How much information is required to well constrain local estimates of future precipitation extremes?

Chao Li, Francis W. Zwiers, Xuebin Zhang, & Guilong Li

2019

Pacific Climate Impacts Consortium (PCIC)

PCIC Publications

Original citation:

Li, C., Zwiers, F. W., Zhang, X., & Li, G. (2019). How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, *7*(1), 11–24. https://doi.org/10.1029/2018EF001001

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Earth's Future

RESEARCH ARTICLE
10.1029/2018EF001001

Key Points:
- Estimates of the scaling of local extreme precipitation with temperature based on local precipitation observations are highly uncertain
- Records with lengths that are many multiples of the length of available observations are needed to well constrain the scaling relation
- Well-constrained scaling relations can reliably estimate climate modeled extreme precipitation changes over almost all of North America

Supporting Information:
- Supporting Information S1

Correspondence to:
X. Zhang,
xuebin.zhang@canada.ca

Citation:
Li, C., Zwiers, F., Zhang, X., & Li, G. (2019). How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, *7*, 11–24. https://doi.org/10.1029/2018EF001001

Received 26 JUL 2018
Accepted 3 DEC 2018
Accepted article online 13 DEC 2018
Published online 8 JAN 2019

How Much Information Is Required to Well Constrain Local Estimates of Future Precipitation Extremes?

Chao Li[1,2,3], Francis Zwiers[1], Xuebin Zhang[4], and Guilong Li[4]

[1]Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada, [2]Now at Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai, China, [3]Now at School of Geographic Sciences, East China Normal University, Shanghai, China, [4]Climate Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada

Abstract Global warming is expected to increase the amount of atmospheric moisture, resulting in heavier extreme precipitation. Various studies have used the historical relationship between extreme precipitation and temperature (*temperature scaling*) to provide guidance about precipitation extremes in a future warmer climate. Here we assess how much information is required to robustly identify temperature scaling relationships, and whether these relationships are equally effective at different times in the future in estimating precipitation extremes everywhere across North America. Using a large ensemble of 35 North American regional climate simulations of the period 1951–2100, we show that individual climate simulations of length comparable to that of typical instrumental records are unable to constrain temperature scaling relationships well enough to reliably estimate future extremes of local precipitation accumulation for hourly to daily durations in the model's climate. Hence, temperature scaling relationships estimated from the limited historical observations are unlikely to be able to provide reliable guidance for future adaptation planning at local spatial scales. In contrast, well-constrained temperature scaling relations based on multiple regional climate simulations do provide a feasible basis for accurately projecting precipitation extremes of hourly to daily durations in different future periods over more than 90% of the North American land area.

Plain Language Summary Global warming is expected to increase the amount of atmospheric moisture, resulting in heavier extreme precipitation. Various studies have used the historical relationship between extreme precipitation and temperature (*temperature scaling*) to provide guidance about precipitation extremes in a future warmer climate. Temperature scaling is useful for this purpose only when it can be robustly identified. Using a large ensemble of 35 North American regional climate simulations of the period 1951–2100, we show that individual climate simulations of length comparable to that of typical instrumental records are unable to identify temperature scaling relationships robustly enough to reliably estimate future extremes of local precipitation accumulation for hourly to daily durations in the model's climate. Hence, temperature scaling relationships estimated from the limited historical observations are unlikely to be able to provide reliable guidance for future adaptation planning at local spatial scales. This also has broader implications for how we account for nonstationarity more generally in historical observations. In contrast, temperature scaling relations based on multiple regional climate simulations do provide a feasible basis for accurately projecting precipitation extremes of hourly to daily durations over more than 90% of the North American land area.

1. Introduction

An integral component of building and bridge construction codes, which set standards for infrastructure design, is guidance on the expected magnitude of local precipitation extremes. The National Building Code of Canada (NCBC, 2005), for example, includes estimates of the 10-year return level for 15-min rainfall accumulation and 50-year return level for 1-day rainfall accumulation for hundreds of communities across Canada. These estimates, which are based on station data, have historically been produced assuming a stationary climate. There is, however, an accumulating body of evidence indicating that the hydrological cycle is changing in response to anthropogenic global warming. Increases in extreme precipitation have been observed (e.g., Donat et al., 2016; Westra et al., 2013) and can be attributed to human influence at large spatial scales (e.g., Fischer & Knutti, 2015; Min et al., 2011; Zhang et al., 2013). The intensity of extreme precipitation is projected to continue to increase in the future (e.g., Kharin et al., 2013, 2018).

Such changes in the climate have led to the recognition that *stationarity is dead* (Milly et al., 2008, 2015). A major challenge the engineering community is facing is therefore the need to reliably project future design values for extreme precipitation. Nevertheless, nonstationarity remains difficult to model (e.g., Serinaldi & Kilsby, 2015) and to detect (e.g., Westra et al., 2013) at local spatial scales due to the dominance of natural variability. As a result, observations or individual climate simulations of length comparable to that of typical observed records (being generally less than 65 years in length) are unlikely to sufficiently constrain nonstationary extreme value models to reliably estimate long-period return levels for the future.

One important ingredient for the occurrence of extreme precipitation is the availability of atmospheric moisture. As the atmosphere warms, its water holding capacity, or saturation water vapor pressure, increases exponentially with temperature rise as described by the Clausius-Clapeyron (CC) relation. As a result, atmospheric moisture (precipitable water) generally increases with warming. Assuming all other factors, such as the circulations driving moisture convergence and precipitation efficiency during extreme events, are equal, extreme precipitation should therefore also increase with warming. Over global land areas, xextreme precipitation as represented by annual maximum 1-day precipitation has increased with warming at a rate consistent with the CC relation (Westra et al., 2013; Zhang et al., 2013). This is also the case for future climates projected with global climate models (e.g., Kharin et al., 2013; Westra et al., 2014). At the continental scale, Prein et al. (2017) reported that changes in subdaily extreme precipitation also scale with temperature change at roughly the CC rate over North America in a convection permitting regional climate model. Thus, exploiting scaling relationships between temperature and extreme precipitation may offer potential solutions to projecting future extreme precipitation as model projections for future temperature are relatively robust. At regional and local spatial scales, however, the estimation of precipitation scaling relations from individual records is highly uncertain (Zhang et al., 2017). Pooling information across space can reduce uncertainty in estimates of the return levels of extreme precipitation (e.g., Hanel et al., 2009; Kendon et al., 2008; Sun & Lall, 2015). For example, the index flood approach (e.g., Hosking & Wallis, 1997) has long been used in flood frequency analysis and has also been extended to the nonstationary world. Combining spatial pooling with temperature scaling may offer a promising approach to projecting future precipitation extremes.

