

Geophysical Research Letters



RESEARCH LETTER

10.1029/2019GL086797

Uncertainties in Future U.S. Extreme Precipitation From Downscaled Climate Projections

Tania Lopez-Cantu¹ , Andreas F. Prein² , and Constantine Samaras¹ 

¹Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, USA, ²National Center for Atmospheric Research, Boulder, CO, USA

Key Points:

- Extreme daily rainfall changes for the United States from five widely used public downscaled climate projection data sets are compared
- Large differences in magnitude and spatial patterns of projected extreme precipitation changes exist between data sets
- High-end extremes (e.g., 100-year event) increase 10–50 % faster than low-end extremes (e.g., 2-year event) at the continental scale

Supporting Information:

- Supporting information S1

Correspondence to:

T. Lopez-Cantu,
tlopez@andrew.cmu.edu

Citation:

Lopez-Cantu, T., Prein, A. F., & Samaras, C. (2020). Uncertainties in future U.S. extreme precipitation from downscaled climate projections. *Geophysical Research Letters*, 47, e2019GL086797. <https://doi.org/10.1029/2019GL086797>

Received 7 JAN 2020

Accepted 30 MAR 2020

Accepted article online 9 APR 2020

Abstract Impacts modelers and stakeholders use publicly available data sets of downscaled climate projections to assess and design infrastructure for changes in future rainfall extremes. If differences across data sets exist, infrastructure resilience decisions could change depending on which single data set is used. We assess changes in U.S. rainfall extremes from 2044–2099 compared with 1951–2005 based on 227 projections under RCP4.5 and RCP8.5 from five widely used data sets. We show there are large differences in the change magnitude and its spatial structure between data sets. At the continental scale, the data sets show different increases, with high-end extremes (e.g. 100-year event) generally increasing more (between 10% and 50%) than low-end extremes (e.g. 5-year). These differences largely contribute to the overall uncertainty for small average recurrence intervals (ARIs) extremes (2- to 10-year), while uncertainties due to short record length dominate large ARIs (25- to 100-year). The results indicate that robust infrastructure planning should consider these uncertainties to enable resilient infrastructure under climate change.

Plain Language Summary Observed extreme rainfall magnitudes have increased since 1950, and climate model projections indicate that these increases will continue throughout the 21st century in many areas in the US. Adapting and designing infrastructure for climate change requires future extreme rainfall projections at high resolutions. Global climate model (GCM) output resolution is generally much coarser, and methods have been developed to create relevant climate information at regional scales. Multiple open data sets exist that provide downscaled projections of future rainfall extremes. We analyze changes in future daily extremes using five widely used data sets in climate change impacts assessments and decision making. We found large differences between data sets in how much and where extreme events will intensify by the end of the 21st century. This and other sources of uncertainty need to be considered when designing resilient infrastructure.

1. Introduction

Extreme rainfall events have intensified in magnitude and become more frequent in many regions of the United States (U.S.) since 1950 (DeGaetano, 2009; Hoerling et al., 2016), and climate model simulations project these increases will continue throughout the 21st century (Prein et al., 2017). Observed and anticipated changes in rainfall extremes affect many sectors, including existing and proposed infrastructure such as stormwater systems. In the United States, stormwater infrastructure is designed using federal, state, and local design standards based on historical extreme rainfall probabilities, and most new infrastructure is similarly not designed to consider future climate change (Lopez-Cantu & Samaras, 2018; Wright et al., 2019). Making infrastructure systems resilient to future conditions involves identifying risks and vulnerabilities to inform climate resilient strategies (IPCC, 2014). However, it is not straightforward to develop these strategies because it requires making decisions under a deeply uncertain future climate (Hallegatte, 2009). Yet, taking no action is also a decision, and large economic damages can occur without adaptation (Martinich & Crimmins, 2019).

Many stakeholders incorporate future climate conditions into their planning and analyses by using downscaled global climate model (GCM) output to inform hydrologic models or update current engineering design standards (Forsee & Ahmad, 2011; Kuo et al., 2015). Although the downscaling process is necessary to match the spatial and temporal resolution to the required resolution by the specific application (Cook et al., 2017), different downscaling methods and resolutions can potentially give different results (Arnbjerg-Nielsen et al., 2015; Cook et al., 2020; Onof & Arnbjerg-Nielsen, 2009; Sunyer et al., 2012;

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Table 1*Some of the Publicly Available Downscaled Climate Products for the United States and an Overview of Their Characteristics*

Data set	GCM models	Emissions scenario	Downscaled spatial resolution (degrees)	Highest temporal resolution	Downscaling technique
BCCAv2	21	2.6, 4.5, 6.0, 8.5	1/8	Daily	Statistical
MACA	20	4.5, 8.5	1/24	Daily	Statistical
LOCA	32	4.5, 8.5	1/16	Daily	Statistical
NA-CORDEX	6 ^b	4.5, 8.5	0.22, 0.44	Hourly	Dynamical

The number of GCM models dynamically downscaled depend on the target resolution and the emissions scenario.

Wu et al., 2019). For the Contiguous United States (CONUS), there are several downscaled climate projection data sets covering at least the late 20th century and the 21st century (Abatzoglou & Brown, 2012; Mearns et al., 2013; Maraun et al., 2010; Pierce et al., 2014) (see Table 1 and Figure S2 and Tables S1 to S3 in the supporting information). These data sets vary in downscaling method used, number of downscaled GCMs, emission scenarios, and horizontal resolution. The total number of available projections from the data sets in Table 1 is 103 for Representative Concentration Pathway (RCP) 4.5 and 124 for RCP8.5. The massive amount of data, limited guidance, or lack of compelling arguments to discard or adopt one or more data sets in Table 1, in addition to different ease of use among these data sets, are often reasons why many impact assessments only use a single data set as input—neglecting possible differences across the data sets and their influence on the specific study area.

In this study, we assess the daily extreme rainfall climate change signal across the U.S. from five downscaled climate projection data sets (Table 1) for RCP4.5 and 8.5. We use five downscaled climate projection data sets (hereafter, *data sets*) consisting of multiple simulations (hereafter, *members*), while the collection of available simulations from all five data sets for each RCP scenario is referred to as *ensemble*. We focus on these data sets because they all downscaled GCMs from the Coupled Model Intercomparison Project Phase 5 (CMIP5), share a common temporal resolution and spatial coverage, comprise two RCP scenarios, are publicly available, and are frequently used as inputs in climate change impacts assessments for infrastructure (Alamdari et al., 2017; Cook et al., 2017; Gelda et al., 2019; Kermanshah et al., 2017).

