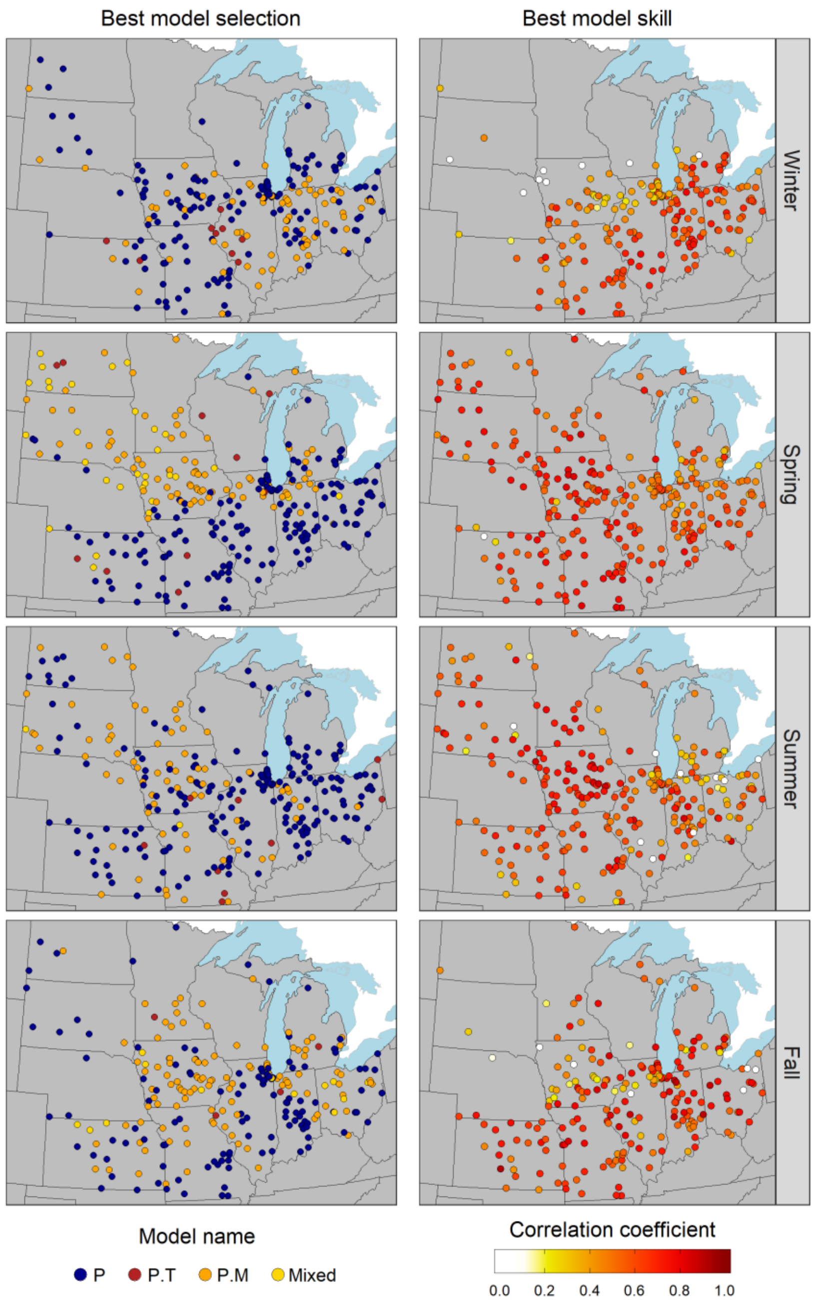


Figure 3 – Map showing the selected best models (four panels on the left) and their skill (four panels on the right) for each season (rows) and for a flood threshold value that gives two peaks per year on average. The blue, brown, orange and yellow circles on the left refer to the P, $P.T$, $P.M$, and $Mixed$ models, respectively. In some stations, no model is selected largely because the observed time series does not have at least five years with a flood count value different from zero. Note that not every best model provides a predicted time series with at least five years with a flood count value different from zero, therefore no correlation coefficient can be computed.

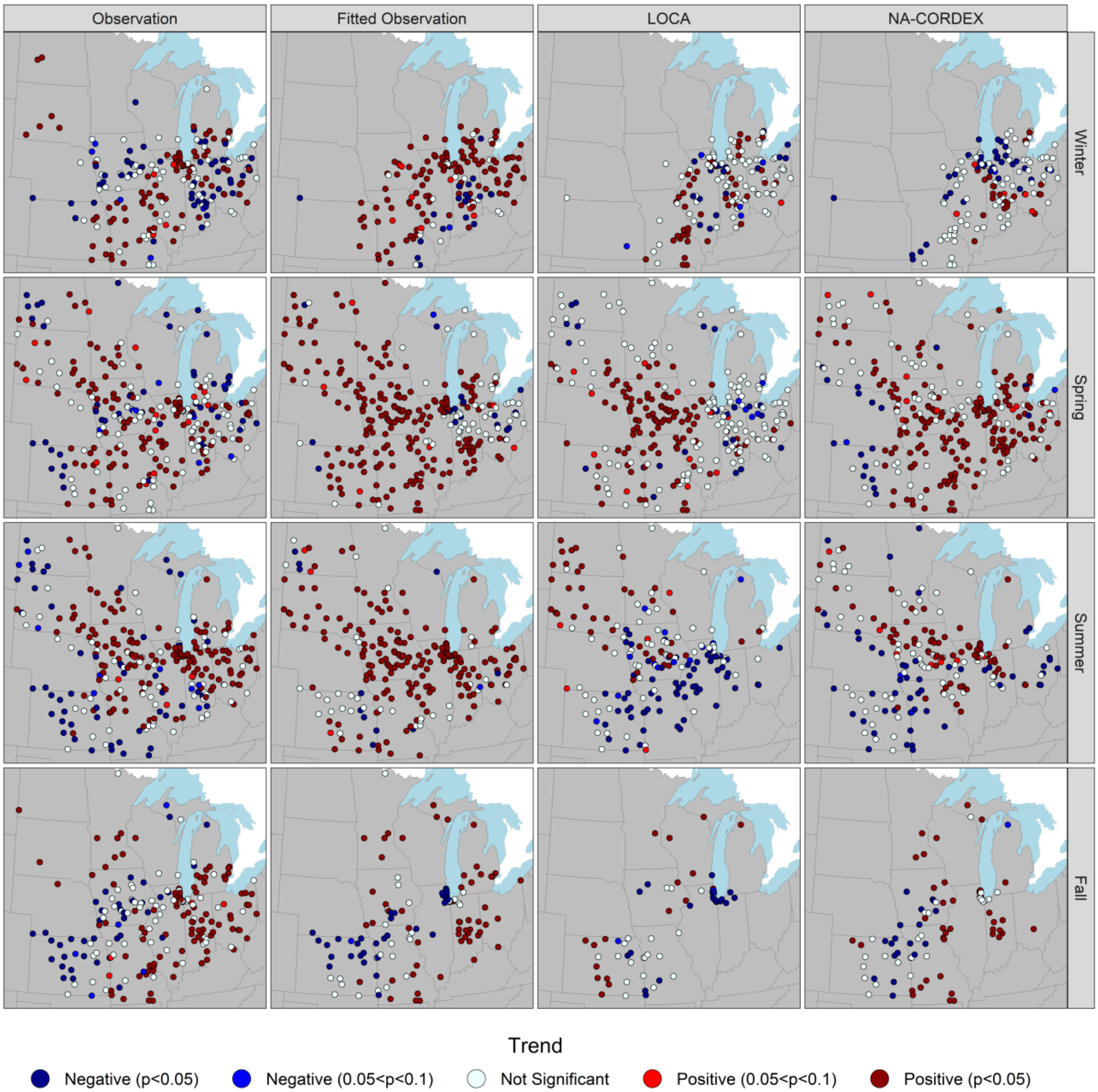


Figure 4 – Map showing the trends of the frequency of flood events during the historical period (i.e., 1950-2005) for the observations and for the median of the Poisson regression models using observations, LOCA and NA-CORDEX. The results refer to the threshold values that gives two peaks per year on average among the observational period 1940-2016 (see Figure S3 of the supplemental material for the other flood threshold values). The four columns represent the four considered datasets, and the rows the four seasons. In each panel, the dark red and dark blue (red and blue) circles indicate the positive and negative trends significant at the 5% (10%) significance level, respectively. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero.

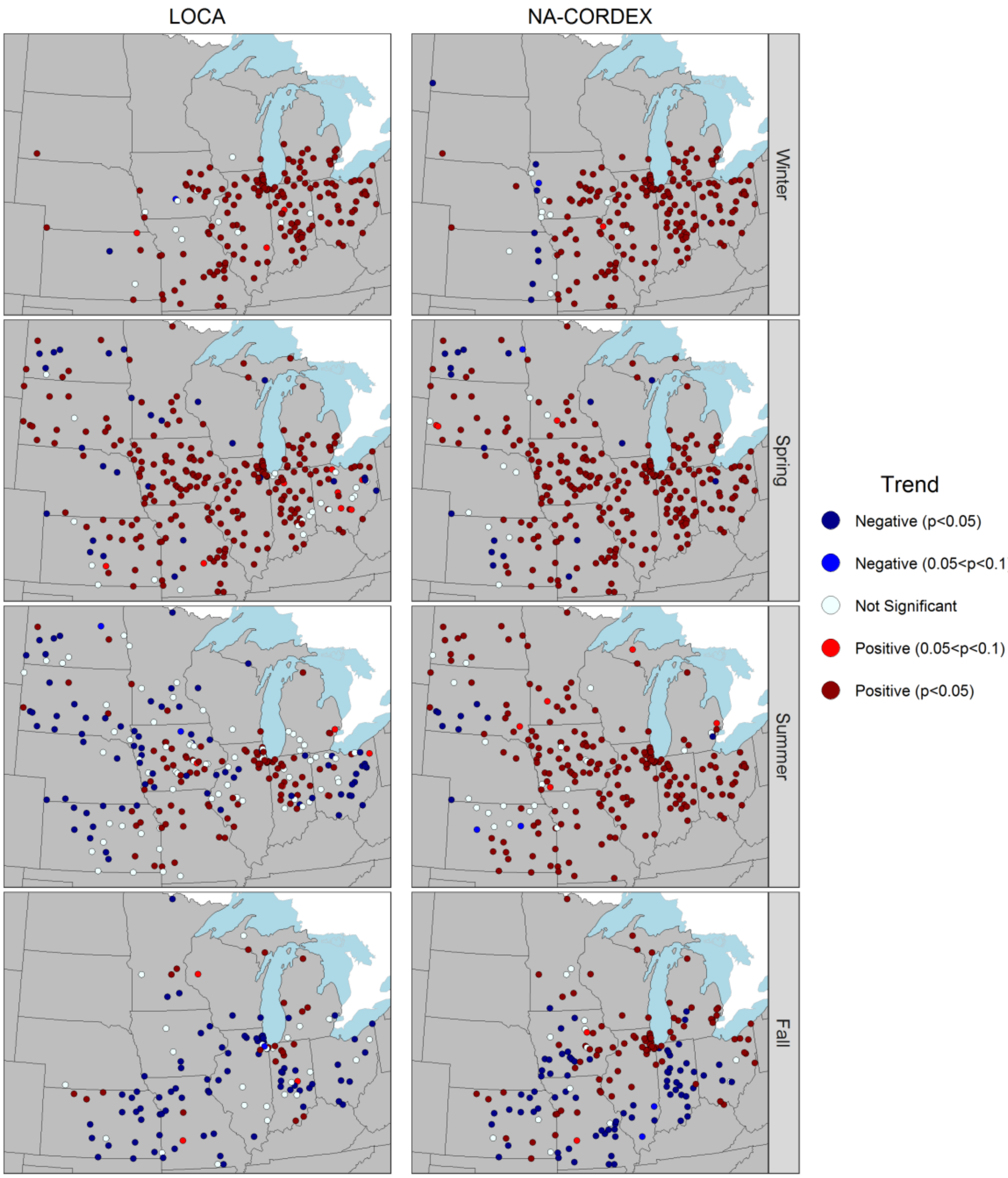


Figure 5 – Map showing the projected trends (2006-2100) in the frequency of flood events based on the LOCA (left column) and NA-CORDEX (right column) for the four seasons and for a flood threshold value that gives two peaks per year on average. The symbol notation is the same as in Figure 4. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero. The same results relative to the other flood thresholds are presented in Figure S4 of the supplemental material.

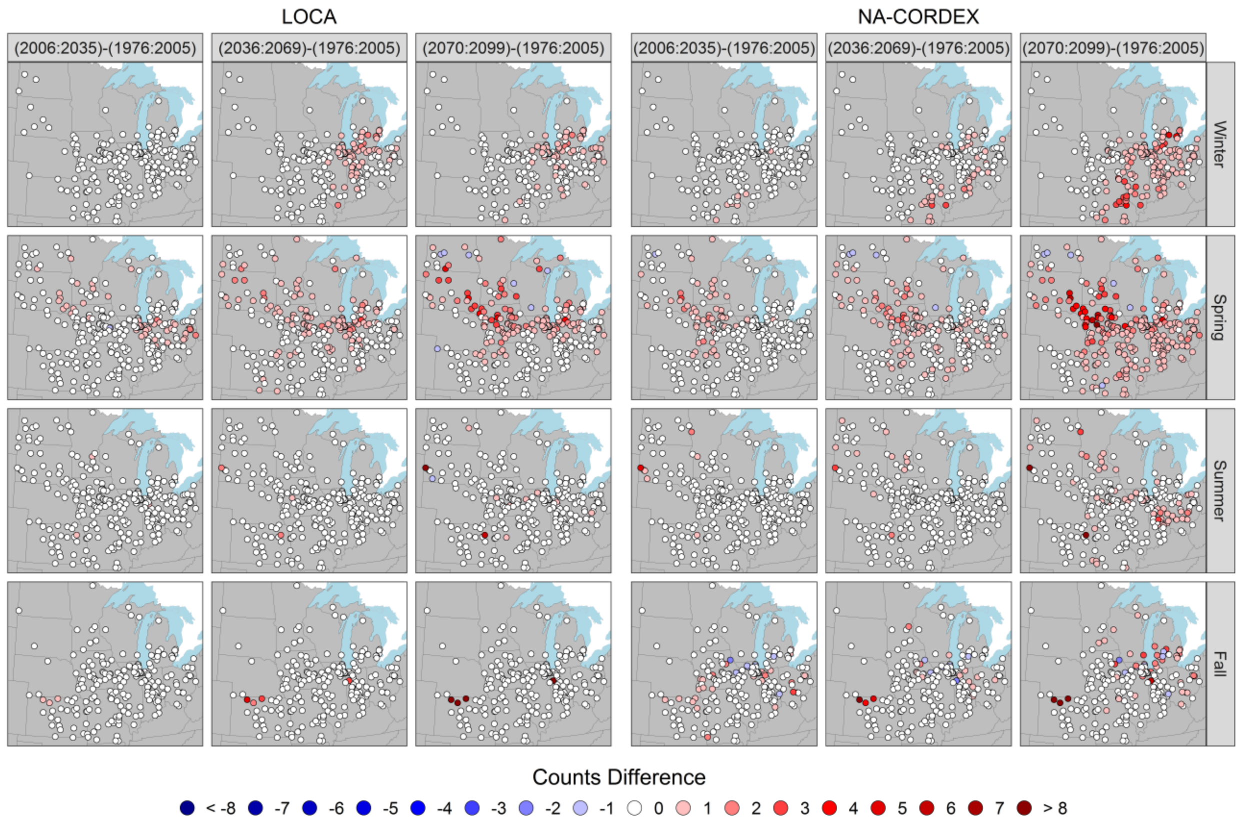


Figure 6 – Map showing the difference between the average value of flood counts during three spans (i.e., 2006:2035, 2036:2069, 2070:2099) of the projection period and the average value of flood counts during the last 30 years of the historical period (i.e., 1976:2005) using the LOCA (left set of panels) and NA-CORDEX (right set of panels) dataset. The results refer to the threshold values that give, two peaks per year on average among the observational period 1940-2016. For each set of panels, the three columns represent the three considered spans, and the rows the four seasons. For the other flood threshold values see Figure S5 and Figure S6 of the supplemental material for the LOCA and NA-CORDEX datasets, respectively.