The main objective of this study is to investigate the amount of information required to well constrain local temperature scaling relationships for extreme precipitation in North America. We investigate this question in an idealized world with a large ensemble of 35 North American regional climate simulations for the period 1951–2100 that has recently become available. We consider both local analyses and spatially pooled analyses using the index flood approach. We find that records that are equivalent in length of many multiples of historically observed records are needed to robustly identify temperature scaling relationships, even when using the index flood approach. Data and methods are described in section 2. Results are presented in section 3, followed by discussion in section 4 and conclusions in section 5.

2. Data and Methods

2.1. Regional Climate Model Simulations

We use hourly precipitation from an ensemble of 35 simulations for the period 1951–2100 covering North America. The simulations are conducted with the Canadian Regional Climate Model (CanRCM4) at a horizontal resolution of ~50 km (Scinocca et al., 2016). These simulations are driven by a corresponding large initial-condition ensemble of Canadian Earth System Model (CanESM2) simulations. The CanESM2 simulations use the historical *all* forcing (solar and volcanic forcing, greenhouse gases, aerosols, ozone, and land use) prescription for the period ending in 2005 and the Representative Concentration Pathway 8.5 emissions scenario for 2006–2100, which has total radiative forcing in 2100 relative to 1750 of approximately 8.5 W/m^2 (van Vuuren et al., 2011). The differences among ensemble members are due to internal variability. As such, to the extent that the CanESM2/CanRCM4 combination of models captures the relevant physical processes for the historical period, the 35 simulations can be considered as 35 plausible realizations of the real world (Deser et al., 2012) and thus at a given location, yield 35 times as much data as an observational record of comparable length. The historical simulations of CanRCM4 have been found to reproduce well the observed spatial patterns, seasonal variations, and the probability distributions of precipitation extremes over most parts of North America, with marked negative biases during May to November, notably in the southeastern United States (Whan & Zwiers, 2016a, 2016b).

2.2. Data Processing

The CanRCM4 large ensemble model output archive includes hourly precipitation accumulation for the entire length of each simulation. We aggregate hourly precipitation for individual grid cells into accumulations of nonoverlapping 6-, 12-, and 24-hr periods. Annual maximum 1-, 6-, 12-, and 24-hr precipitation accumulations (expressed in mm/hr) are then retrieved, leading to 35 sets of annual maxima for each accumulation duration. We use the annual global mean temperature from the 35 driving CanESM2 simulations as a covariate in nonstationary extreme value models of the time series of annual precipitation accumulation maxima. This follows from the physical consideration that the expected response of extreme precipitation may be constrained by the availability of atmospheric moisture (Allen & Ingram, 2002; Trenberth et al., 2003), for which global mean temperature can be a proxy (Sherwood et al., 2010). Similar choices of covariates have been made in other studies (e.g., Hanel et al., 2009; Kharin et al., 2018; Westra et al., 2013).

2.3. Temperature Scaling Estimation

We define $Y(s, t)$ to be a random variable representing annual maximum precipitation of a particular accumulation duration at grid cell $s \in \mathrm{S} := \{1, 2, ..., n_s\}$ in year $t \in \mathrm{T} := \{1, 2, ..., n_t\}$. Statistical theory suggests, via the Fischer-Tippett Theorem, that the generalized extreme value (GEV) distribution should be a good choice for approximating the distribution of $Y(s, t)$ (e.g., Coles, 2001). It has a probability distribution function of the form

$$F_{s,t}(y) = \begin{cases} \exp\left\{-[1 + \xi(s,t)(y - \mu(s,t))/\sigma(s,t)]^{-1/\xi(s,t)}\right\}, & \xi(s,t) \neq 0 \\ \exp\{-\exp[-(y - \mu(s,t))/\sigma(s,t)]\}, & \xi(s,t) = 0 \end{cases}$$

with $\mu(s, t) \in \mathrm{R}$, $\sigma(s, t) > 0$, and $\xi(s, t) \in \mathrm{R}$ being the location, scale, and shape parameters, respectively.

To reduce sampling uncertainty in the GEV parameter estimates at a particular grid cell s, we pool data from neighboring grid cells by means of the *index flood* method. Using the location parameter $\mu(s, t)$ as the index flood, if $Y(s, t)$ follows GEV($\mu(s, t)$, $\sigma(s, t)$, $\xi(s, t)$), then the scaled variable $Y(s, t)/\mu(s, t)$ follows GEV(1, $\gamma(s, t)$, $\xi(s, t)$), with the dispersion coefficient $\gamma(s, t) = \sigma(s, t)/\mu(s, t)$. The index flood method assumes that for a given location s, $\gamma(s, t)$ and $\xi(s, t)$ stay constant across grid cells within a neighborhood U_s of s; that is, $\gamma(s', t) = \gamma(t)$ and $\xi(s', t) = \xi(t)$ for $s' \in \mathrm{U}_s$, $\mathrm{U}_s \subset \mathrm{S}$. It then follows that for points in U_s the distribution of $Y(s', t)$ can be reparameterized as GEV($\mu(s', t)$, $\mu(s', t)\gamma(t)$, $\xi(t)$).

We relate extreme precipitation with temperature by assuming that $\mu(s', t) = \mu_0(s')\exp[\beta_\mu x(t)]$, $\gamma(t) = \gamma_0\exp[\beta_\gamma x(t)]$, and $\xi(t) = \xi_0 + \beta_\xi x(t)$, where $x(t)$ is the annual global mean temperature anomaly in year t relative to the entire analysis period (e.g., 1951–2100), and $\mu_0(s')$, γ_0, ξ_0, β_μ, β_γ, and β_ξ are unknown parameters to be estimated. We use the method of maximum likelihood to estimate these parameters. We chose a complex model with all GEV parameters being dependent on temperature rather than a simpler model with only location, or location and scale being dependent on temperature because overall, the complex model best represents differences in the response to temperature that are seen for different percentiles of precipitation extremes.