We use the generalized extreme value (GEV) distribution to describe the extreme rainfall data and quantify the climate change signal in terms of change factors corresponding to the ratio between the model future rainfall volume (estimated from annual maximum rainfall series for the 2044–2099 period) and model historical volume (1951–2005) of average recurrence interval (ARI) events ranging from 2- to 100-year. Additionally, because these data sets are used in a decision-making context that requires information about the relevant uncertainties (Mullan et al., 2018), we quantify the relative contribution of three sources of uncertainty: (1) the GEV distribution fit uncertainty related to the record length to fit the distribution and its accuracy to describe the extremes, (2) the data set selection uncertainty from the different downscaling methods applied to a set of GCMs, and (3) the uncertainty stemming from the emissions scenario (RCP4.5 and RCP8.5).

2. Materials and Methods

2.1. Publicly Available Downscaled U.S. Climate Projections Data Sets

The data sets compared in this study can be categorized by the type of downscaling process used. The Bias-Corrected Constructed Analogs (BCCA) version 2, Multivariate Adaptive Constructed Analogs (MACA), and the Localized Constructed Analogs (LOCA) data sets are products of statistically downscaling a subset of CMIP5 GCMs (Abatzoglou & Brown, 2012; Maraun et al., 2010; Pierce et al., 2014). The constructed analog (CA) method constitutes the base of each statistical downscaling method. CA involves comparing the GCM control simulations to historical observations, traditionally in terms of anomalies, but all procedures above use absolute values instead. The historical observations are regridded to match the resolution of the GCM simulations. GCM simulations are typically bias-corrected using a quantile mapping (QM) approach against other source of historical observations (Abatzoglou & Brown, 2012; Maraun et al.,

2010; Pierce et al., 2014). For a given GCM day, the goal is to find a set of days (situated within a maximum of 45 days centered at the GCM target day) in the regrided observations that when combined, they approximate the climate conditions on the day simulated by the GCM (Hidalgo et al., 2008).

The main difference between BCCAv2, MACA, and LOCA lies in the searchable domain to find the analog days. For BCCAv2 and MACA, the analog days are selected from any grid within the domain (CONUS), while in LOCA, the analogs are identified at smaller climatically similar regions. Another difference lies in the number of variables that are jointly downscaled. In BCCAv2 and LOCA, variables are independently downscaled, while in MACA a multivariate approach is used for variables other than precipitation (Abatzoglou & Brown, 2012). Also, while the BCCAv2 and MACA methods use a linear combination of 30 days to construct the CA at each grid cell, a single day out of the 30 identified that is closest to GCM target day pattern across the neighboring region to each grid cell is used in the LOCA procedure (Bracken, 2016; Pierce et al., 2014). Finally, the MACA method applies a final bias-correction step (once again QM) to the downscaled output, in contrast to the BCCA and LOCA methods.

The final data set we assess is the North American branch of the Coordinated Regional Downscaling Experiment (NA-CORDEX) data set, which is a dynamically downscaled data set. There are projections available in two different resolutions (see Table 1), and since they include different members (see Tables S2 and S3), we treated each as an individual data set. Dynamical downscaling techniques use physics-based models or regional climate models (RCMs) constrained to a smaller region and are driven by GCM projections at the boundaries to produce higher resolution projections. The NA-CORDEX data set corresponds to a collection of high-resolution projections from different GCM-RCM combinations and emissions scenarios (Mearns et al., 2014). Figure S1 shows a detailed comparison of the methodological steps of the process followed to produce each data set.

In this paper, our focus is extreme rainfall changes, and it is important to highlight the limitations associated with the representation of extreme rainfall processes in each downscaling method. For dynamical downscaling, RCMs may differ from each other in the formulation of the model numerics and physics (e.g., convection and microphysics). Even RCMs with similar parametrization schemes might simulate daily precipitation statistics that are substantially different (Fowler et al., 2007; Prein et al., 2015). On the other hand, the representation of extreme rainfall in statistical methods is limited by the relationship between daily climate model data and regrided observations (Cannon, 2018). Another limitation is caused by poorly constrained relationships at high quantiles of the rainfall distribution due to a lack of a long set of historically observed extreme rainfall events (Maraun, 2016). Furthermore, statistical methods assume that the model biases are stationary and do not change under climate warming, which can result in biased downscaled future climate projections (e.g., Dixon et al., 2016).

2.2. Estimation of Future Changes in Daily Rainfall Extremes

To compare the extreme rainfall climate change signal between data sets, we first regrided all ensemble members to the NA-CORDEX 0.44° grid (i.e., the coarsest grid spacing across data sets used in this study) using a conservative remapping approach (Jones, 1999) to minimize differences stemming from the data sets' output spatial resolution.

For each ensemble member, extreme rainfall was characterized using a stationary GEV distribution fitted to series of annual maxima (Coles, 2001) for two periods, one future period between 2044 and 2099, and a historical period from 1950 to 2005. Assuming that the annual maximum rainfall series, z , follows a GEV distribution, its cumulative distribution function is defined as

$$F(x) = \begin{cases} e^{-\left(1 + \frac{\xi}{\sigma}(x-\mu)^{-\frac{1}{\xi}}\right)} & \xi \neq 0 \\ e^{-e^{-\left(\frac{x-\mu}{\sigma}\right)}} & \xi = 0 \end{cases}, \quad (1)$$

where μ , σ , and ξ are the parameters of the GEV distribution defining its location, scale, and shape, respectively. The best estimate and its 95% confidence interval (CI) of the parameters (computed through a bootstrapping function implemented in the R package extRemes (Gilleland, 2019)) were estimated using the generalized maximum likelihood (GML) method, which extends the maximum likelihood (ML) method by adding a constraint that restricts the shape parameter to take values within a physically coherent range (Adlouni et al., 2007; Martins & Stedinger, 2000). We used the same prior distribution for the shape parameter as defined in (Martins & Stedinger, 2000). When the GML method could not calibrate properly on the

data ($\leq 1\%$ CONUS grid cells per data set member in all data set-RCP scenario combination), we used the L-moments method to estimate the parameters of the GEV distribution (Hosking, 1990).