Bibliography

Ahiablame, L., Sinha, T., Paul, M., Ji, J.H., Rajib, A., 2017. Streamflow response to potential land use and climate changes in the James River watershed, Upper Midwest United States. J. Hydrol. Reg. Stud. 14, 150–166. https://doi.org/10.1016/j.ejrh.2017.11.004

AIDR, 2017. Flood Information to Support Land-use Planning. Australian Institute for Disaster Resilience.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrol. Sci. J. 48, 317–338. https://doi.org/10.1623/hysj.48.3.317.45290

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., Feyen, L., 2017. Global projections of river flood risk in a warmer world. Earth's Futur. 5, 171–182. https://doi.org/10.1002/2016EF000485

Archfield, S.A., Hirsch, R.M., Viglione, A., Blöschl, G., 2016. Fragmented patterns of flood change across the United States. Geophys. Res. Lett. 43, 10,232–10,239. https://doi.org/doi:10.1002/ 2016GL070590

Arnell, N.W., Gosling, S.N., 2016. The impacts of climate change on river flood risk at the global scale. Clim. Change 134, 387–401. https://doi.org/10.1007/s10584-014-1084-5

Brekke, L.D., Kiang, J.E., Olsen, J.R., Pulwarty, R.S., Raff, D.A., Turnipseed, D.P., Webb, R.S., White, K.D., 2009. Climate Change and Water Resources Management - A Federal Perspective: U.S. Geological Survey Circular 1331. U.S. Geological Survey, Reston, Virginia.

Byun, K., Chiu, C.M., Hamlet, A.F., 2019. Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US. Sci. Total Environ. 650, 1261–1277. https://doi.org/10.1016/j.scitotenv.2018.09.063

Byun, K., Hamlet, A.F., 2018. Projected changes in future climate over the Midwest and Great Lakes region using downscaled CMIP5 ensembles. Int. J. Climatol. 38, e531–e553. https://doi.org/10.1002/joc.5388

Chang, W., Chen, X., 2018. Monthly rainfall-runoff modeling at watershed scale: A comparative study of data-driven and theory-driven approaches. Water (Switzerland) 10, 1–21. https://doi.org/10.3390/w10091116

Chien, H., Yeh, P.J.F., Knouft, J.H., 2013. Modeling the potential impacts of climate change on streamflow in agricultural watersheds of the Midwestern United States. J. Hydrol. 491, 73–88. https://doi.org/10.1016/j.jhydrol.2013.03.026

Choi, W., Pan, F., Wu, C., 2017. Impacts of climate change and urban growth on the streamflow of the Milwaukee River (Wisconsin, USA). Reg. Environ. Chang. 17, 889–899. https://doi.org/10.1007/s10113-016-1083-3

Dankers, R., Arnell, N.W., Clark, D.B., Falloon, P.D., Fekete, B.M., Gosling, S.N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble. Proc. Natl. Acad. Sci. 111, 3257–3261. https://doi.org/10.1073/pnas.1302078110

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: Model tuning and validation. J. Hydrol. 270, 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Duethmann, D., Bolch, T., Farinotti, D., Kriegel, D., Vorogushyn, S., Merz, B., Pieczonka, T., Jiang, T., Su, B., Güntner, A., 2015. Attribution of streamflow trends in snow- and glacier melt dominated catchments of the Tarim River, Central Asia. Water Resour. Res. 51, 4727–4750. https://doi.org/10.1002/2014WR016716.Received

Eldho, T.I., Kulkarni, A.T., 2017. Conceptual and Physically Based Hydrological Modeling. Sustain. Water Resour. Manag. 81–118. https://doi.org/10.1061/9780784414767.ch04

François, B., Schlef, K.E., Wi, S., Brown, C.M., 2019. Design considerations for riverine floods in a changing climate – A review. J. Hydrol. 574, 557–573. https://doi.org/10.1016/j.jhydrol.2019.04.068

Giuntoli, I., Villarini, G., Prudhomme, C., Hannah, D.M., 2018. Uncertainties in projected runoff over the conterminous United States. Clim. Change 150, 149–162. https://doi.org/10.1007/s10584-018-2280-5

Giuntoli, I., Villarini, G., Prudhomme, C., Mallakpour, I., Hannah, D.M., 2015. Evaluation of global impact models' ability to reproduce runoff characteristics over the central United States. J. Geophys. Res. Atmos. 120, 9183–9159. https://doi.org/doi:10.1002/2015JD023401

Gosling, S.N., Arnell, N.W., 2011. Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis. Hydrol. Process. 25, 1129–1145. https://doi.org/10.1002/hyp.7727

Gudmundsson, L., Wagener, T., Tallaksen, L.M., Engeland, K., 2012. Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. Water Resour. Res. 48, 1–20. https://doi.org/10.1029/2011WR010911

Haarsma, R.J., Roberts, M.J., Vidale, P.L., Catherine, A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N.S., Guemas, V., Von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L.R., Lu, J., Luo, J.J., Mao, J., Mizielinski, M.S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., Von Storch, J.S., 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geosci. Model Dev. 9, 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Hannah, D.M., Demuth, S., van Lanen, H.A.J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., Tallaksen, L.M., 2010. Large-scale river flow archives: importance, current status and future needs. Hydrol. Process. https://doi.org/10.1002/hyp.7794

Jha, M., Pan, Z., Tackle, E.S., Gu, R., 2004. Impacts of climate change on streamflow in the Upper Mississippi River Basin: A regional climate model perspective. J. Geophys. Res. D Atmos. 109, 1–12. https://doi.org/10.1029/2003JD003686

Kam, J., Sheffield, J., 2016. Changes in the low flow regime over the eastern United States (1962–2011): variability, trends, and attributions. Clim. Change 135, 639–653. https://doi.org/10.1007/s10584-015-1574-0

Kollat, J.B., Kasprzyk, J.R., Thomas, W.O., Miller, A.C., Divoky, D., 2012. Estimating the Impacts of Climate Change and Population Growth on Flood Discharges in the United States. J. Water Resour. Plan. Manag. 138, 442–452. https://doi.org/10.1061/(asce)wr.1943-5452.0000233

Lavell, A., Oppenheimer, M., 2012. Climate Change : New Dimensions in Disaster Risk, Exposure, Vulnerability, and Resilience. Manag. Risks Extrem. Events Disasters to Adv. Clim. Chang. Adapt. 25–64.

Lavers, A.D., Villarini, G., 2013. Atmospheric Rivers and Flooding over the Central United States. J. Clim. 7829–7836. https://doi.org/10.1175/JCLI-D-13-00212.1

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis. Clim. Change 75, 273–299. https://doi.org/10.1007/s10584-006-6338-4

Li, L., Diallo, I., Xu, C.Y., Stordal, F., 2015. Hydrological projections under climate change in the near future by RegCM4 in Southern Africa using a large-scale hydrological model. J. Hydrol. 528, 1–16. https://doi.org/10.1016/j.jhydrol.2015.05.028

Lu, W., Wang, W., Shao, Q., Yu, Z., Hao, Z., Xing, W., Yong, B., Li, J., 2018. Hydrological projections of future climate change over the source region of Yellow River and Yangtze River in the Tibetan Plateau: A comprehensive assessment by coupling RegCM4 and VIC model. Hydrol. Process. 32, 2096–2117. https://doi.org/10.1002/hyp.13145

Mallakpour, I., Villarini, G., 2015. The changing nature of flooding across the central United States. Nat. Clim. Chang. https://doi.org/10.1038/nclimate2516

Maraun, D., 2016. Bias Correcting Climate Change Simulations - a Critical Review. Curr. Clim. Chang. Reports 2, 211–220. https://doi.org/10.1007/s40641-016-0050-x

Mearns, L.O., McGinnis, S., Korytina, D., 2017. The NA-CORDEX dataset.

Nakamura, J., Lall, U., Kushnir, Y., Robertson, A.W., Seager, R., 2013. Dynamical Structure of Extreme Floods in the U.S. Midwest and the United Kingdom. J. Hydrometeorol. 14, 485–504. https://doi.org/10.1175/jhm-d-12-059.1

Natural Resources Canada, 2017. Federal Floodplain Mapping Framework. Government of Canada.

Nayak, M.A., Villarini, G., 2017. A long-term perspective of the hydroclimatological impacts of atmospheric rivers over the central United States. Water Resour. Res. 53, 1144–1166.

https://doi.org/doi:10.1002/ 2016WR019033

Neri, A., Villarini, G., Kaustubh, Salvi, A., Slater, L.J., Napolitano, F., 2019a. On the decadal predictability of the frequency of flood events across the U.S. Midwest. Int. J. Climatol. 3, 1796–1804. https://doi.org/10.1002/joc.5915

Neri, A., Villarini, G., Slater, L.J., Napolitano, F., 2019b. On the statistical attribution of the frequency of flood events across the U.S. Midwest. Adv. Water Resour.

Pierce, D.W., Cayan, D.R., Thrasher, B.L., 2014. Statistical Downscaling Using Localized Constructed Analogs (LOCA)*. J. Hydrometeorol. 15, 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1

PRISM, 2017. Prism Climate Group [WWW Document].

Schlef, K.E., François, B., Robertson, A.W., Brown, C., 2018. A General Methodology for Climate-Informed Approaches to Long-Term Flood Projection—Illustrated With the Ohio River Basin. Water Resour. Res. 54, 9321–9341. https://doi.org/10.1029/2018WR023209

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464. https://doi.org/10.1214/09-AOS712

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., M. Reichstein, A. Sorteberg, C. Vera, and X.Z., 2012. Changes in Climate Extremese and their Impacts on the Natural Physical Environment, Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge University Press, Cambridge, UK and New York, NY, USA. https://doi.org/10.1017/CBO9781139177245.006

Slater, L.J., Villarini, G., 2017. Evaluating the Drivers of Seasonal Streamflow in the U.S. Midwest. Water 9, 695. https://doi.org/10.3390/w9090695

Slater, L.J., Villarini, G., 2016. Recent trends in U.S. flood risk. Geophys. Res. Lett. 43, 12,428-12,436. https://doi.org/10.1002/2016GL071199

Sunde, M.G., He, H.S., Hubbart, J.A., Urban, M.A., 2017. Integrating downscaled CMIP5 data with a physically based hydrologic model to estimate potential climate change impacts on streamflow processes in a mixed-use watershed. Hydrol. Process. 31, 1790–1803. https://doi.org/10.1002/hyp.11150

Swain, S., Hayhoe, K., 2015. CMIP5 projected changes in spring and summer drought and wet conditions over North America. Clim. Dyn. 44, 2737–2750. https://doi.org/10.1007/s00382-014-2255-9

Taylor, K.E., Stouffer, R.J., Meehl, G. a, 2007. A Summary of the CMIP5 Experiment Design. World 4, 1–33. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

van der Knijff, J.M., Younis, J., De Roo, A. P. J., 2010. LISFLOOD: a GIS--based distributed model for river basin scale water balance and flood simulation. Int. J. Geogr. Inf. Sci. https://doi.org/10.1080/13658810802549154

Villarini, G., Scoccimarro, E., White, K.D., Arnold, J.R., Schilling, K.E., Ghosh, J., 2015. Projected Changes in Discharge in an Agricultural Watershed in Iowa. J. Am. Water Resour. Assoc. 51, 1361–1371. https://doi.org/10.1111/1752-1688.12318

Villarini, G., Slater, L.J., 2017. Climatology of flooding in the United States, in: Cutter, S.L. (Ed.), Oxford Research Encyclopedia of Natural Hazard Science. Oxford University Press (OUP), Oxford, pp. 0–36.

Villarini, G., Smith, J.A., Baeck, M.L., Krajewski, W.F., 2011. Examining Flood Frequency Distributions in the Midwest U.S.1. JAWRA J. Am. Water Resour. Assoc. 47, 447–463. https://doi.org/10.1111/j.1752-1688.2011.00540.x

Winkler, J. a, Arritt, R.W., Pryor, S.C., 2012. Climate Projections for the Midwest: Availability, Interpretation and Synthesis. US Natl. Clim. Assess. Midwest Tech. Input Rep. 24. https://doi.org/9781610915113

Wobus, C., Gutmann, E., Jones, R., Rissing, M., Mizukami, N., Lorie, M., Mahoney, H., Wood, A.W., Mills, D., Martinich, J., 2017. Climate change impacts on flood risk and asset damages within mapped 100-year floodplains of the contiguous United States. Nat. Hazards Earth Syst. Sci. 17, 2199–2211. https://doi.org/10.5194/nhess-17-2199-2017

Zheng, H., Chiew, F.H.S., Charles, S., Podger, G., 2018. Future climate and runoff projections across South Asia from CMIP5 global climate models and hydrological modelling. J. Hydrol. Reg. Stud. 18, 92–109. https://doi.org/10.1016/j.ejrh.2018.06.004

Supplemental Figures
Click here to download Supplementary material for on-line publication only: Supplemental_Material_6.pdf

*Revised manuscript with changes marked
Click here to view linked References

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

Andrea Neri[1], Gabriele Villarini[2], and Francesco Napolitano[1]

[1] Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Rome, Italy.

[2] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, Iowa, USA.

Manuscript submitted to

Journal of Hydrology

18 August 2019

Corresponding author address:

Andrea Neri, Department of Civil, Constructional and Environmental Engineering, University of Rome "La Sapienza", Via Eudossiana 18, 00184, Rome, Italy. E-mail: andrea.neri@uniroma1.it. Tel.: +390644585063

ABSTRACT

There is growing empirical evidence that many river basins across the U.S. Midwest have been experiencing an increase in the frequency of flood events over the most recent decades. Albeit these detected changes are important to understand what happened in our recent past, they cannot be directly extrapolated to obtain information about possible future changes in the frequency of flood events. Building on recent statistically-based attribution studies, we project seasonal changes in the frequency of flood events at 286 U.S. Geological Survey gauging stations across the U.S. Midwest using projections of precipitation, antecedent wetness conditions and temperature as drivers. The projections of the covariates are obtained from two datasets obtained by downscaling global circulation models from the Fifth Coupled Model Intercomparison Project (CMIP5). We focus on the representative concentration pathway (RCP) 8.5 and on four different flood thresholds (i.e., from more common to less frequent flood events). We find that the frequency of flood events during the 21st century increases during spring at most of the analyzed gauging stations, with larger changes in the Northern Great Plains and regardless of the flood threshold value. Our findings also point to a projected increasing number of flood events during the winter, especially in the stations in the southern and western part of the domain (Iowa, Missouri, Illinois, Ohio, Indiana and Michigan). A marked change in the frequency of flood events is not projected for the summer and fall.

Keywords: frequency of flood events; projections; statistical modeling; CORDEX; LOCA; CMIP5

1 Introduction

The hydrologic impacts of climate change have been the topic of a growing body of research, and have attracted significant interest from decision makers and stakeholders in terms of their projected changes and related societal and economic impacts. This awareness, together with the evidence of other natural disasters attributed to climate change (e.g., Seneviratne et al., 2012), has recently spread across many countries and pushed governments to react in terms of both adaptation to and management of extreme flood events (e.g., Lavell and Oppenheimer, 2012). For instance, the Government of Canada developed the Federal Floodplain Mapping Framework, which is a document intended to describe the entire process to define reliable flood risk maps and the effect that climate alterations have on them (Natural Resources Canada, 2017); the Australian Disaster Resilience Handbook 7 (AIDR, 2017) describes the main practices and activities for a proper floodplain management, focusing on the land use and development and on how climate change affects flood modeling; in the United States, four federal agencies (U.S. Geological Survey (USGS), U.S. Army Corps of Engineers, Bureau of Reclamation and National Oceanic and Atmospheric Administration) prepared a report (Brekke et al., 2009) which proposes better practices and activities for water resources management by considering the effects of global warming.