The choice of exponential dependence of the location parameter $\mu(s', t)$ on temperature is motivated by the CC relation and has been used in previous studies in the context of extreme precipitation event attribution (e.g., Eden et al., 2016; van Oldenborgh et al., 2015), while the choice of exponential dependence of the dispersion coefficient $\gamma(t)$ on temperature ensures that the dispersion coefficient remains positive. The trend parameters β_μ and β_γ are common to all grid cells in U_s and thus reflect the average response of extreme precipitation to temperature over the pooled grid cells; the data set enlarged by spatial pooling will somewhat reduce uncertainty in the estimates of these trend parameters caused by internal variability and hence will reduce uncertainty in the estimates of temperature scaling.

The intensity of extreme precipitation at a given extreme level (e.g., 99th percentile) scales with temperature change ΔT approximately as $(1 + \hat{r})^{\Delta T}$, where $\hat{r}$ is the estimate of temperature scaling rate for extreme precipitation at that extreme level and can be estimated from the fitted GEV model (see section S1 in the supporting information). We consider extreme precipitation intensity at the 50th and 99th percentiles of the distribution of annual precipitation accumulation maxima, which correspond respectively to the intensity

of precipitation during a moderate 2-year event and a much more extreme 100-year event. The above exponential scaling relation is appropriate for regions where the expected response of extreme precipitation to external forcing is dominantly thermodynamic (e.g., Trenberth et al., 2003; Zhang et al., 2017). Other factors that affect extreme precipitation such as changes in circulation can affect the $\widehat{r}$ value; for example, changes in circulation may result in a reduction in extreme precipitation, and thus a negative scaling rate, or they may enhance extreme precipitation through increased moisture convergence resulting in a super CC scaling rate.

2.4. Spatial Data Pooling

A key aspect of regional frequency analysis is the identification of appropriate homogeneous regions for spatial data pooling. Automated identification approaches are subject to uncertainty from sampling variability and the possibility that statistical procedures for the detection of spatial inhomogeneity may not be very powerful. Thus, we opt to simply use a series of nested neighborhoods of different sizes centered on the grid cell of interest, with regular, spatially isotropic sampling of grid cells within each neighborhood. We consider four configurations for pooling data spatially (Figure S1). They are (1) pooling 5 selected grid cells within a 3×3 grid cell region, (2) pooling all 9 grid cells within a 3×3 region, (3) pooling 9 selected grid cells within a 5×5 grid cell region, and (4) pooling 13 selected grid cells within a 7×7 grid cell region. We focus on land as this is where the impacts of extreme precipitation will be felt most. Model ocean grid cells are excluded from pooling when the land-sea boundary intersects the pooling region. We select a subset of grid cells from the 5×5 and 7×7 regions in order to make computations manageable.

2.5. Evaluation of the Temperature Scaling

We use two metrics, one that measures the uncertainty of a temperature scaling estimate $\widehat{r}$ and a second that evaluates how well temperature scaling projects future precipitation extremes. Uncertainty is measured by the *scaling rate per standard error* (SRSE), which is defined as the ratio between the best estimate of the scaling rate and its standard error. When the 35 simulations are used individually to estimate scaling rates, the best estimate is the mean of the 35 values, and the standard error of the best estimate is approximated by the standard deviation of the 35 values scaled by $1/\sqrt{35}$ (as the ensemble members are independent). When the best estimate is based on multiple simulations, the standard error for that estimate is adjusted accordingly. Assuming that the best estimate is approximately Gaussian distributed with standard deviation equal to the estimated standard error, we consider a scaling rate to be (1) *unconstrainable* when $||\text{SRSE}|| \leq 2$, as there is then no detectable response of extreme precipitation to temperature change at approximately the 5% significance level; (2) *constrainable* when $2 < ||\text{SRSE}|| \leq 5$, which implies that a nonzero response is detectable at the 5% significance level, but the response magnitude is relatively uncertain; and (3) *robustly constrained* when $||\text{SRSE}|| > 5$, suggesting a tightly constrained estimate of temperature scaling rate with a 95% confidence interval that is narrower than $\widehat{r} \pm 0.5\widehat{r}$. A finding of *unconstrainable* scaling can occur either because the link between changes in extreme precipitation and temperature is relatively weak compared to natural variability or because there is no change in extreme precipitation at all. This classification, though somewhat subjective, reflects the strength of climate change signals relative to natural variability similarly to the commonly used signal-to-noise ratio.

To evaluate the performance of temperature scaling in projecting future precipitation extremes, we compare the root-mean-square error (RMSE) for extreme precipitation changes estimated by the scaling relation with the RMSE of changes calculated directly from climate model projections. For a given percentile (e.g., 50th or 99th), 35 estimates of the relative change in that percentile between the historical period 1951–2015 and a future 30-year period (e.g., 2031–2060 or 2071–2100) are estimated based on the temperature scaling relation or calculated directly from climate model projections. The RMSE difference between the 35 estimates of the relative change and a value that we take to be the *true* relative change (described below) is computed, and finally, a ratio comparing the RMSE for the temperature-scaling-based estimates (RMSE_S) with the RMSE for the direct estimates (RMSE_D) is calculated.

The temperature-scaling-based estimates of the relative change in a given percentile are calculated as $(1+\widehat{r})^{\Delta Ti}$, $i = 1, 2, ..., 35$, where $\widehat{r}$ is the scaling rate best estimate, that is, the ensemble mean of the 35 scaling rates estimated by regional analysis with the largest spatial pooling considered and using the full length of the 1951–2100 regional model simulations. The temperatures ΔT_i ($i = 1, 2, ..., 35$) are the changes in annual global mean temperature in future periods relative to the historical period projected by each of the 35 driving CanESM2 global

model projections. For the direct estimates of the relative change, 35 estimates of the given percentile are first computed individually from the 35 CanRCM4 simulations for historical and future periods. At each grid cell, data are pooled for calculating the percentile estimates from neighboring grid cells corresponding to the largest spatial pooling considered for the regional GEV analysis. Estimates of the relative change between the historical period and future periods are calculated from each of the 35 pairs of the percentile estimates as the ratio between the future and historical period percentiles.