For each grid cell and for each data set member, the $1/q$ -year ARI change factor (or climate signal) was defined as the ratio between the future precipitation volume and the historical volume for each ARI quantile, q and was estimated as

$$X_q = \frac{F^{-1}q; \mu_{fut}, \sigma_{fut}, \xi_{fut}}{F^{-1}q; \mu_{ctrl}, \sigma_{ctrl}, \xi_{ctrl}}, \quad (2)$$

where μ_{fut} , σ_{fut} , ξ_{fut} are the GEV parameters of the distribution fitted to the future period (2044–2099), μ_{ctrl} , σ_{ctrl} , ξ_{ctrl} are the GEV parameters of the distribution fitted to the historical (1950–2005) model period, and X_q the change factor computed for q equal to 0.5 (2-year), 0.2 (5-year), 0.1 (10-year), 0.04 (25-year), 0.02 (50-year), and 0.01 (100-year) ARI events.

For the BCCAv2 data set that included multiple downscaled realizations for the same GCM, each individual run change factors were averaged. Also, because the number of members in each RCP scenario are different, members that were not common in both RCP scenarios were discarded (see Tables S1 to S3). Across data sets, the set of members is different, which makes a comparison of the impact of downscaling method challenging. Although here we focus on the complete set of available members per data set to stay consistent with how these data sets are used in the literature, we repeated the analysis by using the CMIP5 GCM model (CanESM2) that is downscaled in all five data sets to assess the effect of using different GCMs in each data set.

Finally, the data set mean climate change signal for the 2-, 5-, 10-, 25-, 50-, and 100-year ARI was defined as the average of equally weighted member change factors in a data set. We specifically chose these ARIs to produce actionable results, given that the selected ARIs are used as standards for designing stormwater infrastructure in the United States (Lopez-Cantu & Samaras, 2018). To test the climate change signal's significance for each data set and ARI, we first constructed a series of 100 artificial rainfall volumes for both the future and the historical period, where each volume was estimated through a GEV distribution whose parameters were randomly selected (with replacement) from the GEV parameters of any data set member estimated in earlier steps. We then compare the future and historical artificial rainfall volumes and detect statistically significant (at the 0.01 level) differences between both periods using the Mann-Whitney rank test.

2.3. Analysis of Uncertainty Sources and their Contribution to the Overall Uncertainty in the Projected Changes in Rainfall Extremes

To understand how important potential differences between data sets are in the overall uncertainty context, we identified two other main sources that contribute to uncertainty in the projected change of rainfall extremes. We define the uncertainty due to data set selection (data set variance), $D(a)$, due to the GEV distribution fitting (distribution fit variance), and due to the emissions scenario (scenario variance) and contextualize these uncertainties in a similar way as Hawkins and Sutton (2009).

The data set variance component by ARI was estimated as

$$D(a) = \frac{1}{N_s} \sum_s \text{var}_d \left(\frac{1}{N_d} \sum_{m=1}^{N_d} x_{asdm} \right), \quad (3)$$

where $D(a)$ is the data set variance component in the a -year event change factors. x_{asdm} is the a -year event change factor in RCP scenario s projected by member m in data set d computed using the GEV best estimate parameters, and N_d is the total number members in data set d , and N_s is the number of scenarios (in our case two, RCP 4.5 and 8.5). $D(a)$ measures the a -year change factor variability across data sets.

The distribution fit variance was estimated using the change factors computed by all combinations of data set members, scenarios, and the lower and upper endpoints of the 95% CI of the estimated GEV parameters. The distribution fit variance component by ARI was estimated as

$$F(a) = \frac{1}{N_s} \sum_s \text{var}_d \left(\frac{\sum_{m=1}^{N_d} (x_{asdm}^U - x_{asdm}^L)}{N_d} \right), \quad (4)$$

where $F(a)$ is the distribution fit variance in the projected a -year event change factors. x_{asdm}^U is the a -year event change factor in scenario s computed using the GEV parameters of the 95% CI upper endpoint, and x_{asdm}^L , the a -year event change factor in scenario s projected by member m in data set d computed using the GEV parameters of the 95% CI lower endpoint. The difference $x_{asdm}^U - x_{asdm}^L$ corresponds to the change factor spread resulting from the uncertainty in the GEV parameters and depends on the confidence level chosen. N_d is the total number of members in data set d , and N_s is the total number of scenarios. $F(a)$ measures how much the a -year change factor varies due to the GEV fitting accuracy on average per member, data set, and scenario.

Finally, the scenario variance component by ARI was estimated as

$$S(a) = \text{var}_s \left(\frac{1}{N_e} \sum_d \frac{1}{N_d} \sum_{m=1}^{N_d} x_{asdm} \right), \quad (5)$$

where $S(a)$ is the scenario variance in the projected a -year event change factors. x_{asdm} is the a -year event change factor in RCP scenario s projected by member m in data set d , N_d is the total number members in data set d , and N_e is the total number of data sets in the ensemble (i.e., five different data sets).

Assuming that the three sources of uncertainty are independent, the total variance in the projected change factors can be written as

$$T(a) = D(a) + F(a) + S(a). \quad (6)$$

The contribution of each source of uncertainty can be found by dividing each by the total variance.

3. Results

3.1. Future Daily Rainfall Extremes Comparison Between Data Sets

The data sets' multimember means show increased extreme precipitation over nearly all CONUS. CONUS-averaged change factors generally increase with ARIs (see Figure 1a), indicating that the highest extreme events increase the most. Between data sets, the spread of the CONUS distribution of change factors is different, and, for each data set, the spread increases with ARIs. For example, the CONUS 5-year change factors from simulations included in the LOCA data set ranges from 0.98 to 1.29 under RCP8.5, while the 50-year change factors range from 0.95 to 1.53. Despite exhibiting the least average bias compared to observed extremes (see Figure S2 to S4), the MACA multimember mean change factors are considerably larger than those of the other data sets. Notably, BCCAv2 shows the lowest multimember mean change factors which are mostly constant across ARIs (Figure 1a and Figures S5b to S10b). We found similar patterns when analyzing the change factors of the downscaled output of the common member across data sets (GCM: CanESM2) (Figure S11). A visual comparison between the data sets change factors and the native-resolution CanESM2 change factors (Figure S11) shows that the data sets preserve the pattern of change signal to some extent, but the magnitude of the signal is lower in BCCAv2, while it is higher in MACA. Figures S12 and S13 show a quantitative comparison between the regionally averaged downscaled climate signals and the GCM climate signals for both RCP4.5 and RCP8.5 scenarios. For most of the United States, the averaged climate signal is close to the GCM signal for short ARIs, while it gradually diverts with increasing ARI. For all regions and large ARIs, the MACA downscaled CanESM2 change signal is considerably higher than the original GCM signal, while the BCCAv.2 signal is mostly lower. Climate signals based on LOCA and NA-CORDEX are similar to the original GCM change signal, with largest differences (i.e., the West region) possibly due to better orographic representation in the dynamically downscaled NA-CORDEX data sets.