Although there is still low confidence about the changes of the frequency and/or magnitude of flood events at the global scale (Seneviratne et al., 2012), these recent actions and strategies for flood risk management have been encouraged by several studies that detect statistically significant trends in flooding at the regional level. For instance, focusing on the continental United States, Mallakpour and Villarini (2015) analyzed daily streamflow records at 774 USGS stream gauge stations across central United States covering the period from 1962 to

2011, and detected statistically significant increases in the frequency of flood events for the majority of the stations (see also Neri et al. (2019b)). Slater and Villarini (2016) showed that the frequency of the water level exceeding the National Weather Service's four flood level categories in 2042 water basins across the United States was subject to significant changes, with different parts of the country exhibiting spatially coherent signals of increasing or decreasing trends. Archfield et al. (2016) analyzed the frequency, duration, magnitude and volume of floods at 345 streamgages across the United States, showing that certain regions present significant changes in these flood properties. For a recent review of the detected changes in flooding across the continental United States, see Villarini and Slater (2017).

Even though these findings represent a key step towards our improved characterization of the changes in flood events over the past several decades, they do not provide useful information about projected changes in flood-related quantities. To address this knowledge gap, different methods have been proposed and developed, which can be classified into two broad classes: hydrological and statistical models (e.g., Chang and Chen, 2018; Eldho and Kulkarni, 2017; François et al., 2019; Giuntoli et al., 2018; Villarini et al., 2015). The former aims to predict future flood conditions by using projections of the hydrologic forcings as input to physical or conceptual equations that describe the main processes regulating the transformation of precipitation into runoff, as for instance through the use of global impact models (e.g., Dankers et al., 2014). For instance, Arnell and Gosling (2016) analyzed the effects of climate change on global flood risk by combining projections of population and of climate variables of 21 global climate models (GCMs) to force the Mac-PDM.09 hydrological model (Gosling and Arnell, 2011) to obtain flood frequency curves at 0.5° resolutions. Alfieri et al. (2017) assessed global projections of the frequency and magnitude of river floods using dynamical downscaled projections of climate variables by seven GCMs as input to the LISFLOOD hydrological model (van der Knijff et al., 2010). Lehner et al. (2006) used temperature, precipitation and water use future projections as input to the WaterGAP (Water – Global Assessment and Prognosis; Alcamo et al. 2003, Döll et al. 2003) hydrological model to analyzes changes in the magnitude and frequencies of floods and droughts in Europe at 0.5° spatial resolution. The second approach focuses on establishing statistical relationships between the hydrological drivers and discharge through the analysis of historical data, and use the projections of the most relevant covariates to project changes in discharge (e.g., Neri et al., 2019a). For instance, Villarini et al. (2015) focused on one watershed in the central United States and considered historical observations of precipitation and agricultural intensity to estimate the parameters of a statistical model; they then used the projections of these predictors to obtain the projected changes in discharge magnitude. Regardless of the selected approach, the overarching methodology involves: 1) establishing relationships (either statistical in nature or based on hydrologic models) between discharge and its drivers; 2) using projected changes in the identified drivers to obtain the projected changes in the discharge-related quantities of interest (Seneviratne et al., 2012; Villarini and Slater, 2017).

The projection studies based on hydrologic models (see also Li et al. 2015; Lu et al. 2018; Zheng et al. 2018) are of great importance for the assessment of the impacts that climate change and land use / land cover (LULC) can have on flood characteristics and they give a global perspective of their most relevant macro-scale patterns. However, ~~being~~ because the outputs of these hydrological models are distributed over grid cells and not necessarily related to the specific gauge station (Giuntoli et al., 2015), their applicability loses strength and reliability for the analyses at the catchment scale (Gudmundsson et al., 2012), and it represents a limitation when specific and localized flood mitigation plans are needed. Focusing on the U.S. Midwest, there are different studies analyzing the future characteristics of floods at specific catchments, but they either consider a single catchment (Ahiablame et al., 2017; Choi et al., 2017; Jha et al., 2004; ~~Schlef et al., 2018;~~ Sunde et al., 2017) or they analyze flood magnitude (Byun et al., 2019; Chien et al., 2013; Kollat et al., 2012; Schlef et al., 2018; Wobus et al., 2017). A regional study about the projected frequency of flood events at the catchment scale is still lacking: this is a particularly important research topic given the increasing trends detected in the frequency of flood events across the U.S. Midwest (e.g., Mallakpour and Villarini, 2015; Slater and Villarini 2016; Neri et al. 2019b). Moreover, given the increasing availability and length of the observed time series of streamflow and climate variables, data-driven statistical attribution (see, e.g., Neri et al., 2019b; Slater and Villarini 2017) represents a more straightforward approach to analyze projected changes in the characteristics of flood events at the catchment scale compared to the physically-based hydrological models, because these models are faster to implement, less time-consuming and less affected by model parameters uncertainties (Duethmann et al., 2015).

Here we adopt a statistical framework similar to the one described in Neri et al. (2019a, 2019b) by using Poisson regression to attribute changes in the frequency of flood events (i.e., the predictand) to changes in precipitation, temperature and antecedent wetness conditions (i.e., the predictors) at the seasonal scale; we then use centennial precipitation and temperature projections from two different ensembles of downscaled and bias-corrected GCMs as input to these statistical models to investigate the projected changes in the frequency of flood events. The research questions we want to answer are: how is the frequency of flood events projected to change across the U.S. Midwest during the 21st century? Are the changes uniform over the region and across seasons, or are there “hotspots” that exhibit a stronger signal of change? Are the changes sensitive to the threshold values used to identify the events?

2. Data

We focus on 286 USGS gauging stations located across the U.S. Midwest (the area includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota and Wisconsin) (Figure 1). The streamflow time series of each station has at least 50 consecutive complete years (we consider a year complete if it has more than 330 daily observations) of data, and is not affected by any regulation (i.e., not classified with code "5" or "6" according to the USGS notation). We create the time series of the frequency of flood events using a peak-over-threshold (POT) approach, counting the number of events with a discharge value greater than a selected threshold during each season (winter: December-February; spring: March-May; summer: June-August; fall: September-November) of every year. The flood threshold value is site-specific and selected to give 1, 2, 3, or 4 events per year on average (see also Neri et al. (2019b); Mallakpour and Villarini (2015)). For instance, if at a given site we focus on two events per year on average over the 1940-2016 period (i.e., 77 years), we set a threshold so that we select the top 144 events, making sure that each event is separate by 5 days plus the logarithm of the drainage area (in square miles) (Lang et al., 1999); this threshold varies from site to site, and decreases as we move from 1 to 4 events per year on average. The time series obtained using these thresholds represent the predictand for our statistical models. It is worth mentioning that the daily values are smaller than the instantaneous peak values, especially at small basins; however, given that we work with events exceeding a threshold, as long as the ratio between daily averages and instantaneous peaks is constant for a given basin, the selected flood events would be the same.

Observed precipitation and temperature records are derived from the PRISM dataset (PRISM, 2017), which provides monthly values across the entire United States at a resolution of ~4km. For each gauging station, we compute the basin-averaged value of both variables for each month, and then we aggregate it at the seasonal time scale to obtain the basin-averaged seasonal precipitation and temperature time series. We focus on the period starting from 1940 to 2016.

Centennial projections of precipitation and temperature are derived from two different datasets: the North America Coordinated Regional Downscaling Experiment (NA-CORDEX) (Mearns et al., 2017) and the Localized Constructed Analogs (LOCA) (Pierce et al., 2014). NA-CORDEX provides outputs of regional climate models (RCMs) using boundary conditions from GCMs from the Coupled Model Intercomparison Project phase five (CMIP5) (Taylor et al., 2012) archive, covering most of North America at a resolution of 0.22° and monthly. The LOCA dataset provides daily time series of climate variables across North America at a resolution of 1/16th of a degree obtained by means of statistically downscaling the CMIP5 GCMs. Here we focus on the historical simulations of precipitation and temperature covering the 1950-2005 period, and the representative concentration pathway (RCP) 8.5 for the projections from 2006 to 2100. We consider ten members of the NA-CORDEX obtained by using five GCMs providing initial and boundary conditions to five RCMs (not all the RCMs are used for each GCM). LOCA has 32 members obtained by downscaling 32 GCMs. Similar to the observations, we use the GCM outputs to compute the basin-averaged time series of seasonal precipitation and temperature. The third predictor, i.e., the antecedent wetness conditions, is defined as the accumulated precipitation during the three months prior to the analyzed season (e.g., Kam and Sheffield, 2016; Neri et al., 2019a, 2019b; L. Slater and Villarini, 2017).

-We then estimate the ensemble mean for each basin-averaged driver, with each GCM member having the same weight. To correct for the biases in the ensemble mean of the LOCA and NA-CORDEX, we use the delta-change bias-correction approach (Maraun, 2016) with a modification that adjusts the variability of the historical and projected time series according to the observations. The correction of the mean is simply accomplished by shifting the time series by the difference between the average of the simulated and observed variable over 1950-2005 (i.e., the historical period). The correction of the variance is performed in two steps. First we compute the difference between the shifted time series and a moving average, which allows us to estimate the variability of the time series locally; then we multiply this difference by a factor which is estimated in such a way that the standard deviation of the GCM outputs over the historical period matches the one from the observation (over the same period). Figure 2 shows an example (USGS station 07014500; Meramec River near Sullivan, Missouri) of the type of time series we create for each site and for precipitation and temperature based on observations and bias-corrected GCM outputs.

3. Methodology

Our methodology builds on the approach described by Neri et al. (~~2019a,~~ 2019b) and here we provide just a brief overview. Neri et al. (~~2019a,~~ 2019b) used Poisson regression to relate the occurrence of flood events to six different predictors: precipitation, antecedent wetness conditions, temperature, population density (as a proxy for urbanization) and agricultural intensity (i.e., combined harvested corn and soybean acreage). They found that precipitation (x_P), wetness conditions (x_M) and temperature (x_T) are the most important drivers across the study region, and this is why we only consider these predictors in this study. We combine these three variables to build four different statistical models relating the parameter of the Poisson distribution to these covariates as described in Table 1. Model P only considers precipitation (x_P) as covariate; model $P.T$ considers precipitation (x_P) and temperature (x_T); model $P.M$ considers precipitation (x_P) and wetness conditions (x_M); model $Mixed$ considers all the three drivers. In this last model, which is not used for the winter season, the value of temperature changes according to the analyzed season: during spring, the temperature is the average temperature for March and April ($X_{T_{Mar-Apr}}$), as a simple way to account for the generation of flood peaks caused by snowmelt and/or rain-on-snow processes; during summer and fall, it considers the average temperature over the summer months, as a proxy for the effects of evapotranspiration during summer and drying soils during fall.

Similar to Neri et al. (2019a), we fit the four models over the observational period from 1940 to 2016 (pending data availability) to each station, season and flood threshold value, and perform the model selection using the Bayesian Information Criterion (Schwarz, 1978). We estimate the α, β, γ and δ parameters (Table 1) for each of the best models over the 1940-2005 period and we evaluate their skill in reproducing the observed time series by computing the correlation coefficient between the observed and simulated flood count time series. We then use the NA-CORDEX/LOCA outputs as inputs to these Poisson regression models to describe the projected changes in flood counts, in a similar way as Neri et al. (2019a) used decadal predictions as input to the models to investigate the future conditions of the frequency of flood events with a lead time up to ten years.

To quantify the temporal changes in the frequency of flood events during the historical and projection period, we use Poisson regression with time as predictor, and focus on the sign and significance (i.e., 5% and 10% level) of the slope coefficient. We show these values only if

Formatted: Italian (Italy)

the time series has at least five years of non-zero flood counts. Furthermore, we compute the difference between the average number of flood events during three consecutive periods of the 21st century (i.e., 2005-2035, 2036-2069, 2070-2099) and the historical period (i.e., 1976-2005) to quantify the magnitude of these changes.

4. Results and Discussion

In this section we focus on the results based on a threshold that gives two event per year on average because the results for the other threshold values are similar (see Supplemental Material).

Figure 3 shows the best models that were selected at every gauging station and the corresponding correlation coefficient between observed and predicted flood counts time series for each season and for a flood threshold value that gives two peaks per year on average. The same results for all the flood threshold values are shown in Figure S1 and Figure S2 of the supplemental material. The P and $P.M$ models are selected at most of the stations, suggesting that precipitation and antecedent wetness conditions are the two most important drivers of the frequency of flood events. The $Mixed$ model is selected only in the northern stations during spring, where temperature and antecedent wetness conditions are crucial for snow-related flood-generating processes. Lastly, the $P.T$ model is the best model only in few stations, with no significant consistency in space or season. These results are consistent with previous similar analyses (Neri et al., 2019a, 2019b; Slater and Villarini, 2016, 2017), which show that the frequency of flood events during spring at stations in the Northern Great Plains is driven by a combination of temperature and antecedent wetness conditions and that precipitation is an important driver, particularly during summer. Because across much of the study area the only drivers responsible for the frequency of flood events are precipitation and antecedent wetness conditions, it is clear that changes in this flood hazard during the 21st century are mostly driven by projected changes in precipitation rather than temperature. The skill of the Poisson regression statistical models in reproducing the observed flood counts is overall good, with an average correlation coefficient among the different seasons of 0.56 (Figure 3) (consult Neri et al. (2019b) for a more detailed evaluation of the model performance).

The trends of the seasonal frequency of flood events during the historical period for the flood threshold value that gives two peaks per year on average and according to the observations and to the median of the Poisson regression model when using observations, LOCA and NA-CORDEX datasets are shown in Figure 4 (Figure S3 of the supplemental material shows the same results also for the other threshold values). In general, the Poisson regression models using the observed precipitation and temperature as predictors are able to well reproduce the trends in the observed number of flood events (compare the first and second columns of Figure 4). These trends in the frequency of flood events are consistent with those obtained in Mallakpour and Villarini (2015) and Neri et al. (2019b) with respect to a comparable historical period. Moreover, these findings further support what mentioned in the introduction, i.e., that it is the frequency of flood events, rather than its magnitude (Villarini et al., 2011; Mallakpour and Villarini, 2015), which presents significant trends. The statistical models forced with the NA-CORDEX well reproduce the observed positive and negative trends during all seasons except for winter, where many trends in central Indiana, northern Illinois and southern Michigan are different. The LOCA dataset also performs comparatively well, even though it presents some trends which are discordant with the observations. The two datasets behave similarly with respect to the spring season, where most of the gauging stations present positive trends in agreement with the observations. These findings are similar also for the trends obtained using a flood threshold value that gives 3 and 4 peaks per year on average (Figure S3 of the supplemental material). The acceptable skill of the statistical models in reproducing the observed frequency of flood events, when forced with climate observations and the ensemble of the historical runs by the GCMs, enables us to use the same models to project future changes in the frequency of flood events up to the end of the 21st century.

Formatted: Superscript

Figure 5 shows the trends in the frequency of flood events over the 2006-2100 period based on the LOCA and NA-CORDEX dataset. These results suggest that flood events are projected to become more frequent during the 21st century across much of the U.S. Midwest during winter and spring. The fall season presents, instead, spatially consistent negative trends. With respect to the summer season, no reliable conclusions can be drawn because the two datasets provide discordant results. To quantify the magnitude of these changes, Figure 6 shows the difference between the average flood counts during three different future periods (i.e., 2006-2035, 2036-2069 and 2070-2099) compared to the last 30 years of the historical period (1976-2005) according to the LOCA and NA-CORDEX ensemble. At a very general level, we project a considerable increase in the frequency of flood events during winter and spring, with larger changes as we move towards the end of the 21st century. In particular, the largest increases in the frequency of flood events occur in the stations located in the northern Great Plains during spring, suggesting that projected precipitation during the wintertime and temperature play an important role in driving the future changes of the frequency of flood events in the context of snowmelt and potential changes in the seasonality of precipitation. The winter season is subject to a considerable increase in flood events, especially at stations located in the south-eastern part of the domain that experience flooding associated with atmospheric rivers and extratropical cyclones (e.g., Lavers and Villarini, 2013; Nakamura et al., 2013; Nayak and Villarini, 2017). With regards to summer and fall, the results obtained using LOCA and NA-CORDEX (Figure 6) suggest that there is not a strong change (at least compared to spring) in terms of flood counts during the 21st century. These last findings appear to be in contrast with many studies showing that precipitation is projected to slightly decrease during summer and fall in the U.S. Midwest (i.e. Byun and Hamlet, 2018; Swain and Hayhoe, 2015), which should lead to a reduction in the number of flood events. One way to reconcile these discrepancies is by considering that the projected decrease in precipitation is small during these seasons (see also Winkler et al. (2012)); therefore, at the stations where the model with precipitation as the only predictor is selected, there are minor or no changes in the frequency of flood events, leading to a muted response. Moreover, some of the positive trends can be due to the possible increase in precipitation towards the end of the 21st century which leads to an increase in the frequency of these events, because those are the years that exert a significant leverage in terms of detected trends.~~; with that said, we still expect years with more events alternating to more quiet ones~~.