Calculation of both RMSEs requires reference to a true value for the relative change in the percentile between historical and future periods. We obtain this value by deriving the percentiles in question from the largest sample of annual maxima that can be assembled from the 35-member ensemble by pooling across ensemble members and neighborhoods (e.g., see Figure S2 for the true relative changes in the 50th and 99th percentiles of annual maximum 12-hr precipitation). At each grid cell, data are pooled from neighborhoods corresponding to the largest spatial pooling considered for the regional GEV analysis. We consider the resulting relative change as the true value in the sense that it is based on very large samples of annual maxima at each location (samples sizes are $35 \times 65 \times 13 = 29{,}575$ and $35 \times 30 \times 13 = 13{,}650$ for the historical and future periods, respectively).

Squared RMSE (i.e., mean-square error) can be decomposed into a squared bias component and a variance component. Since the direct estimates of the relative change should be essentially unbiased, RMSE_D is effectively an estimate of the standard deviation of the direct estimates. In contrast, RMSE_S reflects the combination of the standard deviation of the scaling-based estimates and their bias (i.e., the mean difference between the 35 scaling-based estimates and the *true value* of the relative change). An RMSE_S/RMSE_D ratio that is less than 1 at the 5% significance level (evaluated by a bootstrap approach) indicates that there is a significant advantage for a well-constrained temperature scaling relation in projecting future precipitation extremes over direct estimation from climate model projections. In this case, the variance reduction that is enabled by temperature scaling overcomes any bias that it might have introduced. A ratio that is significantly greater than 1 is an indication of a large bias in the scaling-based estimates that inflates their RMSE.

3. Results

3.1. Temperature Scaling Inferred From Limited Historical Records Can Be Highly Uncertain

Figure 1 shows the estimated temperature scaling rates for the 50th percentile of annual maximum 12-hr precipitation obtained by at-site analysis (involving one grid cell) from two individual simulations during a historical period 1951–2015 (Figures 1a and 1b) and a future period 2036–2100 (Figures 1c and 1d). We chose these two simulations as they have the smallest and largest spatial averages of estimated temperature scaling rates during the historical period over North America as a whole among the 35 simulations.

The spatial patterns of the scaling rates estimated from individual simulations are rather fragmented, with neighboring grid cells that often have very different scaling rates, sometimes even with opposite signs (Figure 1a or 1b). Spatial patterns among individual simulations are hardly consistent, particularly during the historical period when the external forcing is relatively weak (Figure 1a or 1b). Estimated scaling rates in different periods of the same simulation can also have different signs (Figure 1a vs. 1c and Figure 1b vs. 1d). For example, negative estimates of scaling rates for annual maximum 12-hr precipitation dominate over central and southwestern North America in the historical 1951–2015 period (Figure 1a), while positive estimates prevail in the future 2036–2100 period (Figure 1c). Regional analysis (e.g., with the largest spatial pooling considered) results in more coherent patterns in the spatial distributions of the estimated scaling rates (Figure S3). There is, however, still substantial variation between spatial patterns obtained from different simulations. Also, it is evident that spatial pooling cannot adequately sample low-frequency large-scale internal variability as represented by intersimulation differences (Figure S3a vs. S3b). Increasing the degree of spatial pooling will sample internal variability somewhat more completely, but at the cost of the loss of spatial details such that the estimated scaling rates are less informative at local spatial scales. The situation for higher percentiles (e.g., the 99th percentile) will be even worse since substantially more samples are needed for estimating rare extremes. These results are in line with previous findings that internal variability is the dominant source of uncertainty for the intensification of local-to-regional extreme precipitation by midcentury and even further (Deser et al., 2012, 2014; Fischer et al., 2013, 2014; Hawkins & Sutton, 2011).

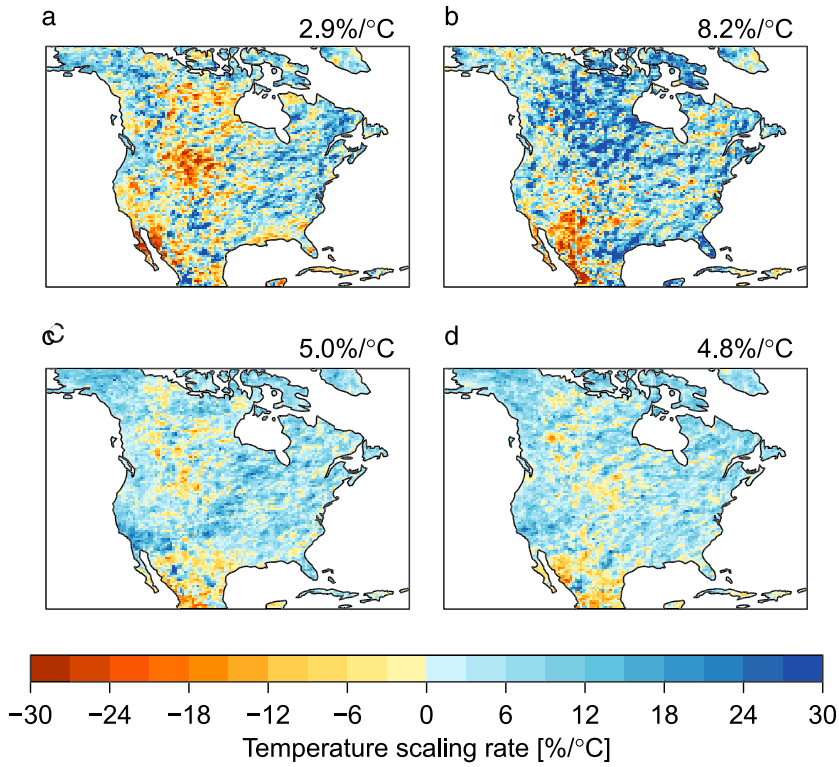

Figure 1. Temperature scaling rate estimates for extreme subdaily precipitation based on a single simulation of a length of typical observational records are highly uncertain. Panels show the scaling rates (in %/°C) for the 50th percentile of annual maximum 12-hr precipitation inferred by at-site analysis from two of the 35 Canadian Regional Climate Model version 4 (CanRCM4) simulations for a (a and b) historical (1951–2015) and a (c and d) future (2036–2100) period. The two simulations have the smallest and largest scaling rate estimates during the historical period spatially averaged over the North American land area among the 35 simulations. Numbers on the top of the panels indicate the spatial averages of the scaling rate estimates. See Figure S2 for results based on regional analysis with the largest spatial pooling considered.