Figure 1a also shows that the data sets agree that daily rainfall extremes intensify more under RCP8.5 than under RCP4.5 on continental scales, consistent with previous studies comparing both scenarios (Fix et al., 2018). However, the difference between RCP scenarios varies across data sets (see Figure 1b). There exist little differences between the RCP4.5 and 8.5 scenarios in the BCCAv2 data set, while there is a larger difference in the low-resolution NA-CORDEX data set.

In terms of the spatial structure of these changes, the data sets roughly agree on higher percent increases west of the continental divide and the Ohio River Basin and Northeast regions (see Figure S12 for a graphical representation of the regions' location) for ARIs between 2- to 10-year. In contrast, for larger ARIs (25- to 100-year), there is no clear spatial agreement due to an increase in noisiness, yet the signal is statistically

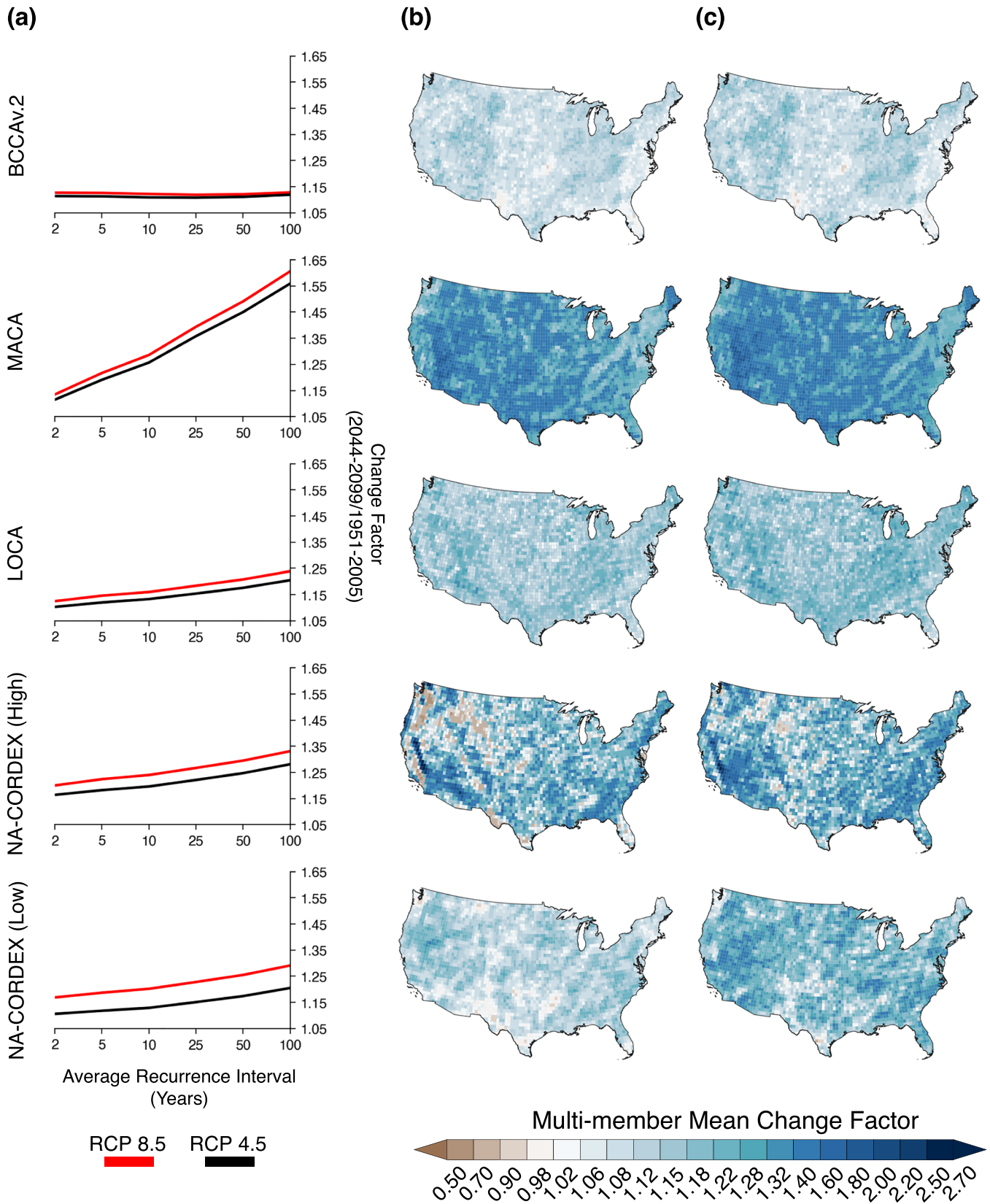


Figure 1. Future (2044–2099) extreme rainfall volumes are increasing in all data sets with larger increases for large ARI events and under RCP 8.5 compared to the past (1951–2005). Multimember mean extreme rainfall signal averaged across the United States for different ARIs (a), spatial distribution of the 25-year event change factors under RCP 4.5 (b), and RCP 8.5 (c). Each data set member was weighted equally in the multimember mean computation.