It is worth pointing out that these results are based on the assumption that the regression coefficients of the drivers of the best models estimated on the 1940-2005 period are the same also for the 2006-2100 projection period. To gain insights with respect to the validity of this assumption, we use a splitting-sample validation approach: we calibrate the statistical models on the 1940-1977 period and then we estimate the median of the Poisson distribution on the 1978-2016 period (i.e., the validation period). Figure S7 of the supplemental material shows the correlation coefficient between the observed and modeled flood counts over the validation period. The models present good skill in reproducing the interannual variability of flood counts, suggesting that the parameters of the best drivers obtained during the calibration period are also representative of the rainfall-runoff processes of the following period. Even though this test provides encouraging results regarding the reliability and performance of our statistical models over the observational period, it does not ensure the same robustness with respect to the projection period, because we do not know how the future hydrological system is going to behave. The uncertainties associated with the above-mentioned assumption represent therefore a limitation of our approach, which is however “an attribute of information and therefore does not mean lack of knowledge” (Blöschl and Montanari, 2010). We took these uncertainties into consideration in our results given that we developed probabilistic models rather than deterministic outputs.

5. Conclusions

In this study we used a statistical approach to investigate the projected changes in the seasonal frequency of flood events during the 21st century at 286 USGS station across the U.S. Midwest. The results are based on downscaled and bias corrected GCM outputs and the RCP 8.5. The selection of the flood events is carried out through a peak-over-threshold approach and considering four different flood threshold values. Compared to previous studies, here we provide a regional perspective of the projected changes in the frequency of flood events. The main findings of this study can be summarized as follows:

- The trends over the historical period (1950-2005) based on the NA-CORDEX reproduce reasonably well those from the observations, especially during spring. The LOCA dataset also performs well, with the exception of the summer season where most of the trends have opposite signs with respect to the observations.

- Our findings suggest that the spring season is projected to experience a substantial increase in the frequency of flood events during the 21st century across much of the study region, and in particular across the Northern Great Plains. The average number of flood events is also projected to increase in the winter, especially in the south-eastern part of the domain which is within the storm track of the extratropical cyclones. Despite summer and fall present statistically significant trends, the change of the average number of flood counts is negligible for most of the gauging stations.
- It is worth reminding that these results are based on the extrapolation of the modeling results for the historical period to the future; this means that we assumed that the relationship between the response variable and the predictor(s) is expected to remain constant. Moreover, we also assumed that the performance of the GCMs for the historical period is a reflection of their performance in the future.
- This framework provides a simple and rapid methodology to assess projected changes in flood events, which can be further updated and improved with new and higher resolution GCMs (e.g., Haarsma et al., 2016).

Acknowledgements:

This study was supported in part by the Broad Agency Announcement (BAA) Program and the Engineer Research and Development Center (ERDC)–Cold Regions Research and Engineering Laboratory (CRREL) under Contract No. W913E5-16-C-0002, USACE Institute for Water Resources, and by the National Science Foundation under CAREER Grant AGS-1349827. The comments and suggestions by two anonymous reviewers are gratefully acknowledged.

363

Table 1 – List of the four statistical models used to relate the seasonal occurrence of flood events to the three drivers: precipitation (x_P), wetness conditions (x_M) and temperature (x_T).

Model Name	Dependence
$\boldsymbol{P}$	$log\lambda_1 = \alpha_1 + \beta_1 x_P$
$\boldsymbol{P.T}$	$log\lambda_2 = \alpha_2 + \beta_2 x_P + \gamma_2 x_T$
$\boldsymbol{P.M}$	$log\lambda_3 = \alpha_3 + \beta_3 x_P + \gamma_3 x_M$
$\boldsymbol{Mixed}$	$log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Mar-Apr}}$ $log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Summer}}$

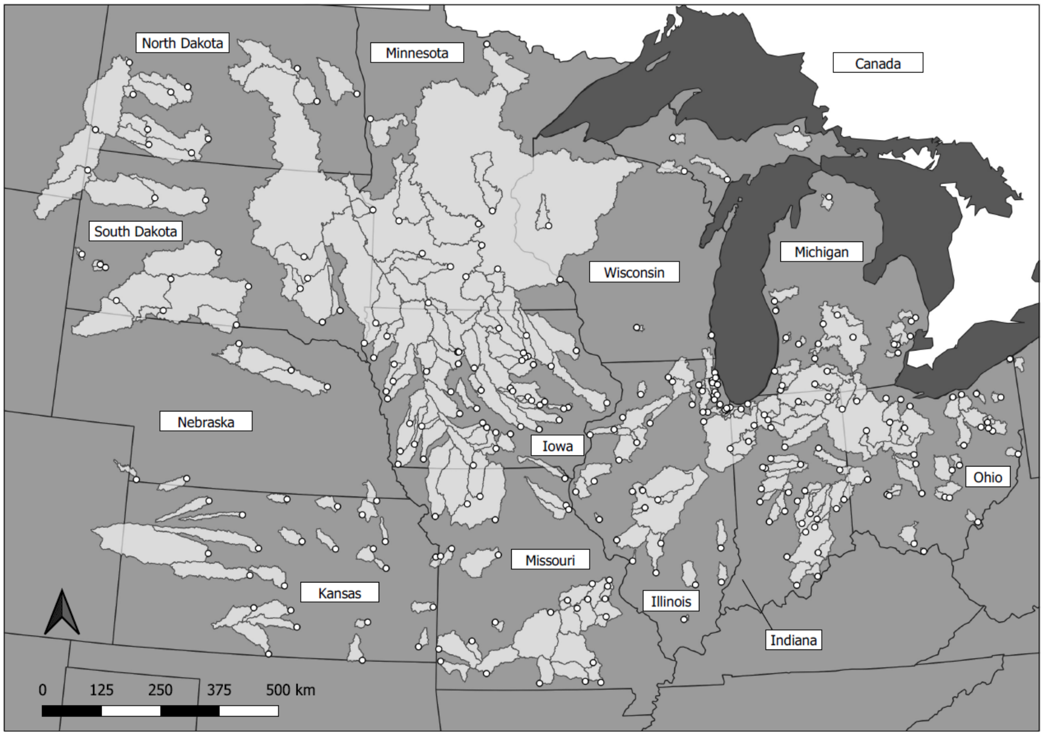


Figure 1 - Location of the 286 USGS gauging stations (white circles) and the relative upstream drainage area (light gray polygons).

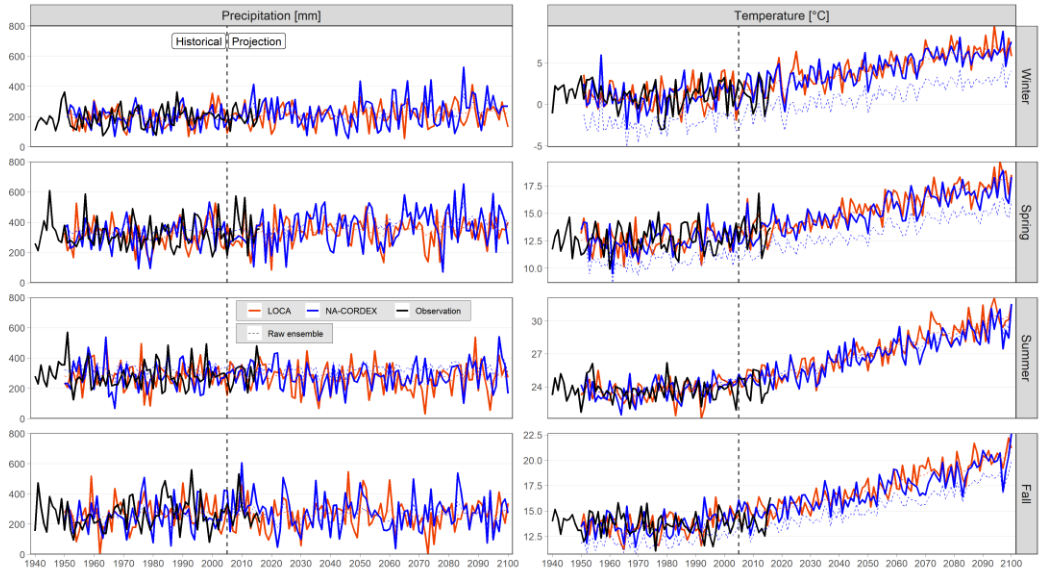


Figure 2 – Observed, historical and projected basin-averaged bias-corrected precipitation (left panels) and temperature (right panels) seasonal time series for USGS station 07014500 (Meramec River near Sullivan, Missouri). The black line represents the observed values, while the red and blue solid (dotted) lines the values based on bias corrected (raw) LOCA and NA-CORDEX, respectively.

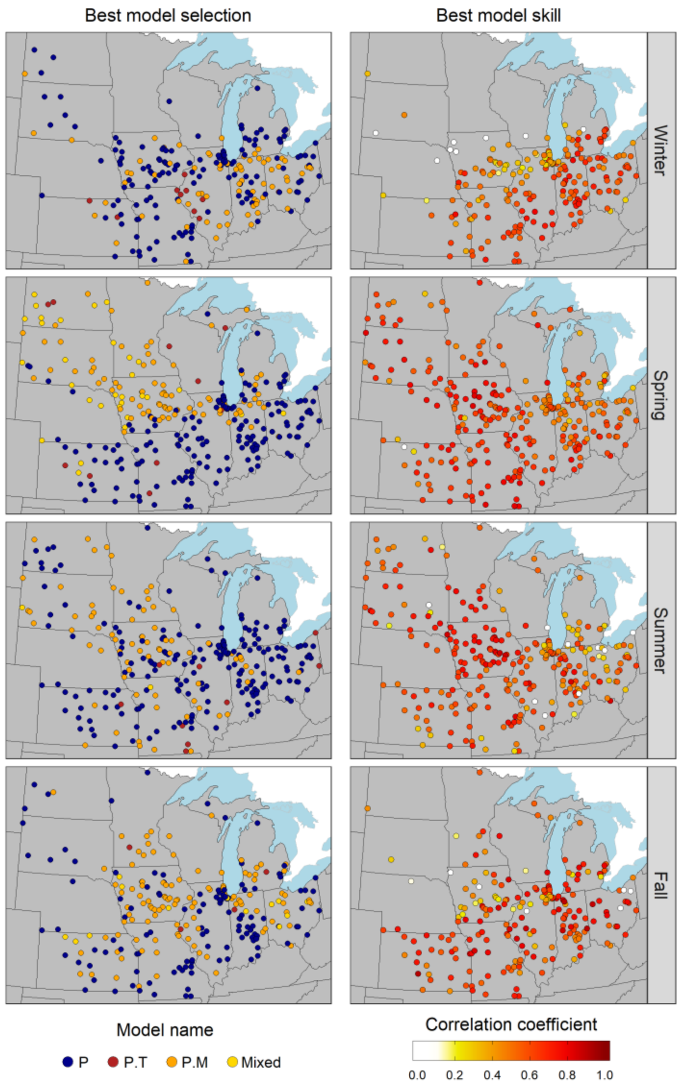
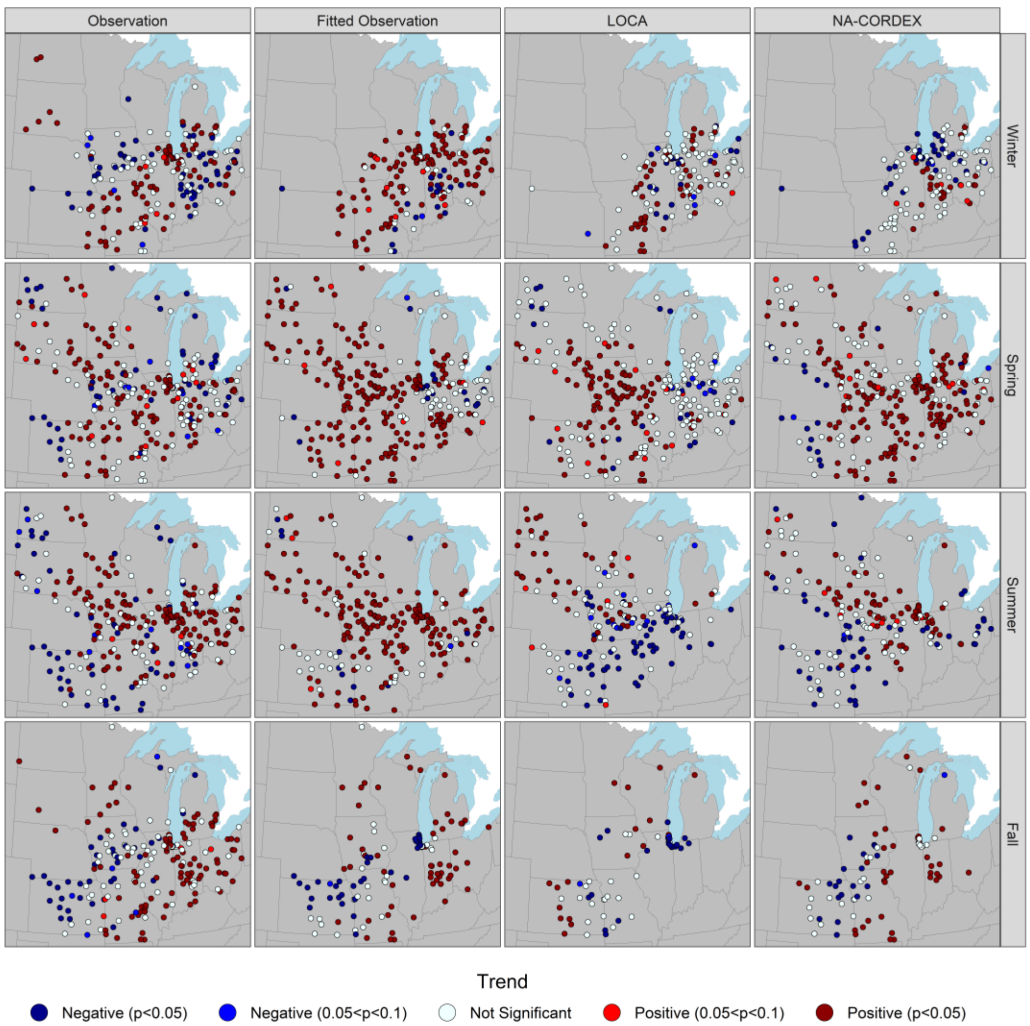


Figure 3 – Map showing the selected best models (four panels on the left) and their skill (four panels on the right) for each season (rows) and for a flood threshold value that gives two peaks per year on average. The blue, brown, orange and yellow circles on the left refer to the P, $P.T$, $P.M$, and $Mixed$ models, respectively. In some stations, no model is selected largely because the observed time series does not have at least five years with a flood count value different from zero. Note that not every best model provides a predicted time series with at least five years with a flood count value different from zero, therefore no correlation coefficient can be computed.