To summarize, these findings show clearly that it is unlikely that the response of local extreme precipitation to anthropogenic forcing can be well constrained based on a single simulation of a length that is representative of the currently available long-term observational subdaily precipitation records. This further means that temperature scaling derived from historical observations at individual locations is unlikely to produce robust projections for future precipitation extremes. Stationarity may be dead (Milly et al., 2008), but the relationships that characterize nonstationarity in extreme daily and subdaily precipitation are unlikely to be robustly quantifiable using observational records or single simulations of lengths that are representative of those of currently available observational records.

3.2. Multiple Long Initial-Condition Climate Change Simulations Are Needed

The use of a much larger sample that is obtained by combining multiple initial-condition simulations should improve the estimation of scaling rates. This can be done in two ways. One is to aggregate data from multiple simulations and then fit one GEV distribution (either by at-site analysis or by regional analysis of the aggregated data). Another is to fit GEV distributions to individual simulations and then average estimated scaling rates from these simulations. We found that these two approaches provide almost identical results (not shown). For simplicity, we report only results based on the latter approach.

Figure 2 shows the fraction of the North American land area where temperature scaling for the 50th and 99th percentiles of annual maximum 12-hr precipitation can be robustly constrained (i.e., $||\mathrm{SRSE}|| > 5$) as a function of the number of available simulations. Results for extreme precipitation of other durations are qualitatively the same (Figures S4–S6). When using information for the whole 1951–2100 period along with regional analysis with the largest spatial pooling considered, an ensemble of 15 simulations or more allows

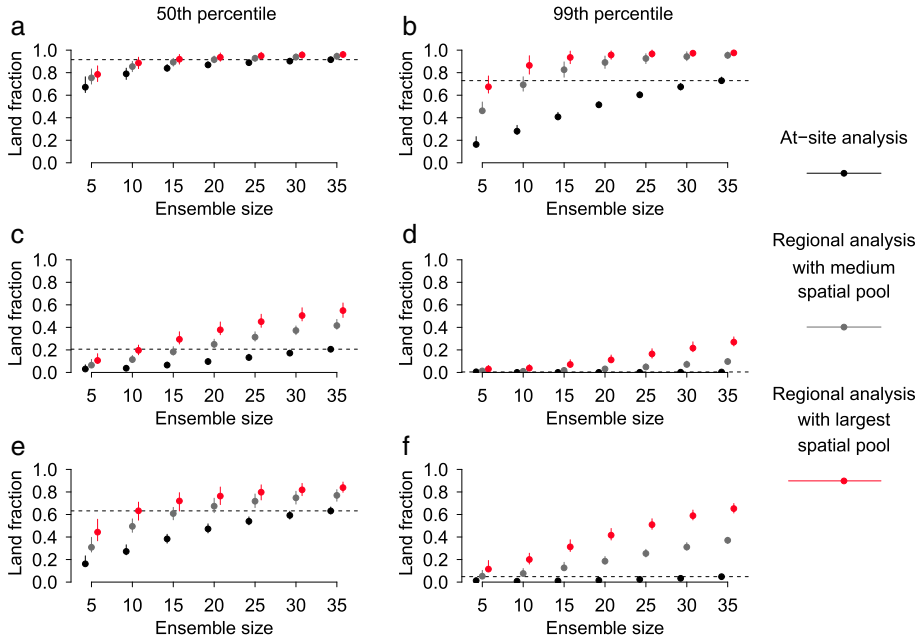

Figure 2. Multiple long simulations with strong external forcings are needed to robustly constrain temperature scaling rate estimates. Panels show the fraction of the North America having ||SRSE|| ≥ 5 for the 50th and 99th percentiles of annual maximum 12-hr precipitation as a function of the number of available simulations during period (a and b) 1951–2100, (c and d) 1951–2015, and (e and f) 2036–2100. Scaling rates are estimated respectively by at-site analysis (black), regional analysis with medium spatial pooling (over 3 × 3 grid cells; gray), and regional analysis with the largest spatial pooling considered (red). The whiskers show 5–95% uncertainty ranges obtained by a bootstrap approach. The dashed horizontal lines mark the fractions corresponding to at-site analysis using 35 simulations. See Figures S3–S5 for annual maximum precipitation of 1-, 6-, and 24-hr accumulation durations.

robust scaling rate estimates for both percentiles over most of North America (>90%); including more than 15 simulations does not necessarily lead to substantially more improvement (red bars in Figures 2a and 2b and S4–S6, panels a and b). On the other hand, scaling rate estimates based on the shorter 1951–2015 period may still not be robust even with 35 simulations and regional analysis, with robustly constrained temperature scaling (i.e., ||SRSE|| > 5) being found over less than 60% of North America (red bars in Figures 2c and 2d and S4–S6, panels c and d). The situation improves when information for the future 2036–2100 period is used because of the greater warming resulting from the intensified external forcing in this period (red bars in Figures 2e and 2f and S4–S6, panels e and f).

As expected, regional analysis by spatial pooling noticeably improves scaling rate estimates compared to at-site analysis. Given the same number of simulations of the same period, the scaling rate can be robustly estimated over a substantially greater area with regional analysis than with at-site analysis (red and gray bars vs. black bars in Figures 2 and S4–S6). This is particularly evident when the historical 1951–2015 simulations are used to estimate the temperature scaling for the extreme 99th percentile. With at-site analysis, it is unlikely to obtain robust scaling estimates for such a high percentile anywhere across the domain (black bars in Figure 2d), representing a marked contrast to regional analysis with the largest spatial pooling considered, which produced robust estimates over ~30% of the domain (red bars in Figure 2d).