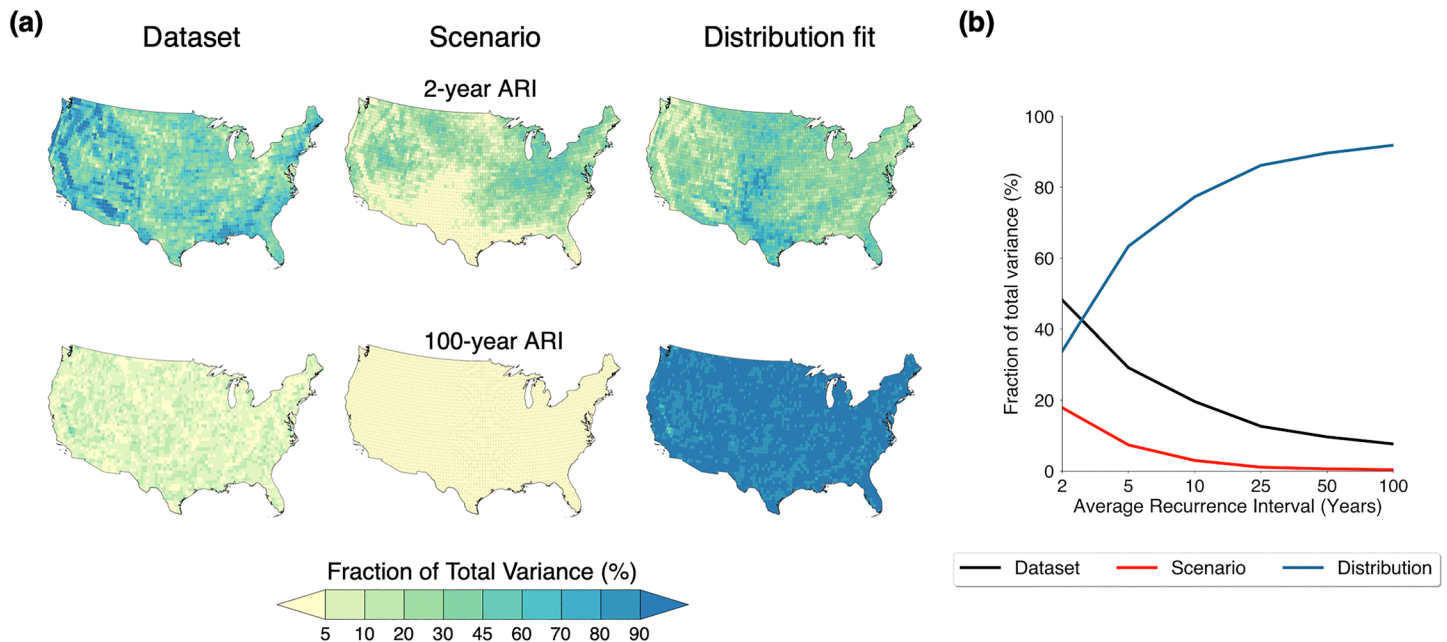


Figure 2. Data set selection uncertainties dominate future extreme rainfall changes for small ARI changes, while distribution fit uncertainties dominate larger ARI changes. Contribution of each uncertainty source to the total variance at the grid scale (a) and on average for the CONUS (b). High percentages indicate a larger difference in the projected change factors for a given source, compared to the difference from the other sources.

significant at the 0.01 level in all data sets for at least more than 40% of CONUS. For the 25-year ARI, the change signal is statistically significant in more than 50% of the grid cells in CONUS for RCP4.5 and more than 73% for RCP8.5 (see Table S5 for statistical significance of other ARIs).

In general, there are noticeable differences in change factors of the dynamically downscaled NA-CORDEX data sets compared to the statistically downscaled, particularly over orographic regions (See section S3).

Figures 1b and 1c show the magnitude and the spatial structure of the 25-year ARI multimember mean climate change signal which considerably differs across the five data sets (See Figures S5 to S10 for other ARIs). For the 25-year ARI on average for CONUS, the multimember mean change signal shows an increase of 11%, 15%, 36%, 22%, and 15% (BCCAv2, LOCA, MACA, NA-CORDEX high resolution, and NA-CORDEX low resolution, respectively) in the 25-year rainfall volume under RCP4.5 emissions scenario, and 12%, 18%, 39%, 27%, and 23% under RCP8.5 by the end of the century. We found negligible percent change differences between scenarios on average for CONUS for the 25-year event in BCCAv2 and between 3% and 7% for other data sets.

Additional information about the distribution of change factors in each data set (with respect to their individual members) can be obtained by examining the interquartile range (IQR), which is shown in Figures S14 to S19 for each ARI. For small ARIs, the IQR is generally small across CONUS and increases for larger ARIs. The NA-CORDEX high-resolution IQR is consistently large across ARIs over the Sierra Nevada mountains. The large IQR might be affected by the smaller number of downscaled simulations in the NA-CORDEX project at this resolution and the differences in RCM characteristics (i.e., dynamics, land surface model, etc).

3.2. Uncertainty in Future Rainfall Extremes Changes

In section 2.3, we distinguish three different sources of uncertainty in the projected change in rainfall extremes: data set selection, RCP scenario, and distribution uncertainty, and we assess the individual contribution of uncertainty for each source relative to the other sources of uncertainty. Figure 2 shows the contribution of the three sources of uncertainty to the total variance in the projected percent change in rainfall volume for each grid cell.

High percentages indicate a larger difference in the projected change factors for a given source, compared to the difference of the other sources. There are marked differences in the dominant uncertainty with increasing ARI across the United States. For the 2-year event, the data set selection uncertainty is largest, reaching