389


Observation
Fitted Observation
LOCA
NA-CORDEX
Winter
Spring
Summer
Fall
Trend
Negative (p<0.05)
Negative (0.05<p<0.1)
Not Significant
Positive (0.05<p<0.1)
Positive (p<0.05)

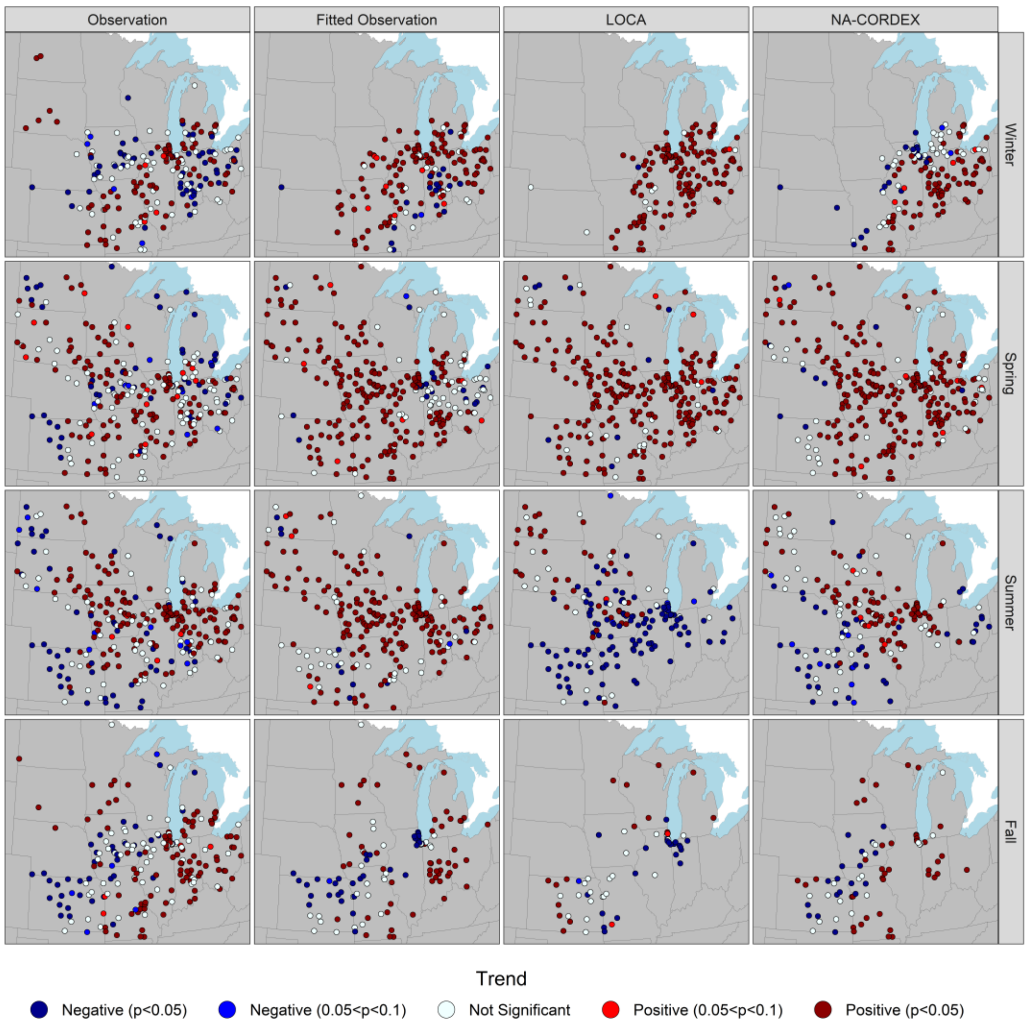


Figure 4 – Map showing the trends of the frequency of flood events during the historical period (i.e., 1950-2005) for the observations and for the median of the Poisson regression models using observations, LOCA and NA-CORDEX. The results refer to the threshold values that gives two peaks per year on average among the observational period 1940-2016 (see Figure S3 of the supplemental material for the other flood threshold values). The four columns represent the four considered datasets, and the rows the four seasons. In each panel, the dark red and dark blue (red and blue) circles indicate the positive and negative trends significant at the 5% (10%) significance level, respectively. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero.

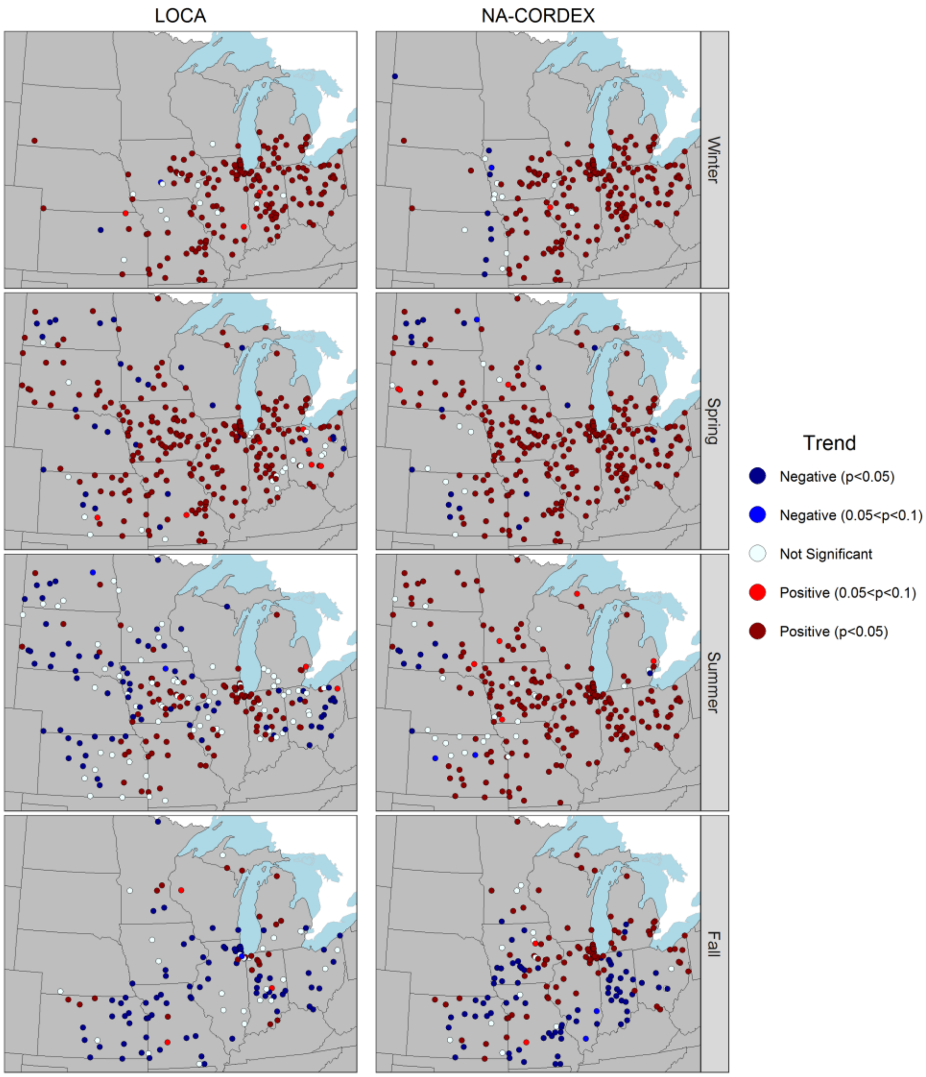


Figure 5 – Map showing the projected trends (2006-2100) in the frequency of flood events based on the LOCA (left column) and NA-CORDEX (right column) for the four seasons and for a flood threshold value that gives two peaks per year on average. The symbol notation is the same as in Figure 4. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero. The same results relative to the other flood thresholds are presented in Figure S4 of the supplemental material.

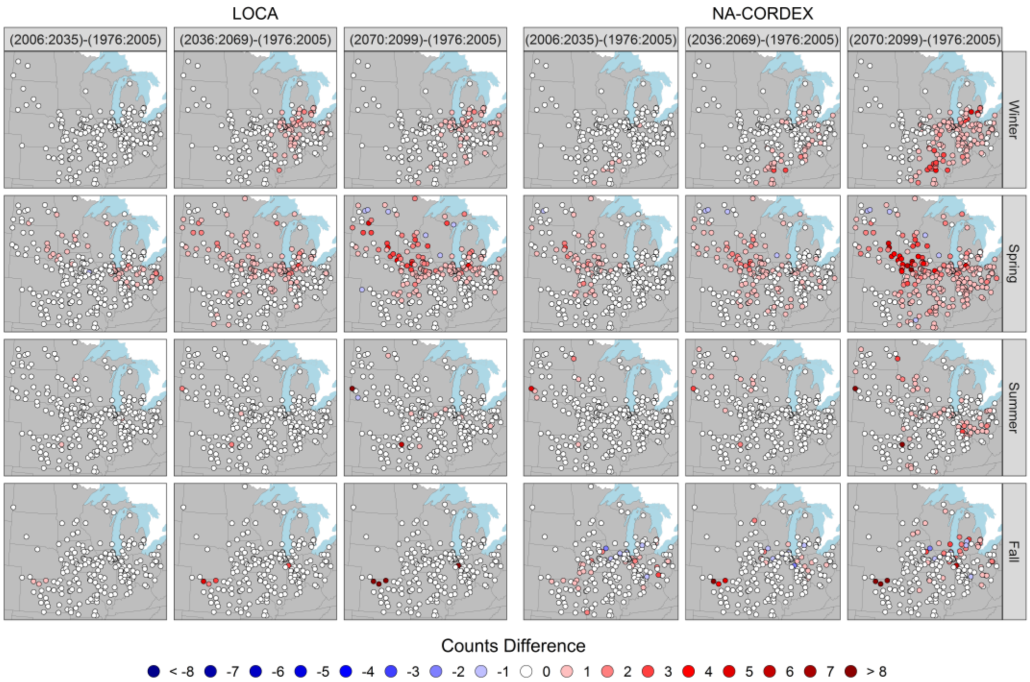


Figure 6 – Map showing the difference between the average value of flood counts during three spans (i.e., 2006:2035, 2036:2069, 2070:2099) of the projection period and the average value of flood counts during the last 30 years of the historical period (i.e., 1976:2005) using the LOCA (left set of panels) and NA-CORDEX (right set of panels) dataset. The results refer to the threshold values that give, two peaks per year on average among the observational period 1940-2016. For each set of panels, the three columns represent the three considered spans, and the rows the four seasons. For the other flood threshold values see Figure S5 and Figure S6 of the supplemental material for the LOCA and NA-CORDEX datasets, respectively.

Bibliography

Ahiablame, L., Sinha, T., Paul, M., Ji, J.H., Rajib, A., 2017. Streamflow response to potential land use and climate changes in the James River watershed, Upper Midwest United States. J. Hydrol. Reg. Stud. 14, 150–166. https://doi.org/10.1016/j.ejrh.2017.11.004

AIDR, 2017. Flood Information to Support Land-use Planning. Australian Institute for Disaster Resilience.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrol. Sci. J. 48, 317–338. https://doi.org/10.1623/hysj.48.3.317.45290

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., Feyen, L., 2017. Global projections of river flood risk in a warmer world. Earth's Futur. 5, 171–182. https://doi.org/10.1002/2016EF000485

Archfield, S.A., Hirsch, R.M., Viglione, A., Blöschl, G., 2016. Fragmented patterns of flood change across the United States. Geophys. Res. Lett. 43, 10,232–10,239. https://doi.org/doi:10.1002/ 2016GL070590

Arnell, N.W., Gosling, S.N., 2016. The impacts of climate change on river flood risk at the global scale. Clim. Change 134, 387–401. https://doi.org/10.1007/s10584-014-1084-5

Brekke, L.D., Kiang, J.E., Olsen, J.R., Pulwarty, R.S., Raff, D.A., Turnipseed, D.P., Webb, R.S., White, K.D., 2009. Climate Change and Water Resources Management - A Federal Perspective: U.S. Geological Survey Circular 1331. U.S. Geological Survey, Reston, Virginia.

Byun, K., Chiu, C.M., Hamlet, A.F., 2019. Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US. Sci. Total Environ. 650, 1261–1277. https://doi.org/10.1016/j.scitotenv.2018.09.063

Byun, K., Hamlet, A.F., 2018. Projected changes in future climate over the Midwest and Great Lakes region using downscaled CMIP5 ensembles. Int. J. Climatol. 38, e531–e553. https://doi.org/10.1002/joc.5388

Chang, W., Chen, X., 2018. Monthly rainfall-runoff modeling at watershed scale: A comparative study of data-driven and theory-driven approaches. Water (Switzerland) 10, 1–21. https://doi.org/10.3390/w10091116

Chien, H., Yeh, P.J.F., Knouft, J.H., 2013. Modeling the potential impacts of climate change on streamflow in agricultural watersheds of the Midwestern United States. J. Hydrol. 491, 73–88. https://doi.org/10.1016/j.jhydrol.2013.03.026

Choi, W., Pan, F., Wu, C., 2017. Impacts of climate change and urban growth on the streamflow of the Milwaukee River (Wisconsin, USA). Reg. Environ. Chang. 17, 889–899. https://doi.org/10.1007/s10113-016-1083-3

Formatted: Italian (Italy)

Dankers, R., Arnell, N.W., Clark, D.B., Falloon, P.D., Fekete, B.M., Gosling, S.N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble. Proc. Natl. Acad. Sci. 111, 3257–3261. https://doi.org/10.1073/pnas.1302078110

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: Model tuning and validation. J. Hydrol. 270, 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Duethmann, D., Bolch, T., Farinotti, D., Kriegel, D., Vorogushyn, S., Merz, B., Pieczonka, T., Jiang, T., Su, B., Güntner, A., 2015. Attribution of streamflow trends in snow- and glacier melt dominated catchments of the Tarim River, Central Asia. Water Resour. Res. 51, 4727–4750. https://doi.org/10.1002/2014WR016716.Received

Eldho, T.I., Kulkarni, A.T., 2017. Conceptual and Physically Based Hydrological Modeling. Sustain. Water Resour. Manag. 81–118. https://doi.org/10.1061/9780784414767.ch04

François, B., Schlef, K.E., Wi, S., Brown, C.M., 2019. Design considerations for riverine floods in a changing climate – A review. J. Hydrol. 574, 557–573. https://doi.org/10.1016/j.jhydrol.2019.04.068

Giuntoli, I., Villarini, G., Prudhomme, C., Hannah, D.M., 2018. Uncertainties in projected runoff over the conterminous United States. Clim. Change 150, 149–162. https://doi.org/10.1007/s10584-018-2280-5

Giuntoli, I., Villarini, G., Prudhomme, C., Mallakpour, I., Hannah, D.M., 2015. Evaluation of global impact models' ability to reproduce runoff characteristics over the central United States. J. Geophys. Res. Atmos. 120, 9183–9159. https://doi.org/doi:10.1002/2015JD023401

Gosling, S.N., Arnell, N.W., 2011. Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis. Hydrol. Process. 25, 1129–1145. https://doi.org/10.1002/hyp.7727

Gudmundsson, L., Wagener, T., Tallaksen, L.M., Engeland, K., 2012. Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. Water Resour. Res. 48, 1–20. https://doi.org/10.1029/2011WR010911

Haarsma, R.J., Roberts, M.J., Vidale, P.L., Catherine, A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N.S., Guemas, V., Von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L.R., Lu, J., Luo, J.J., Mao, J., Mizielinski, M.S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., Von Storch, J.S., 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geosci. Model Dev. 9, 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Hannah, D.M., Demuth, S., van Lanen, H.A.J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., Tallaksen, L.M., 2010. Large-scale river flow archives: importance, current status and future needs. Hydrol. Process. https://doi.org/10.1002/hyp.7794

Jha, M., Pan, Z., Tackle, E.S., Gu, R., 2004. Impacts of climate change on streamflow in the Upper Mississippi River Basin: A regional climate model perspective. J. Geophys. Res. D Atmos. 109, 1–12. https://doi.org/10.1029/2003JD003686

Kam, J., Sheffield, J., 2016. Changes in the low flow regime over the eastern United States (1962–2011): variability, trends, and attributions. Clim. Change 135, 639–653. https://doi.org/10.1007/s10584-015-1574-0

Kollat, J.B., Kasprzyk, J.R., Thomas, W.O., Miller, A.C., Divoky, D., 2012. Estimating the Impacts of Climate Change and Population Growth on Flood Discharges in the United States. J. Water Resour. Plan. Manag. 138, 442–452. https://doi.org/10.1061/(asce)wr.1943-5452.0000233

Lavell, A., Oppenheimer, M., 2012. Climate Change : New Dimensions in Disaster Risk, Exposure, Vulnerability, and Resilience. Manag. Risks Extrem. Events Disasters to Adv. Clim. Chang. Adapt. 25–64.