Overall, the advantage of regional analysis in sampling internal variability is particularly obvious when the data length is short, the number of available simulations is small, the external forcing is weak, and for rare extremes. Nevertheless, even with regional analysis, the period analyzed needs to include a strongly evolving signal for the temperature scaling relation to be successfully constrained.

3.3. Robust Climate Model Simulated Extreme Precipitation Response to Forcing

Ensemble average scaling rate estimates based on the largest spatial pooling considered are geographically well organized and positive almost everywhere for both the 50th and 99th percentiles of annual

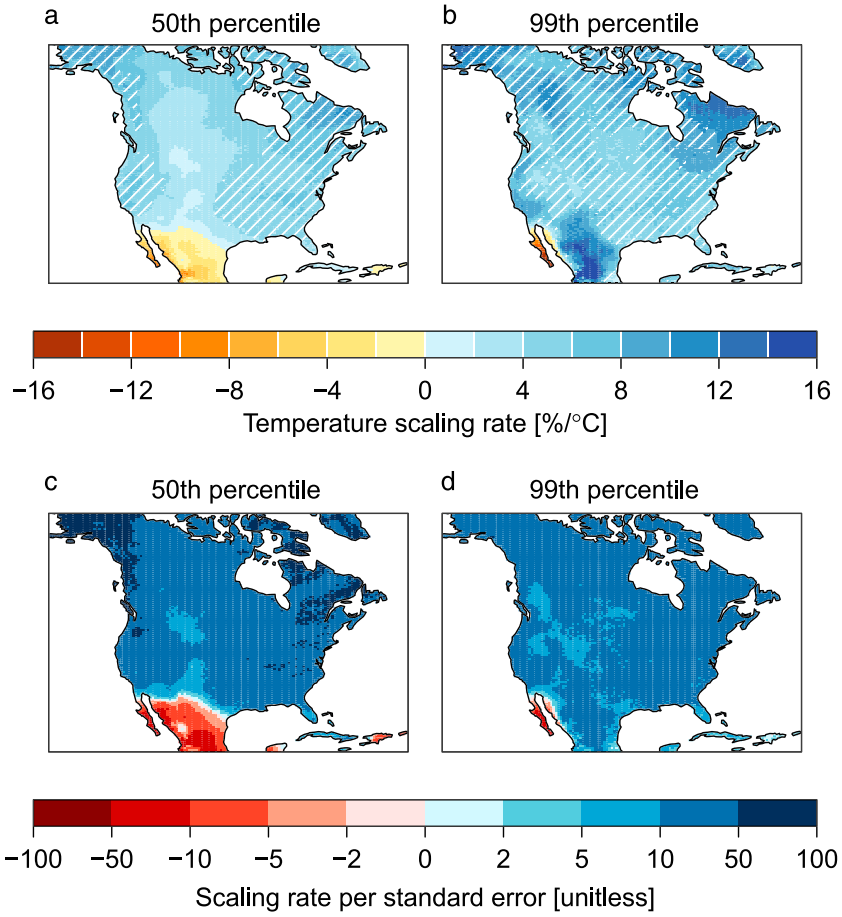

Figure 3. Extreme subdaily precipitation response to external forcing. (a and b) Ensemble mean of scaling rate estimates (in %/°C) for the 50th and 99th percentiles of annual maximum 12-hr precipitation based on scaling rates estimated by regional analysis with the largest spatial pooling for 1951–2100. Hatching highlights regions where the scaling rate estimates are consistent with the Clausius-Clapeyron relation (5.7–8.1%/°C for the central 95% range of CanRCM4 modeled North American temperatures). To plot the hatching, the scaling rate estimates are normalized with respect to annual latitudinal mean temperatures over North America. See Figure S6 for annual maximum precipitation of 1-, 6-, and 24-hr accumulation durations. (c and d) Uncertainties (unitless) in the scaling rate estimates shown in (a) and (b) expressed as *scaling rate per standard error*.

maximum 12-hr precipitation, except for southern subtropical subsidence regions where scaling rate estimates for the 50th percentile are negative (Figures 3a and 3b). Extreme precipitation of 1-, 6-, and 24-hr duration exhibits similar spatial patterns of temperature scaling, but with a tendency that shorter duration extreme precipitation appears to increase faster with global warming than longer duration extreme precipitation (Figure S7). We find that the domain average of the estimated temperature scaling rates for the 99th percentile are greater than that for the 50th percentile regardless of accumulation durations. For example, the 99th percentile of annual maximum 12-hr precipitation is 2.9%/°C larger than that for the 50th percentile (7.5%/°C for the former and 4.6%/°C for the latter). Such an asymmetric response suggests that high-impact low-frequency precipitation events are likely to intensify more substantially in a warmer climate in North America.

We find that for the moderate 50th percentile, the positive temperature scaling tends to converge to a limit predicted by the thermodynamic CC relation based on regional temperature trends (which varies from 5.7 to 8.1%/°C for the central 95% range of CanRCM4 modeled North American temperatures; Figure 3a). We also see that temperature scaling for the extreme 99th percentile may exceed the CC rate in parts of the Northeast (Figure 3b), especially for shorter duration extreme precipitation (see Figures S7b and S7e for 1- and 6-hr precipitation extremes). The positive feedback between precipitation formation, latent heat release, and low-

level moisture convergence may lift the temperature scaling for these high-intensity short-duration precipitation extremes to super CC levels (e.g., Trenberth et al., 2003). It should be noted that the estimated scaling rate for the 99th percentile in the southernmost subtropics is large compared to direct estimates from the climate model (as will be further discussed in the following section). It appears that even the most complex GEV model, which has six parameters, is too rigid to represent differences in the response of extreme precipitation to temperature for all percentiles in this particular region, at least as simulated in the CanRCM4 large ensemble. One could speculate that circulation change, such as a broadening of the subtropical subsidence region might be responsible since such a broadening would replace a circulation regime with low-level moisture convergence where there is a connection between temperature and precipitable water with a regime with low-level divergence where there is little, if any, connection.