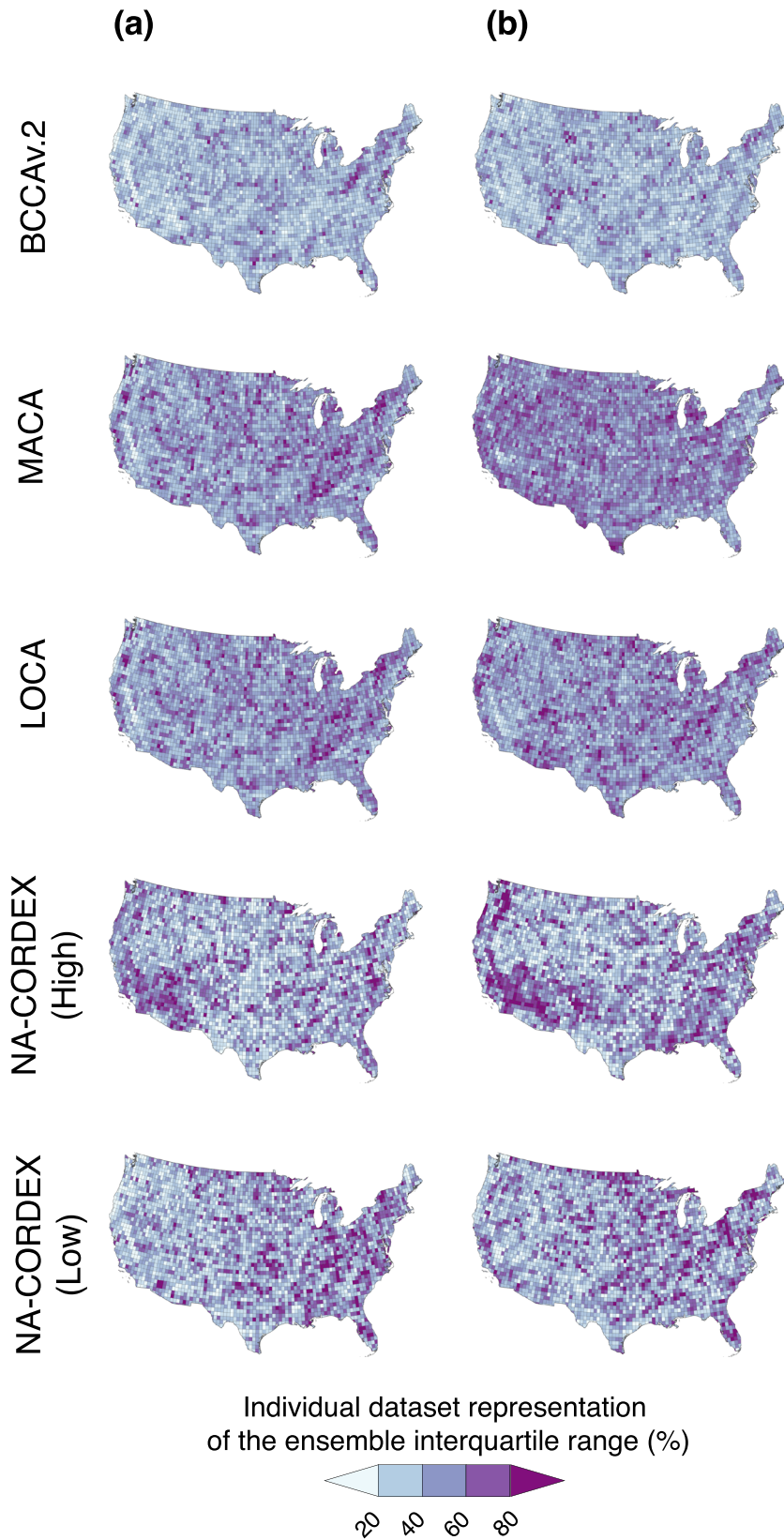


Figure 3. Percent covered by each individual data set IQR of the ensemble IQR defined by the largest 75th and smallest 25th percentiles of all data sets for the 25-year event under (a) RCP 4.5 and (b) RCP 8.5. Darker purple means that the data set IQR spans a large portion of the overall data set uncertainty range defined above.

very large percentages (~60%) in several grid cells. The data set uncertainty is especially high in coastal areas in the Northeast and the Southeast, as well as in the mountainous regions in the West (see Figure S12 for a graphical representation of the regions' location). For the 100-year event, the data set selection uncertainty is no longer prominent in comparison to the distribution fit uncertainty, which is now dominant (some cells reach about 70%). Larger contributions to the total variance from the scenario uncertainty occur at the 2-year than at the 100-year.

In most of the Ohio River Basin, the scenario uncertainty contribution is approximately between 40% and 50%, but for the rest of the United States, the percent is not higher than approximately 20%. On average (Figure 2b), we found that for small ARIs (<5 years), the data set selection uncertainty is largest followed by the distribution fit and scenario uncertainty. The contribution of uncertainty from the distribution fit increases with ARI and becomes dominant for large ARIs (>5 years). Scenario uncertainty is largest (~20%) for short ARIs and is negligible for longer ARIs. Note that the scenario uncertainty would increase by including lower emission scenarios in the analysis (e.g., RCP2.6). However, this is currently not possible due to a lack of downscaled data products over the CONUS.

Across data sets, the difference in the change factors computed using the GEV 95% CI upper- and lower-end parameters was similar. For any given data set and long ARIs, we found large variability in the projected change factors due to the GEV distribution fit, which is consistent with previous studies that found that the record length strongly affects the estimate of the GEV parameters (Papalexiou & Koutsoyiannis, 2013).

4. Discussion and Conclusions

We showed that the magnitude and spatial pattern of the climate change signal in U.S. daily rainfall extremes greatly differs across the most commonly used public downscaled climate model output data sets for the second half of this century. While the set of CMIP5 GCMs downscaled in each data set is different, we found that the differences between data sets are likely due to methodological choices within the downscaling framework rather than selected GCMs. We compared the climate change signal from the downscaled CanESM2 GCM simulations (common in all data sets), to the native resolution CanESM2 signal. In general, downscaling modifies the original GCM change signals with MACA simulations enhancing the original GCM trend for large ARIs (> 5 years). Conversely, the BCCAv.2 data lowers change signals. Previous studies have shown the original GCM trend can be modified by the downscaling method and/or bias correction steps due to improved surface representation and capability to solve physical processes at higher resolution in dynamical downscaling, while statistical downscaling can lead to an inflated signal artifact of the method because of correcting events that have occurred with very low frequency (Eum & Cannon, 2017). The magnitude of the difference in the projected change factors between data sets and RCP scenarios is in most cases non-negligible and might represent additional stress to infrastructure systems designed for historical conditions (Guentchev et al., 2016; Niemczynowicz, 1989).

When modelers and stakeholders are evaluating critical areas and decisions in resilience assessments, it is prudent to acknowledge the differences between data sets in the projected changes in daily rainfall extremes, their multimember distribution (i.e., the distribution of individual members' change factors), and the implications of using each data set. The varying magnitude of uncertainty sources with ARIs also emphasizes the need to properly manage and propagate uncertainties in climate change impacts assessments and decision making support (Arnbjerg-Nielsen et al., 2015).