Lavers, A.D., Villarini, G., 2013. Atmospheric Rivers and Flooding over the Central United States. J. Clim. 7829–7836. https://doi.org/10.1175/JCLI-D-13-00212.1

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis. Clim. Change 75, 273–299. https://doi.org/10.1007/s10584-006-6338-4

Li, L., Diallo, I., Xu, C.Y., Stordal, F., 2015. Hydrological projections under climate change in the near future by RegCM4 in Southern Africa using a large-scale hydrological model. J. Hydrol. 528, 1–16. https://doi.org/10.1016/j.jhydrol.2015.05.028

Lu, W., Wang, W., Shao, Q., Yu, Z., Hao, Z., Xing, W., Yong, B., Li, J., 2018. Hydrological projections of future climate change over the source region of Yellow River and Yangtze River in the Tibetan Plateau: A comprehensive assessment by coupling RegCM4 and VIC model. Hydrol. Process. 32, 2096–2117. https://doi.org/10.1002/hyp.13145

Mallakpour, I., Villarini, G., 2015. The changing nature of flooding across the central United States. Nat. Clim. Chang. https://doi.org/10.1038/nclimate2516

Maraun, D., 2016. Bias Correcting Climate Change Simulations - a Critical Review. Curr. Clim. Chang. Reports 2, 211–220. https://doi.org/10.1007/s40641-016-0050-x

Mearns, L.O., McGinnis, S., Korytina, D., 2017. The NA-CORDEX dataset.

Nakamura, J., Lall, U., Kushnir, Y., Robertson, A.W., Seager, R., 2013. Dynamical Structure of Extreme Floods in the U.S. Midwest and the United Kingdom. J. Hydrometeorol. 14, 485–504. https://doi.org/10.1175/jhm-d-12-059.1

Natural Resources Canada, 2017. Federal Floodplain Mapping Framework. Government of Canada.

Nayak, M.A., Villarini, G., 2017. A long-term perspective of the hydroclimatological impacts of atmospheric rivers over the central United States. Water Resour. Res. 53, 1144–1166.

Formatted: Italian (Italy)

https://doi.org/doi:10.1002/ 2016WR019033

Neri, A., Villarini, G., Kaustubh, Salvi, A., Slater, L.J., Napolitano, F., 2019a. On the decadal predictability of the frequency of flood events across the U.S. Midwest. Int. J. Climatol. 3, 1796–1804. https://doi.org/10.1002/joc.5915

Neri, A., Villarini, G., Slater, L.J., Napolitano, F., 2019b. On the statistical attribution of the frequency of flood events across the U.S. Midwest. Adv. Water Resour.

Pierce, D.W., Cayan, D.R., Thrasher, B.L., 2014. Statistical Downscaling Using Localized Constructed Analogs (LOCA)*. J. Hydrometeorol. 15, 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1

PRISM, 2017. Prism Climate Group [WWW Document].

Schlef, K.E., François, B., Robertson, A.W., Brown, C., 2018. A General Methodology for Climate-Informed Approaches to Long-Term Flood Projection—Illustrated With the Ohio River Basin. Water Resour. Res. 54, 9321–9341. https://doi.org/10.1029/2018WR023209

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464. https://doi.org/10.1214/09-AOS712

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., M. Reichstein, A. Sorteberg, C. Vera, and X.Z., 2012. Changes in Climate Extremese and their Impacts on the Natural Physical Environment, Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge University Press, Cambridge, UK and New York, NY, USA. https://doi.org/10.1017/CBO9781139177245.006

Slater, L.J., Villarini, G., 2017. Evaluating the Drivers of Seasonal Streamflow in the U.S. Midwest. Water 9, 695. https://doi.org/10.3390/w9090695

Slater, L.J., Villarini, G., 2016. Recent trends in U.S. flood risk. Geophys. Res. Lett. 43, 12,428-12,436. https://doi.org/10.1002/2016GL071199

Sunde, M.G., He, H.S., Hubbart, J.A., Urban, M.A., 2017. Integrating downscaled CMIP5 data with a physically based hydrologic model to estimate potential climate change impacts on streamflow processes in a mixed-use watershed. Hydrol. Process. 31, 1790–1803. https://doi.org/10.1002/hyp.11150

Swain, S., Hayhoe, K., 2015. CMIP5 projected changes in spring and summer drought and wet conditions over North America. Clim. Dyn. 44, 2737–2750. https://doi.org/10.1007/s00382-014-2255-9

Taylor, K.E., Stouffer, R.J., Meehl, G. a, 2007. A Summary of the CMIP5 Experiment Design. World 4, 1–33. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Formatted: Italian (Italy)

Formatted: Italian (Italy)

van der Knijff, J.M., Younis, J., De Roo, A. P. J., 2010. LISFLOOD: a GIS--based distributed model for river basin scale water balance and flood simulation. Int. J. Geogr. Inf. Sci. https://doi.org/10.1080/13658810802549154

Villarini, G., Scoccimarro, E., White, K.D., Arnold, J.R., Schilling, K.E., Ghosh, J., 2015. Projected Changes in Discharge in an Agricultural Watershed in Iowa. J. Am. Water Resour. Assoc. 51, 1361–1371. https://doi.org/10.1111/1752-1688.12318

Villarini, G., Slater, L.J., 2017. Climatology of flooding in the United States, in: Cutter, S.L. (Ed.), Oxford Research Encyclopedia of Natural Hazard Science. Oxford University Press (OUP), Oxford, pp. 0–36.

Villarini, G., Smith, J.A., Baeck, M.L., Krajewski, W.F., 2011. Examining Flood Frequency Distributions in the Midwest U.S.1. JAWRA J. Am. Water Resour. Assoc. 47, 447–463. https://doi.org/10.1111/j.1752-1688.2011.00540.x

Winkler, J. a, Arritt, R.W., Pryor, S.C., 2012. Climate Projections for the Midwest: Availability, Interpretation and Synthesis. US Natl. Clim. Assess. Midwest Tech. Input Rep. 24. https://doi.org/9781610915113

Wobus, C., Gutmann, E., Jones, R., Rissing, M., Mizukami, N., Lorie, M., Mahoney, H., Wood, A.W., Mills, D., Martinich, J., 2017. Climate change impacts on flood risk and asset damages within mapped 100-year floodplains of the contiguous United States. Nat. Hazards Earth Syst. Sci. 17, 2199–2211. https://doi.org/10.5194/nhess-17-2199-2017

Zheng, H., Chiew, F.H.S., Charles, S., Podger, G., 2018. Future climate and runoff projections across South Asia from CMIP5 global climate models and hydrological modelling. J. Hydrol. Reg. Stud. 18, 92–109. https://doi.org/10.1016/j.ejrh.2018.06.004

Response to the Editor's comments on:

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

by

ANDREA NERI, GABRIELE VILLARINI, AND FRANCESCO NAPOLITANO

(Note: In the text below we have copied the Editor's comments verbatim.)

General Comments:

Dear authors,
The two reviewers of the original version of your article have now sent their reports for the revised version of the manuscript. As you may see below, both reviewers agree that your article improved in several ways. They both have, however, some minor to moderate comments and suggestions before acceptance of your article for publication in JoH. I would especially attract your attention on the comment from reviewer #1 that somehow claims that you have only partly, or poorly, address several of his/her comments and suggestions. I thus suggest the authors to carefully address all reviewers comments and suggestions, which I believe will eventually lead to improve an already very good manuscript.
B.

Response:

We thank the Editor for handling our manuscript. We have addressed the Reviewers' comments point-by-point below. Moreover, because of a error in selecting the source file of Figure 4, we replaced it with its correct version; this change does not significantly affect the results of the paper. We do believe that the review process has led to an improved manuscript, and we hope that it is now ready for publication.

Response to Reviewer 1's comments on:

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

by

ANDREA NERI, GABRIELE VILLARINI, AND FRANCESCO NAPOLITANO

(Note: In the text below we have copied the Reviewer's comments verbatim.)

General Comments:
I appreciate the authors' revision of the manuscript. While I think some of my original comments have been satisfactorily addressed, I do not think that is the case for all of them. My remaining comments are as follows.

Response:

We thank the Reviewer for his/her feedback and comments, which we addressed below point by point.

Comment 1)
Regarding the issue of bias correcting the GCMs; most of what is described sounds fine to me, but in lines 182-197 of the tracked changes manuscript, it is unclear to me whether the bias correction was applied to each GCM individually and then averaged? Or to the average across the 32 GCMs? If the former, then it seems there would still be the issue of smoothing out the extremes. Please clarify in the document.

Response:

Thank you for pointing this out. The bias correction was applied to the ensemble mean of the GCMs. In other words, we first calculated the average across the CGMs, and then we applied the bias-correction approach, so that we could still catch the variability of the climate variables. In revising the manuscript, we added some wording at lines 180-181 of the revised manuscript to clarify this issue.

Comment 2)
Regarding the phrase "with that said, we still expect years with more events alternating to more quiet ones". I appreciate that you explained it to me, but it did not change in the manuscript - and hence is likely to still be confusing to readers. I suggest to clean up some of the wording in your response to the comment and replace the original sentence.

Response:

Thank you for pointing this out. In revising the manuscript, we removed that sentence not to confuse the reader.

Comment 3)

Finally, I am quite unsatisfied with the authors' response regarding appropriate literature. I would like to see a significant improvement in both the introduction and either discussion or conclusion, that goes beyond simply the addition of a few sentences. Some particular comments regarding this issue are below.

I do not think that the literature review that was added to the introduction has been well-considered. For example, it is true that Jha et al. (2004) model only one catchment, but that one catchment covers a significant number of the catchments considered in your study - so to state that it is part of a group a studies that focuses on "floods at single specific catchments" is not really a fair assessment. I could make a similar comment for Schlef et al. (2018), which actually builds a model for 26 gages locations, spread over multiple states. This issue gives the impression that the authors did not take adequate time to understand the literature. I grant that this study on changes in frequency at the regional scale, using statistical models, is, to the best of my knowledge, novel and useful - but I am basing some of that understanding on the limited literature review provided by the authors. (Also in the introduction, regarding the different methods for long-term projection of floods, the recent review by François et al. 2019 in JoH is a good reference.)

There still is no reference to the literature, and how the results from this study fit into what is already known, in the discussion/conclusion. There are ample areas of discussion even from studies published by the same authors. For example, do the results of this study indicate similar or different trends to those seen in the historical period in Mallakpour & Villarini (2015)? Or do they align with what we know about changing frequency of heavy rainfall in Villarini et al. (2011)? Or, how does the short-term view of the next 10 years as discussed in Neri et al. (2019) in IJoC (with very similar methodology, just on the scale of 1-10 years, rather than 100 years, which should be probably noted in the introduction) compare with the longer-term view out to 2100 discussed in this work? (And, what about the issue of suppressed skill from basin-averaged precipitation predictions noted at the end of the abstract of Neri et al., 2019?) And, from other studies for example, do the projected changes make sense with what we know regarding general climate changes in Byun & Hamlet (2018)? These are just some questions that can and should be addressed. I would hope and expect the authors to do due diligence to address any other pertinent questions and literature of similar nature not listed here.

Response:

Thank you for highlighting these issues. The literature provided in the previous version of the manuscript has been revised significantly. We included the work of Jha et al. (2004) among those focusing on “specific single catchments” because it does focus on the projection with respect to a single catchment (Mississippi river Grafton, Illinois): even though it includes some of the water basins considered in our study, it focuses exclusively on that gage station and it does not allow to provide insights on the spatial variability of the projected changes of streamflow across the U.S. Midwest.

We agree with the reviewer that the wording "single catchment" does not properly reflect what presented in Schlef et al. (2018). We included this reference in that context because the work focuses on water basins belonging to a small area of the Ohio River basin and thus lacking the capability of analyzing different hydrological processes driven by different climatic conditions. In revising the manuscript, however, we removed this reference from the group of papers cited at lines 113-114 of the original manuscript because not completely congruent with them, and moved it to the second group of references at lines 114-115 of the revised manuscript, given that it focuses on flood magnitude rather than the frequency of flood events. Overall, the message we wanted to communicate to the reader by citing the two groups of papers at lines 113-114 and at lines 114-115 of the original manuscript is that the literature lacks of studies focusing on the catchment-specific projections of the frequency of flood events at a regional scale.

Moreover, in revising the manuscript, we also added the suggested reference of François et al. (2019) in the part of the introduction where we summarize the approaches used for flood events prediction.

We also agree with the Reviewer that the manuscript can be improved in terms of a proper discussion of how the present work advances the scientific knowledge about the projected changes in the frequency of flood events and of the similarities with respect to published papers. In revising the manuscript, we improved our discussions. Here we summarize the additions to the manuscript and the responses to the second part of the Reviewer's comment:

- At lines 261-266 of the revised manuscript we refer to some studies which show that observed trends in the frequency of flood events in the U.S. Midwest are similar to those obtained in the present work and we also recall the fact that it is the frequency of flood events, rather than its magnitude, which presents stronger time trends;
- At lines 244-248 of the revised manuscript we cite some studies to confirm the outcomes with respect to the analysis of the best drivers responsible for the variability in the frequency of flood events in the U.S. Midwest;
- With respect to the similarities between this work and the one investigating decadal predictions (Neri et al., 2019a), we already mentioned this aspect at the end of the introduction (line 123 of the original manuscript) and in the Method Section (line 205 of the original manuscript). To further stress the similarities, we added a sentence at lines 220-222 of the revised manuscript;
- It is quite hard to compare the results of the work focusing on decadal predictions (Neri et al., 2019a) with those of the present work for several reasons: (i) for the projections, we do not have observations to evaluate the skill of our future prediction, which is instead assessed in Neri et al. (2019a); (ii) this work focuses on the projected changes in the frequency of flood events, not on the skill of the

models in reproducing the observations (which is addressed in Neri et al. (2019a) and in Neri et al. (2019b)); (iii) the projections are based on possible future scenarios and pathways (like for instance the increasing emissions of carbon dioxide) and change during the investigation period, while in the decadal predictions the forecasts are initialized based on the observed state of the atmosphere. Therefore, we did not include a detailed comparisons between the two studies.

- We analyzed the work of Byun & Hamlet (2018), related to the projected changes of precipitation and temperature in the U.S. Midwest. They show that according to ten statistically downscaled GCMs, the U.S. Midwest seasonal temperature increases during all four seasons and precipitation increases during all seasons except summer, which shows a decreasing trend. These results are in accordance with our climate ensemble obtained from the LOCA and NA-CORDEX datasets. However, by looking at the projected trends in the frequency of flood events, the summer season is characterized by positive significant trends in most of the stations (even though the average increase in flood counts does not appear to be very strong; Figure 6). This is driven by two main reasons. The first reason is that in some stations during summer the best model selected is the P.M model, i.e., the model considering precipitation during summer and precipitation during spring (i.e., the antecedent wetness conditions) as predictors. Because the increase of precipitation during spring is higher compared to the decrease of precipitation during summer, the antecedent wetness conditions predictor has more of an effect on the flood counts, resulting in increasing trends in the frequency of flood events. The second reason is related to the stations where the best model is the P model, i.e., the model considering only precipitation as predictor: in these stations the trend in the projected frequency of flood events is positive, even though the projected seasonal precipitation shows a negative trend. However, given the fact that the decrease in precipitation is quite small and the variability is high, the discrete Poisson regression model does not catch the overall pattern, but only extreme events, which actually generate positive trends. In revising the manuscript, we discussed this issue in the Results and Discussion Section at lines 298-307 of the revised manuscript.