We find that the 50th percentile of extreme precipitation is estimated to increase at roughly the CC rate (with respect to annual latitudinal mean temperatures over North America) over 40–50% of North American land area (50%, 48%, 45%, and 40% for annual maximum 1-, 6-, 12-, and 24-hr precipitation, respectively; see hatching in Figure 3a for annual maximum 12-hr precipitation and in Figure S7 left panel for extreme precipitation of other durations). In contrast, the estimated scaling rate for the extreme 99th percentile is consistent with CC over more widespread regions (65%, 67%, 69%, and 70% for annual maximum 1-, 6-, 12-, and 24-hr precipitation, respectively; see hatching in Figure 3b for annual maximum 12-hr precipitation and in Figure S7, right panel, for extreme precipitation of other durations). This contrast is in line with the expectation that high-intensity precipitation extremes are largely controlled by the enhancement of low-level atmospheric moisture content, which increases with warming at roughly the CC rate (e.g., Trenberth et al., 2003). There are substantial inland regions where the 50th percentile of extreme precipitation increases at less than the CC rate. The lower scaling rates in these regions are likely related to limits to the ability of the atmosphere to transport moisture to these locations (Pall et al., 2007), dynamic factors such as a weakening of upward winds during extreme precipitation events (Pfahl et al., 2017), an incorrect representation of land-atmosphere feedback processes (e.g., Lorenz et al., 2016), or combinations thereof.

Our findings of negative scaling for the 50th percentile but positive scaling for the 99th percentile of extreme precipitation in the southernmost subtropics are consistent with previous studies with global climate models, which have reported decreases in the mean of the distribution of annual maximum 24-hr precipitation in the driving CanESM2 simulations and in Community Earth System Model simulations (e.g., Fischer et al., 2014), and increases in the ensemble mean of the Coupled Model Intercomparison Project Phase 5 (CMIP5) projected 95th percentile (i.e., 20-year return level) of annual maximum 24-hr precipitation in this region (Kharin et al., 2013).

3.4. Temperature Scaling Can Project Changes in Precipitation Extremes

Figure 4 presents the bias for the projected relative changes in the 50th and 99th percentiles of annual maximum 12-hr precipitation during a near future 2031–2060 period relative to a historical 1951–2015 period estimated by the scaling relation (using the scaling rates shown in Figures 3a and 3b; top panels) and the ratios of RMSE_S/RMSE_D for these relative changes (bottom panels). For both percentiles, bias in the estimates obtained via temperature scaling is less than 15% over most of the continent, except for some localized subtropical regions south of California where the positive bias for the projected 99th percentile can be as large as 50% (Figures 4a and 4b). For the 50th percentile, RMSE_S is significantly less than RMSE_D at the 5% level or comparable to RMSE_D almost everywhere across the continent (Figure 4c), whereas for the 99th percentile RMSE_S significantly exceeds RMSE_D in the southernmost region and along the Northwest Coast (cross-hatching in Figure 4d). The later areas are generally coincident with areas where the bias in the scaling-based estimates is also large (Figure 4b vs. 4d).

Looking to a further future 2071–2100 period, the scaling estimates for the 50th percentile exhibit continued good performance as in the near future period, with small bias and RMSE that is significantly less than or comparable to that for the direct estimates (Figures 5a and 5c). For the 99th percentile, however, areas with RMSE_S/RMSE_D ratios that are significantly greater than 1 expand to more regions (~7.3%; cross-hatching in Figure 5d) due to increased positive bias in the scaling-based estimates (Figure 5b). Similar results are found for extreme precipitation of other durations (Figures S8 and S10). Overall, for both future periods, no more than 9% of North America would see RMSE_S significantly exceeding RMSE_D for both percentiles

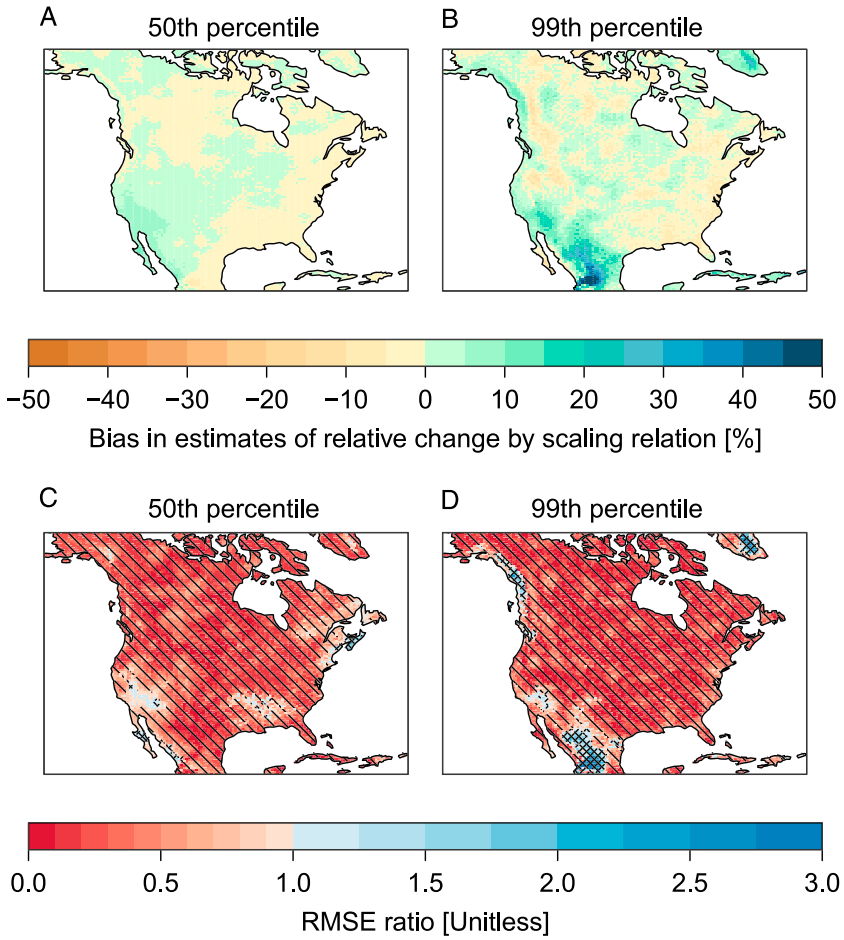

Figure 4. A well-constrained temperature scaling relation can project changes in precipitation extremes. (a and b) Bias for relative changes in the 50th percentile and 90th percentiles of annual maximum 12-hr precipitation during a near future period 2031–2060 relative to 1951–2015 estimated by the temperature scaling relation (using the scaling rates shown in Figure 3). (c and d) The ratio between the root-mean-square error for the relative changes estimated by the temperature scaling relation and that calculated directly from climate model projections. Cross-hatching shows where the ratio is significantly greater than 1 at the 5% level evaluated by a bootstrap approach, while hatching shows where the ratio is significantly less than 1at the 5% level. See Figures S7–S9 for annual maximum precipitation of 1-, 6-, and 24-hr accumulation durations.

of extreme precipitation of all considered accumulation durations. This shows the advantages of well-constrained temperature scaling in projecting future precipitation extremes up to the end of the century over direct estimation of these extremes from climate model projections for the majority of North American continent.