Selecting a specific data set implies constraining the range of future change factors to a range that might not be representative of the full spread of uncertainties. Figure 3 shows how much of the ensemble IQR range is represented by the IQR of each individual data set's members at the grid scale level for the 25-year ARI event (see Figures S20 to S25 for other ARIs). A darker purple color means that the IQR from simulations within one data set is similar to the projected IQR from the ensemble and vice versa. Some data set representation percentages stand out in particular regions; for example, NA-CORDEX high-resolution downscaled output provides a better representation of the range of change factors overall (for all ARIs) at topographically complex regions. If computational resources are limited and the decision maker needs to select projections from one data set, selecting the data set with the darkest purple color for a given grid cell could be an option to at least partially represent signals from other data sets. However, this might not be the most robust decision, as Figure 3 also shows large spatial noise. Additionally, since the multimember mean is different between

data sets, the change factors of the data set with the largest ensemble IQR representation might be considerably larger or smaller than other data sets' change factors. In the case of the former, it could be considered by the decision maker as an added layer of safety or robustness; however, it could be problematic in the latter case. A modeler or decision maker can interpret Figures 1 and 3 to understand the consequences of selecting a data set based on different metrics. For instance, selecting the MACA data set because of its low average bias will lead to large climate change signals and, depending on the location of interest, might not be representative of other data sets since there is large spatial noise (see Figure 3).

In summary, selecting a specific downscaled climate data set for decision making and neglecting others can result in potentially omitting considerable uncertainty in the range of rainfall extremes that inform resilience decisions. While the contribution to the overall uncertainty in rainfall extremes from data set, distribution fit and emission scenario vary by ARI, given the differences we found, using more than a single data set and considering distribution fit uncertainty is preferable for robust decision making. However, one limitation in these conclusions lies in the comparability of the downscaling method given that only a single GCM output (CanESM2) was downscaled across all data sets. The differences found in data set statistics (mean and IQR) highlight the need for conducting future downscaling efforts on a controlled set of GCM for ease of comparability. Additionally, we show that there is also a need for longer records to facilitate more robust estimates of very large ARIs. Large ensemble data sets could be used to provide more robust estimates for very large ARIs (e.g., 100-year return levels) from model data to reduce the distribution fit uncertainty (Tandon et al., 2018). Another promising approach for reducing data set uncertainties in future extreme precipitation estimates are kilometer-scale climate models that better simulate extreme precipitation than state-of-the-art coarser-resolution models (Prein et al., 2015).

In case the computational resources exist to perform impacts assessments under many scenarios based on the change factors from all data sets, our data can facilitate the process of generating future scenarios for analyses and we provide a grid-cell level data set across CONUS for the climate change signal in each of the five data sets (Lopez-Cantu et al., 2020). For specific applications, decision makers can use our data and results, coupled with their expertise and risk preferences, to inform resilience assessments and decisions.

Acknowledgments

This research was partially supported by the National Science Foundation (NSF Collaborative Award Number CMMI 1635638/1635686) and the UCAR Next Generation Fellowship. UCAR and NCAR are sponsored by the National Science Foundation. Student support was also provided by Consejo Nacional de Ciencia y Tecnología (CONACYT).

We acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5. We are grateful to the climate modeling groups that produced and made available their downscaled climate projections listed in Table S1. We also acknowledge the U.S. Department of Defense ESTCP for its support of the NA-CORDEX data archive. Computer resources were provided by the Computational and Information Systems Laboratory, 2017. Cheyenne: HPE/SGI ICE XA System (NCAR Community Computing). Boulder, CO: National Center for Atmospheric Research. doi:10.5065/D6RX99HX.

References

- Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32(5), 772–780. <https://doi.org/10.1002/joc.2312>
- Adlouni, S. E., Ouara, T. B. M. J., Zhang, X., Roy, R., & Bobée, B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43, W03410. <https://doi.org/10.1029/2005WR004545>
- Alamdari, N., Sample, D. J., Steinberg, P., Ross, A. C., & Easton, Z. M. (2017). Assessing the effects of climate change on water quantity and quality in an urban watershed using a calibrated stormwater model. *Water*, 9(7), 464. <https://doi.org/10.3390/w9070464>
- Arnbjerg-Nielsen, K., Leonardsen, L., & Madsen, H. (2015). Evaluating adaptation options for urban flooding based on new high-end emission scenario regional climate model simulations. *Climate Research*, 64(1), 73–84.
- Bracken, C. (2016). *Downscaled CMIP3 and CMIP5 Climate Projections—Addendum*. Washington, DC: Reclamation.
- Cannon, A. J. (2018). Multivariate quantile mapping bias correction: An N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, 50(1), 31–49. <https://doi.org/10.1007/s00382-017-3580-6>
- Coles, S. (2001). *An introduction to statistical modeling of extreme values* (2001 edition). London; New York: Springer.
- Cook, L. M., Anderson, C. J., & Samaras, C. (2017). Framework for incorporating downscaled climate output into existing engineering methods: Application to precipitation frequency curves. *Journal of Infrastructure Systems*, 23(4), 04017027. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000382](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000382)
- Cook, L. M., McGinnis, S., & Samaras, C. (2020). The effect of modeling choices on updating intensity-duration-frequency curves and stormwater infrastructure designs for climate change. *Climatic Change*, 158, 289–308. <https://doi.org/10.1007/s10584-019-02649-6>
- DeGaetano, A. T. (2009). Time-dependent changes in extreme-precipitation return-period amounts in the continental United States. *Journal of Applied Meteorology and Climatology*, 48(10), 2086–2099. <https://doi.org/10.1175/2009JAMC2179.1>
- Dixon, K. W., Lanzante, J. R., Nath, M. J., Hayhoe, K., Stoner, A., Radhakrishnan, A., Balaji, V., & Gaitán, C. F. (2016). Evaluating the stationarity assumption in statistically downscaled climate projections: Is past performance an indicator of future results? *Climatic Change*, 135(3), 395–408. <https://doi.org/10.1007/s10584-016-1598-0>
- Eum, H.-I., & Cannon, A. J. (2017). Intercomparison of projected changes in climate extremes for South Korea: Application of trend preserving statistical downscaling methods to the CMIP5 ensemble. *International Journal of Climatology*, 37(8), 3381–3397. <https://doi.org/10.1002/joc.4924>
- Fix, M. J., Cooley, D., Sain, S. R., & Tebaldi, C. (2018). A comparison of U.S. precipitation extremes under RCP8.5 and RCP4.5 with an application of pattern scaling. *Climatic Change*, 146(3), 335–347. <https://doi.org/10.1007/s10584-016-1656-7>
- Forsee, W. J., & Ahmad, S. (2011). Evaluating Urban Storm-Water Infrastructure Design in Response to Projected Climate Change. *Journal of Hydrologic Engineering*, 16(11), 865–873. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000383](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000383)
- Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12), 1547–1578. <https://doi.org/10.1002/joc.1556>
- Gelda, R. K., Mukundan, R., Owens, E. M., & Abatzoglou, J. T. (2019). A Practical Approach to Developing Climate Change Scenarios for Water Quality Models. *Journal of Hydrometeorology*, 20, 1197–1211. <https://doi.org/10.1175/JHM-D-18-0213.1>
- Gilleland, E. (2019). extRemes: Extreme Value Analysis.