Response to Reviewer 2's comments on:

Statistically-based projected changes in the frequency of flood events across the U.S. Midwest

by

ANDREA NERI, GABRIELE VILLARINI, AND FRANCESCO NAPOLITANO

(Note: In the text below we have copied the Reviewer's comments verbatim.)

General Comments:
I had reviewed a previous version of this manuscript. I liked the manuscript then, and the authors' responses have greatly improved the manuscript. The authors have explored historical and future (based on GCMs) trends in the frequency of historic peak daily average streamflow. The methods are well-documented, and the discussion does a great job of exploring the implications and limitations of the results. My additional comments are minor and are intended to deepen the impact of the work.

Response:

We thank the Reviewer for his/her comments and feedback. We have addressed the comments point by point.

Comment 1)
On line 141, the authors discuss the data of their peaks over threshold approach. I wonder if they could add a comment on the impact of using daily average streamflow rather than instantaneous peaks. I don't think their work is invalidated in anyway, but I think it important to point out that characterizing a high daily mean streamflow as a peak is different than traditional flood-frequency analysis that uses instantaneous peaks. (The daily averages will be lower, right?).

Response:

Thank you for pointing this out. In our analysis we use daily discharge records from 287 USGS gauging stations and we select the flood events using a peak-over-threshold approach. We agree with the Reviewer that the daily discharge peak is lower than the instantaneous peak (especially for small watersheds), but we think that since we are focusing on the number of events, the magnitude of the peak (whether daily or instantaneous) should not significantly impact the flood count time series. In other words, the way the flood events are selected (lines 140-149 of the original manuscript) implies that the x highest discharge values are selected; assuming that the ratio between daily and instantaneous peak is constant for all the flood events for a given basin, the selected flood events using either the daily or instantaneous time series should be the same. In revising the manuscript, we have added text on lines 152-156 to clarify this issue.

Comment 2)
Reading around line 165, I am reminded of the author's claim circa line 104. Around line 104, it is claimed that grid-based data lacks applicability to basins and gages. Here, around line 165, the authors seem to be using grid-based products for their methods. What is the impact of gridded products? How is the problem identified around line 104 surmounted here (around 165)?

Response:

Thank you for pointing this out. The limitations discussed at lines 107-117 of the original manuscript do not refer to the climate data used as input to the models, but to their outputs, i.e., the variable of interest. The works mentioned at lines 83-93 and at lines 104-105 of the original manuscript provide projections of the quantity of interest (i.e., flood frequency curves, discharge quintiles) in each pixel of a gridded map, which is often quite coarse because of computational requirements. Through our statistical models, instead, we are able to assess projections relative to a specific location (i.e., the gauging station) rather than to a pixel, being therefore more accurate and more useful for specific local flood mitigation plans. In other words, we are not questioning the usefulness of the gridded climate data, but rather that of the gridded projections outputs of hydrological models.

Comment 3)
On line 172, the authors note that they take the ensemble mean of GCM projections. While this is quite common, I wonder if the authors could add a statement on how this reduces the variability of the analysis. Given that we are looking at trends, that variability, which is currently averaged out, may be very important. Could averaging have been done later in the methods? At a minimum, and it's discussed in the authors' response to previous comments, I think it important to discuss this loss of variability.

Response:

Thank you for pointing this out. At lines 180-193 of the revised manuscript we explain the approach used to bias-correct the GCM ensembles. This approach, apart from correcting the average of the time series, also corrects its variability. By averaging the GCM members, in fact, the variability of the time series reduces substantially. In order to deal with this loss of variability, we also correct the standard deviation of both the historical and projection period, so that extreme events could be preserved in our precipitation and temperature basin-averaged time series.

Comment 4)
The sentence on line 107 with "However…" needs to be revised. There are several connecting verbs missing. While I understand the intent of the sentence, I can't quite figure out what is wrong with it. One addition: "…these hydrological models [are] distributed over …"

Response:

Thank you for pointing this out. We have revised the sentence to read: "because the outputs of these hydrological models are distributed over grid cells and not necessarily related to the specific gauge station (Giuntoli et al., 2015), their applicability loses…"

Comment 5)
The sentence that traverses line 120 needs some additional words. For example, it should probably be "...because [they are] faster to implement, less time-consuming ...".

Response:

Thank you for pointing this out. In revising the manuscript we have changed the sentence to "because these models are faster to implement…"

Comment 6)
On line 238, please provide a specific, quantifiable metric that leads the authors to the conclusion that the results are "overall good".

Response:

Thank you for pointing this out. Figure 3 shows the values of the correlation coefficient computed between observed and modeled flood counts time series. The map suggests higher or lower values depending on the location and/or on the season, but considering all the stations together, the R values are satisfactorily high. Because evaluating the model performance is not the main objective of the work, we used the word "overall" to assess the model performance in a general way, referring the reader to a previous study for a more detailed evaluation of the statistical framework skill. In revising the manuscript, we have added text to say that the average correlation coefficient among the different seasons is 0.56.

Comment 7)
On line 254, please add an explicit sentence summarize your conclusions. Despite disagreement, you feel comfortable proceeding to projection. Please state that explicitly, with a specific quantifiable metric to support it, if possible.

Response:

Thank you for pointing this out. We agree that a sentence introducing the projection analysis is required at the end of the historical simulations assessment. In revising the manuscript, we added a phrase (at lines 273-277 of the revised manuscript) which introduces to the projection results.

Comment 8)
On the paragraph ending on line 294: No change needed here, just wanted to say that this is a really excellent job of capturing and discussing this limitation.

Response:

Thank you!

Comment 9)
In the conclusions, please consider adding a sentence to remind the reader of the novelty of this work.

Response:

Thank you for pointing this out. In revising the manuscript, we have added a sentence in the Conclusions to summarize the novelty of this work.

Table 1
Click here to download high resolution image

Model Name	Dependence
P	$log\lambda_1 = \alpha_1 + \beta_1 x_P$
P.T	$log\lambda_2 = \alpha_2 + \beta_2 x_P + \gamma_2 x_T$
P.M	$log\lambda_3 = \alpha_3 + \beta_3 x_P + \gamma_3 x_M$
Mixed	$log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Mar-Apr}}$ $log\lambda_4 = \alpha_4 + \beta_4 x_P + \gamma_4 x_M + \delta_4 x_{T_{Summer}}$