- Guentchev, G. S., Rood, R. B., Ammann, C. M., Barsugli, J. J., Ebi, K., Berrocal, V., et al. (2016). Evaluating the appropriateness of downscaled climate information for projecting risks of Salmonella. *International Journal of Environmental Research and Public Health*, *13*(3), 267. <https://doi.org/10.3390/ijerph13030267>
- Hallegatte, S. (2009). Strategies to adapt to an uncertain climate change. *Global Environmental Change*, *19*(2), 240–247. <https://doi.org/10.1016/j.gloenvcha.2008.12.003>
- Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. <https://doi.org/10.1175/2009BAMS2607.1>
- Hidalgo, H. G., Dettinger, M. D., & Cayan, D. R. (2008). Downscaling with constructed analogues: Daily precipitation and temperature fields over the United States. California Energy Commission, CEC-500-2007-123.
- Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X.-W., Wolter, K., & Cheng, L. (2016). Characterizing recent trends in U.S. heavy precipitation. *Journal of Climate*, *29*(7), 2313–2332. <https://doi.org/10.1175/JCLI-D-15-0441.1>
- Hosking, J. R. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(1), 105–124.
- IPCC (2014). *Climate change 2014: Impacts, adaptation, and vulnerability. Summaries, frequently asked questions, and cross-chapter boxes. A contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* Edited by Field, C. B., Barros, V. R., Dokken, D. J., Mach, K. J., Mastrandrea, M. D., et al. Geneva, Switzerland: World Meteorological Organization.
- Jones, P. (1999). First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, *127*(9), 2204–2210.
- Kermanshah, A., Derrible, S., & Berkelhammer, M. (2017). Using climate models to estimate urban vulnerability to flash floods. *Journal of Applied Meteorology and Climatology*, *56*(9), 2637–2650.
- Kuo, C.-C., Gan, T. Y., & Gizaw, M. (2015). Potential impact of climate change on intensity duration frequency curves of central Alberta. *Climatic Change*, *130*(2), 115–129. <https://doi.org/10.1007/s10584-015-1347-9>
- Lopez-Cantu, T., Prein, A. F., & Samaras, C. (2020). Change factors for the 2- to 100-year daily (24-hour) extreme rainfall storms for the Continental United States from downscaled climate projections. <https://doi.org/10.1184/R1/12148932>
- Lopez-Cantu, T., & Samaras, C. (2018). Temporal and spatial evaluation of stormwater engineering standards reveals risks and priorities across the United States. *Environmental Research Letters*, *13*(7), 074,006. <https://doi.org/10.1088/1748-9326/aac696>
- Maraun, D. (2016). Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports*, *2*(4), 211–220. <https://doi.org/10.1007/s40641-016-0050-x>
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, *48*, RG3003. <https://doi.org/10.1029/2009RG000314>
- Martinich, J., & Crimmins, A. (2019). Climate damages and adaptation potential across diverse sectors of the United States. *Nature Climate Change*, *9*(5), 397–404. <https://doi.org/10.1038/s41558-019-0444-6>
- Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, *36*(3), 737–744. <https://doi.org/10.1029/1999WR900330>
- Mearns, L. O., Bukovsky, M., Pryor, S. C., & Magaña, V. (2014). Downscaling of climate information. In G. Ohring (Ed.), *Climate Change in North America* (pp. 201–250). Regional Climate Studies. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-03768-45>
- Mearns, L. O., Sain, S., Leung, L. R., Bukovsky, M. S., McGinnis, S., Biner, S., et al. (2013). Climate change projections of the North American Regional Climate Change Assessment Program (NARCCAP). *Climatic Change*, *120*(4), 965–975. <https://doi.org/10.1007/s10584-013-0831-3>
- Mullan, M., Danielson, L., Lasgargues, B., Chrisna Morgado, N., & Perry, E. (2018). Climate-resilient infrastructure. OECD, 14.
- Niemczynowicz, J. (1989). Impact of the greenhouse effect on sewerage systems—Lund case study. *Hydrological Sciences Journal*, *34*(6), 651–666. <https://doi.org/10.1080/02626668909491373>
- Onof, C., & Arnbjerg-Nielsen, K. (2009). Quantification of anticipated future changes in high resolution design rainfall for urban areas. *Atmospheric Research*, *92*(3), 350–363. <https://doi.org/10.1016/j.atmosres.2009.01.014>
- Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, *49*, 187–201. <https://doi.org/10.1029/2012WR012557>
- Pierce, D. W., Cayan, D. R., & Thrasher, B. L. (2014). Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), 2558–2585. <https://doi.org/10.1175/JHM-D-14-0082.1>
- Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., et al. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*, 323–361. <https://doi.org/10.1002/2014RG000475>
- Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, *7*(1), 48–52. <https://doi.org/10.1038/nclimate3168>
- Sunyer, M. A., Madsen, H., & Ang, P. H. (2012). A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change. *Atmospheric Research*, *103*, 119–128. <https://doi.org/10.1016/j.atmosres.2011.06.011>
- Tandon, N. F., Zhang, X., & Sobel, A. H. (2018). Understanding the dynamics of future changes in extreme precipitation intensity. *Geophysical Research Letters*, *45*, 2870–2878. <https://doi.org/10.1002/2017GL076361>
- Wright, D. B., Bosma, C. D., & Lopez-Cantu, T. (2019). U.S. hydrologic design standards insufficient due to large increases in frequency of rainfall extremes. *Geophysical Research Letters*, *46*, 8144–8153. <https://doi.org/10.1029/2019GL083235>
- Wu, S., Markus, M., Lorenz, D., Angel, J. R., & Grady, K. (2019). A comparative analysis of the historical accuracy of the point precipitation frequency estimates of four data sets and their projections for the Northeastern United States. *Water*, *11*(6), 1279. <https://doi.org/10.3390/w11061279>