Figure 1
Click here to download high resolution image

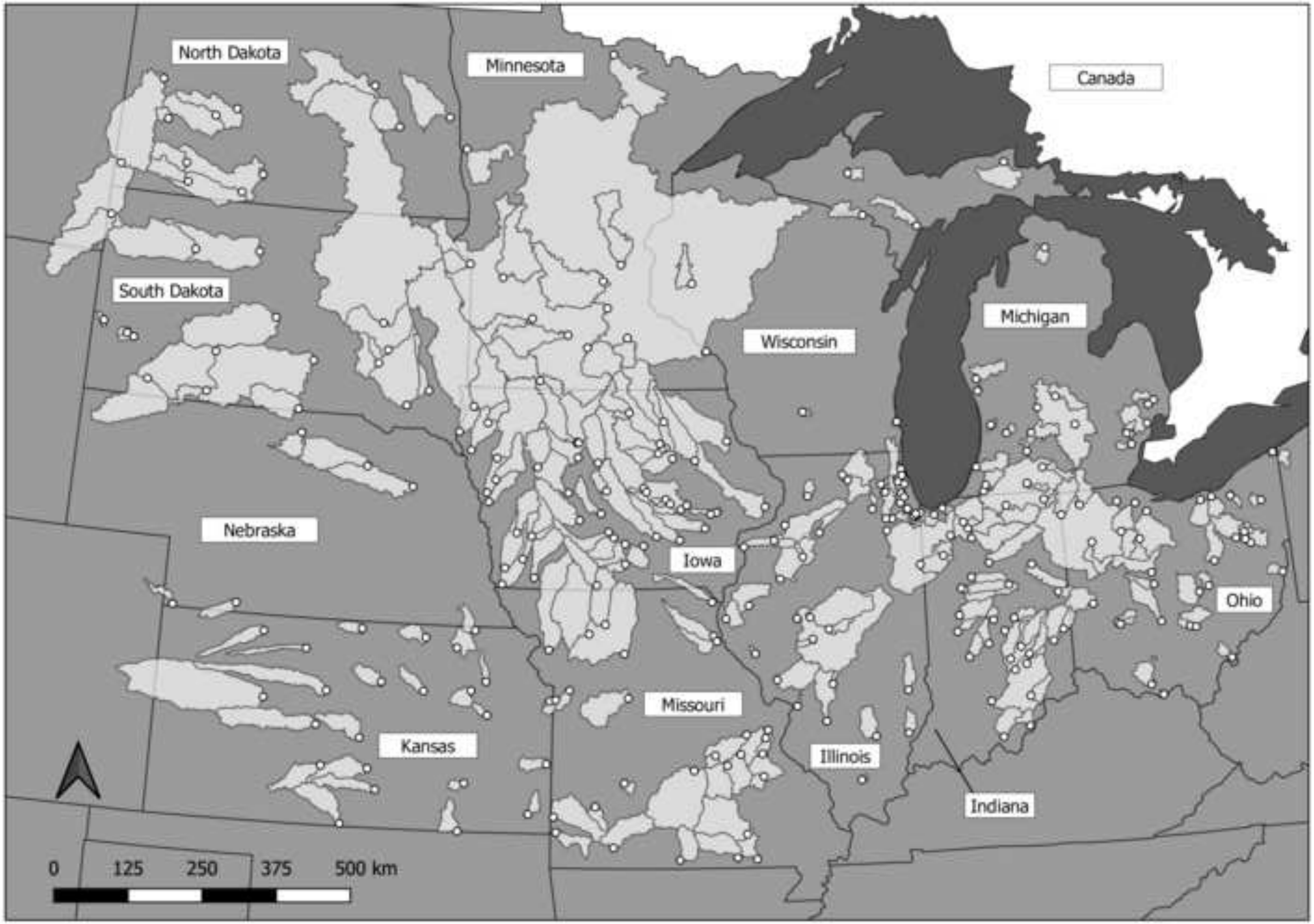


Figure 2
Click here to download high resolution image

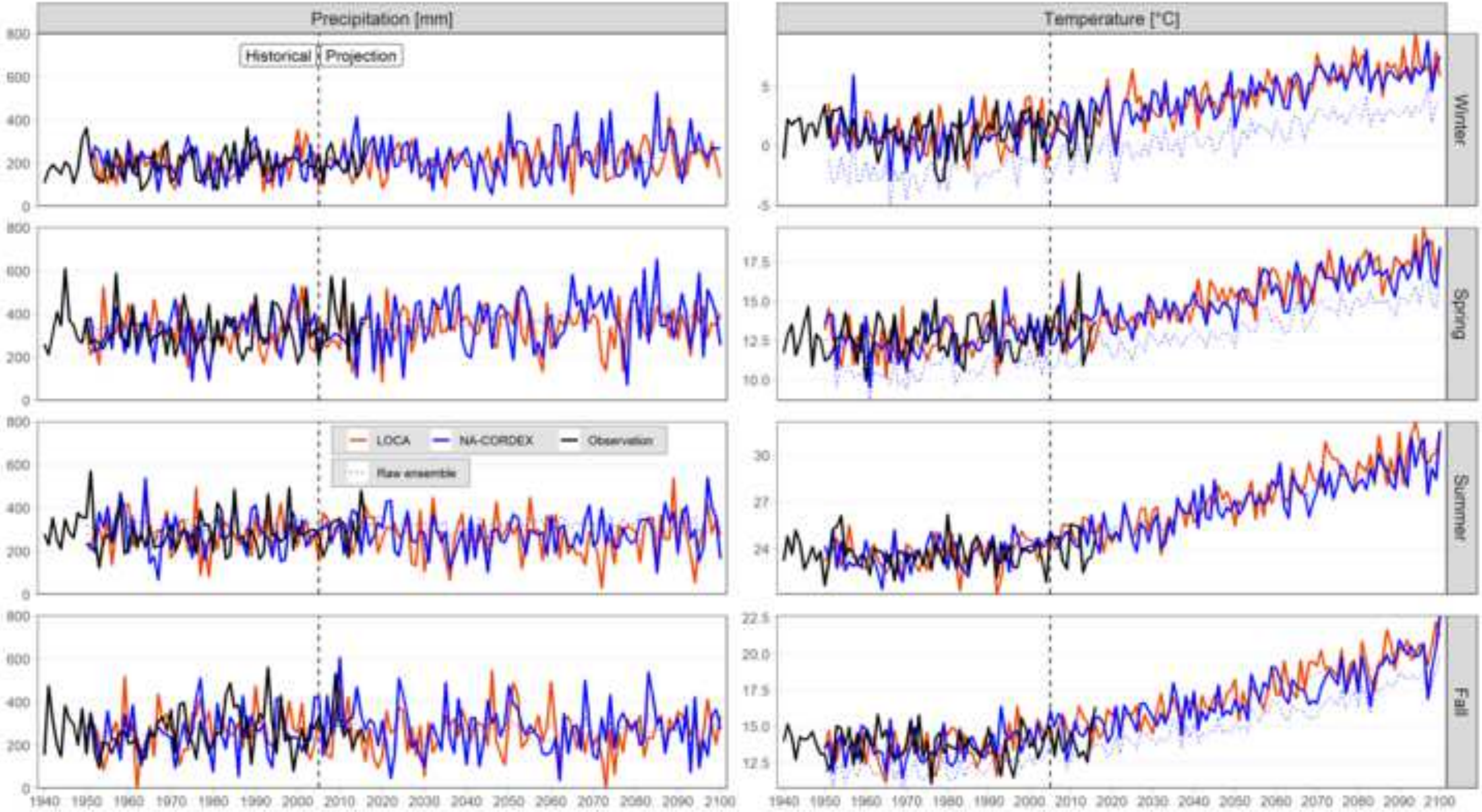


Figure 3
Click here to download high resolution image

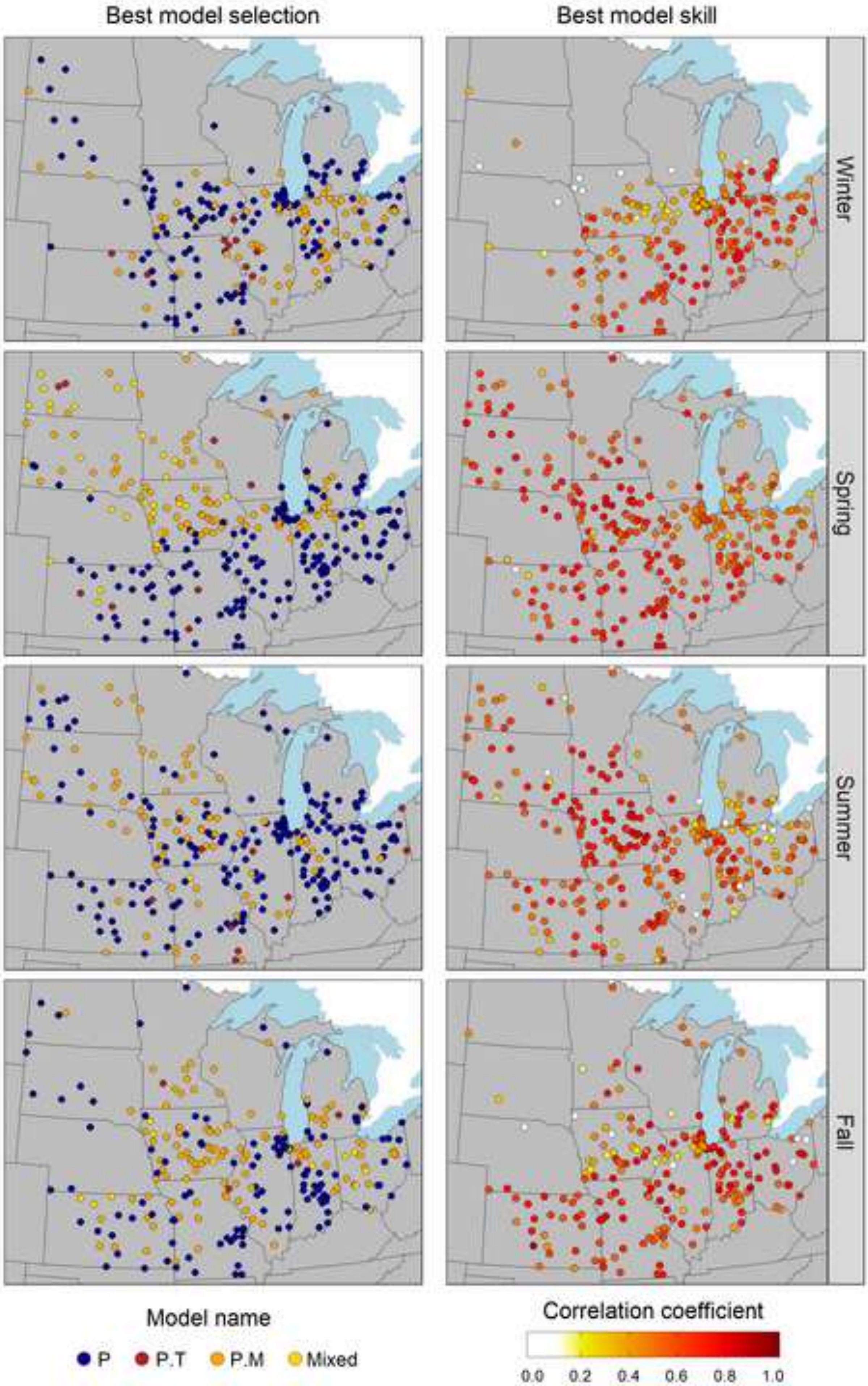


Figure 4
Click here to download high resolution image

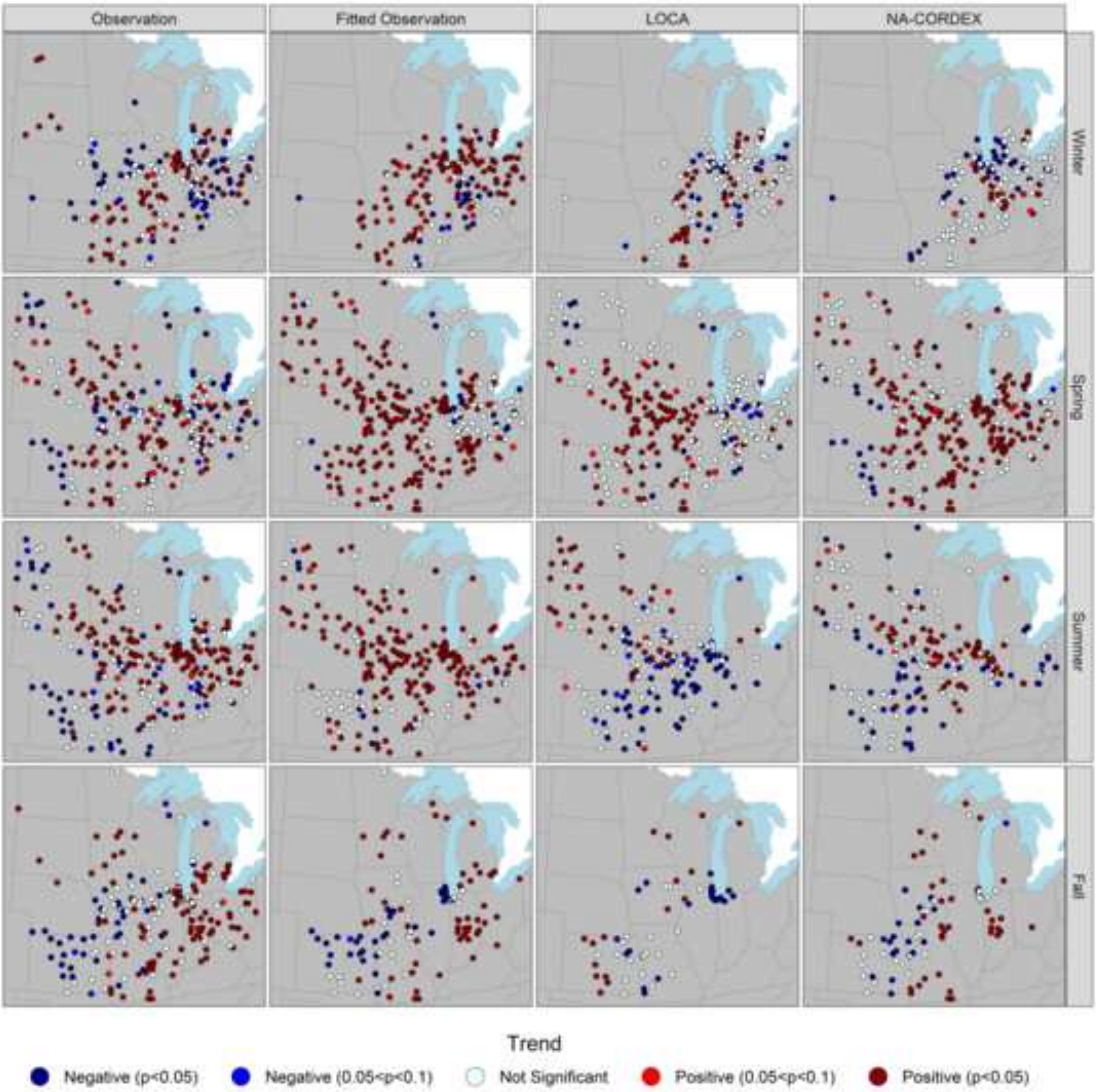


Figure 5
Click here to download high resolution image

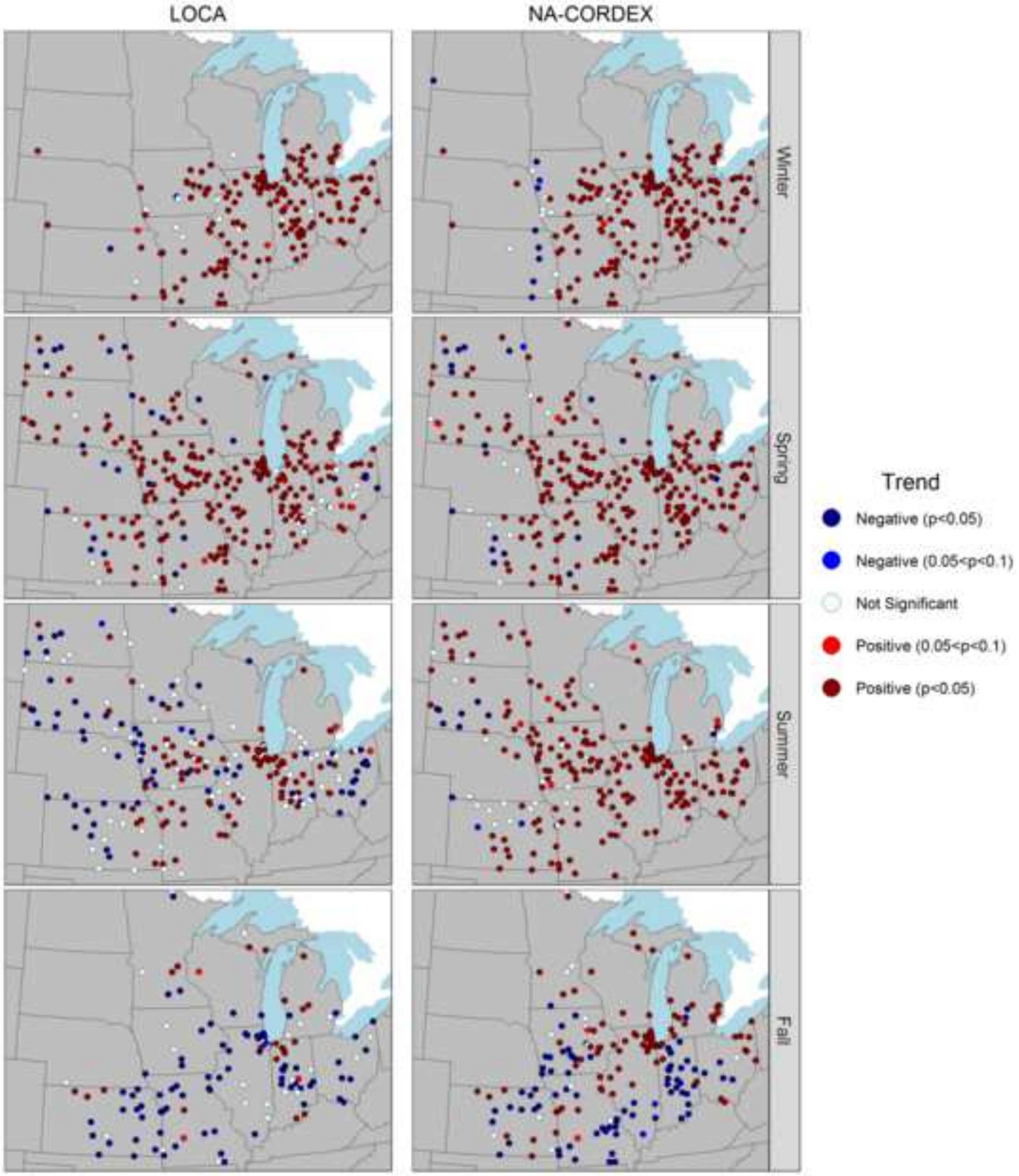
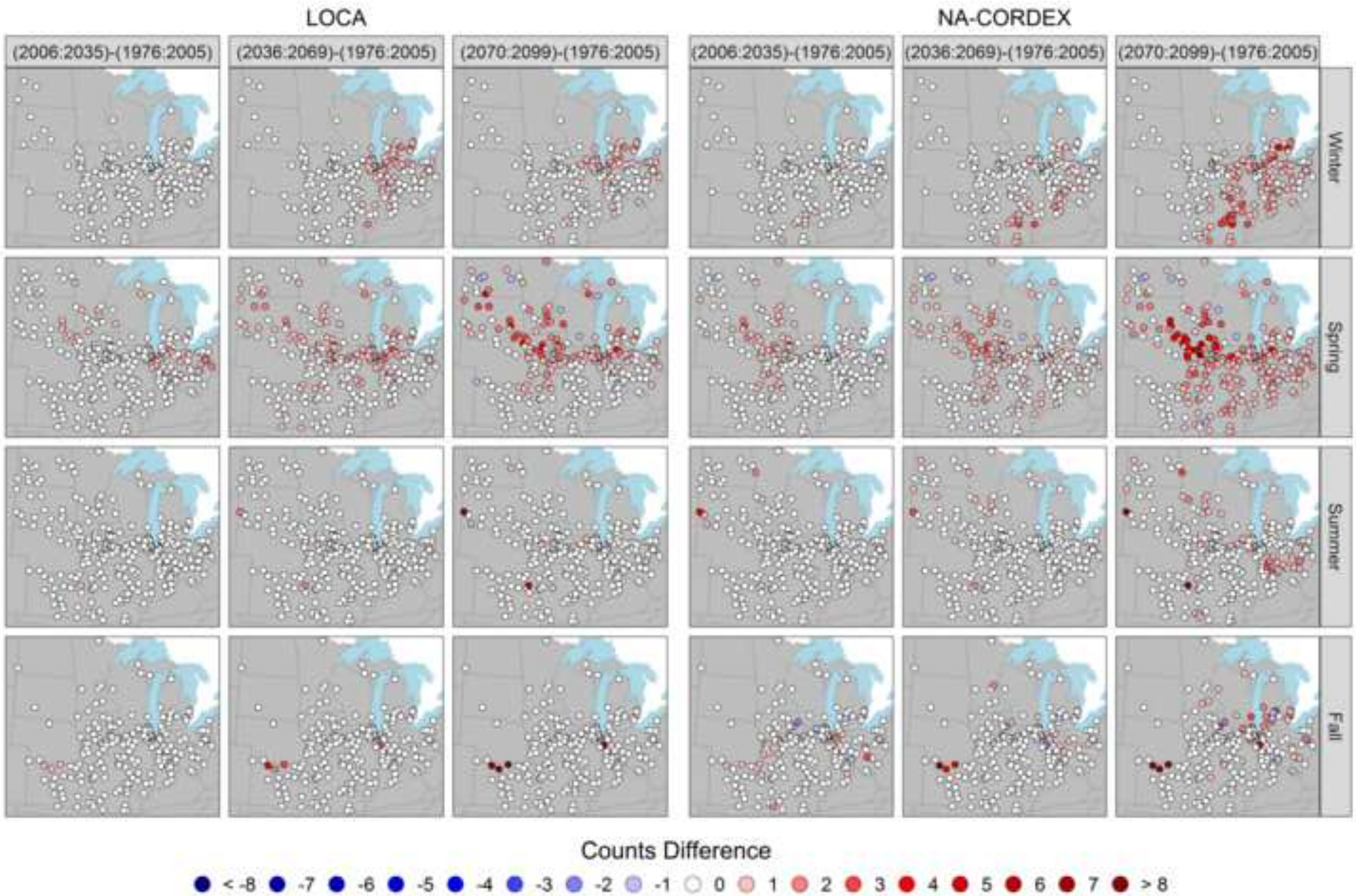


Figure 6
Click here to download high resolution image



Figures and table captions

Table 1 – List of the four statistical models used to relate the seasonal occurrence of flood events to the three drivers: precipitation (x_P), wetness conditions (x_M) and temperature (x_T).

Figure 1 - Location of the 286 USGS gauging stations (white circles) and the relative upstream drainage area (light gray polygons).

Figure 2 – Observed, historical and projected basin-averaged bias-corrected precipitation (left panels) and temperature (right panels) seasonal time series for USGS station 07014500 (Meramec River near Sullivan, Missouri). The black line represents the observed values, while the red and blue solid (dotted) lines the values based on bias corrected (raw) LOCA and NA-CORDEX, respectively.

Figure 3 – Map showing the selected best models (four panels on the left) and their skill (four panels on the right) for each season (rows) and for a flood threshold value that gives two peaks per year on average. The blue, brown, orange and yellow circles on the left refer to the P, $P.T$, $P.M$, and $Mixed$ models, respectively. In some stations, no model is selected largely because the observed time series does not have at least five years with a flood count value different from zero. Note that not every best model provides a predicted time series with at least five years with a flood count value different from zero, therefore no correlation coefficient can be computed.

Figure 4 – Map showing the trends of the frequency of flood events during the historical period (i.e., 1950-2005) for the observations and for the median of the Poisson regression models using observations, LOCA and NA-CORDEX. The results refer to the threshold values that gives two peaks per year on average among the observational period 1940-2016 (see Figure S3 of the supplemental material for the other flood threshold values). The four columns represent the four considered datasets, and the rows the four seasons. In each panel, the dark red and dark blue (red and blue) circles indicate the positive and negative trends significant at the 5% (10%) significance level, respectively. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero.

Figure 5 – Map showing the projected trends (2006-2100) in the frequency of flood events based on the LOCA (left column) and NA-CORDEX (right column) for the four seasons and for a flood threshold value that gives two peaks per year on average. The symbol notation is the same as in Figure 4. The trend in some stations is not estimated because the predicted time series does not have at least five years with a flood count value different from zero. The same results relative to the other flood thresholds are presented in Figure S4 of the supplemental material.

Figure 6 – Map showing the difference between the average value of flood counts during three spans (i.e., 2006:2035, 2036:2069, 2070:2099) of the projection period and the average value of flood counts during the last 30 years of the historical period (i.e., 1976:2005) using the LOCA (left set of panels) and NA-CORDEX (right set of panels) dataset. The results refer to the threshold values that give, two peaks per year on average among the observational period 1940-2016. For each set of panels, the three columns represent the three considered spans, and the rows the four seasons. For the other flood threshold values see Figure S5 and Figure S6 of the supplemental material for the LOCA and NA-CORDEX datasets, respectively.