This result can only be obtained when the scaling relation is well constrained. To demonstrate this, we also estimated the quantiles of extreme precipitation for future periods using scaling rate estimates obtained from individual simulations of the historical 1951–2015 period, mimicking the challenge that is faced when attempting to use scaling relations diagnosed from historical observations to constrain projected changes in precipitation extremes. Estimates of precipitation extremes for the 2031–2060 period obtained in this way show bias that is comparable to, or larger than, those obtained with well-constrained scaling estimates (Figures S11a and S11b). More importantly, these estimates are highly uncertain, as indicated by very large relative RMSEs for both the 50th and 99th percentiles (Figures S11c and S11d). The RMSE of extreme precipitation changes estimated based on these individual scaling rate estimates is very significantly greater than RMSE_D (and thus RMSE_S) for both percentiles and throughout the domain. The situation is far worse in the further future 2071–2100 period (not shown). This confirms that temperature scaling derived

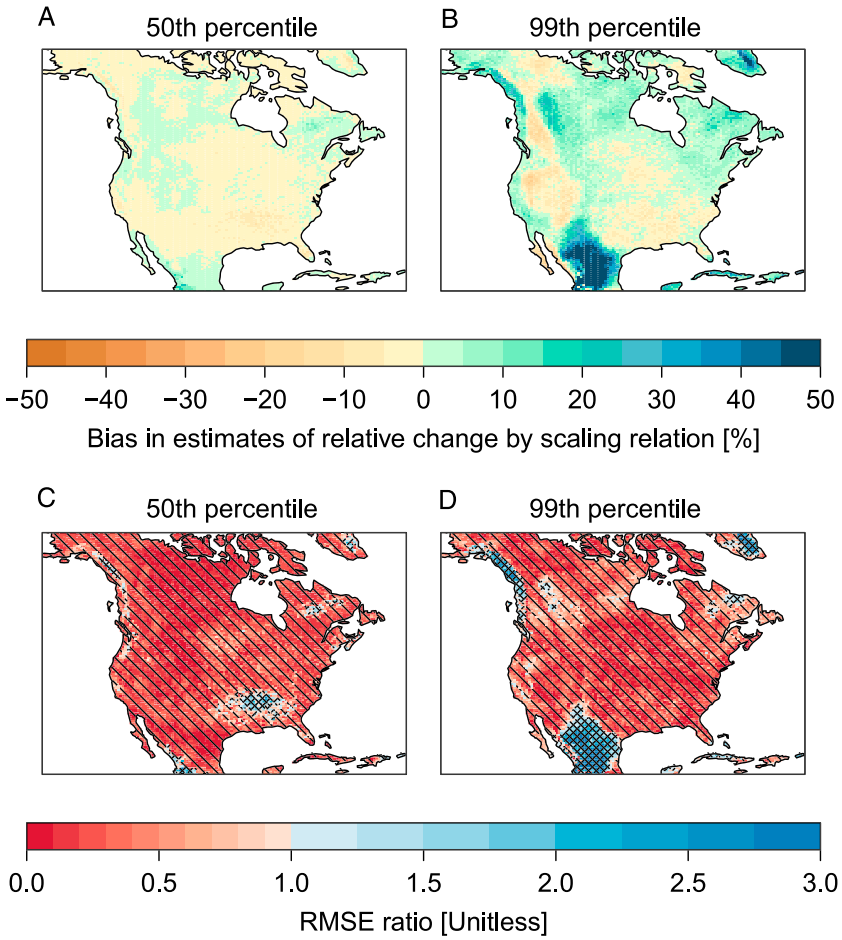

Figure 5. A well-constrained temperature scaling relation can project changes in precipitation extremes. As in Figure 4 but for a further future period 2071–2100. See Figures S7–S9 for annual maximum precipitation of 1-, 6-, and 24-hr accumulation durations.

from individual climate simulations of length comparable to that of typical instrumental records at individual locations cannot produce robust projections for future precipitation extremes, which in turn implies that temperature scaling relationships estimated from the limited historical observations are unlikely to be able to provide reliable guidance for future adaptation planning at local spatial scales.

Overall, these results show clearly that well-constrained temperature scaling relations are able to provide useful guidance about precipitation extremes in a warmer climate over most of North America and that they can do so at a lower overall cost than direct estimation from a large ensemble of climate change simulations. Temperature scaling may be less useful in areas such as the subtropical subsidence regions, where the thermodynamic response is complicated by the dynamic response such as a widening of the tropics, leading to temporally nonuniform precipitation change. We stress that statistical methods that are used for temperature scaling estimation should be flexible enough to resolve the scaling rates for the full range of precipitation extremes since it is possible that extreme precipitation of different severity levels responds differently to temperature warming.

4. Discussion

4.1. The Utility of Temperature Scaling in Constraining Future Extreme Precipitation From Observations

We have shown the advantages of properly constrained temperature scaling in projecting future precipitation extremes over direct estimation from climate model projections for most of North America.

In this case, temperature scaling would be of considerable utility in constraining estimates of extreme precipitation percentiles from observations to inform future changes. This can be done by adjusting the observed extreme precipitation percentile y in a historical period via the temperature scaling relation $y\times(1+\widehat{r})^{\Delta T}$, where $\widehat{r}$ is the estimated temperature scaling rate for that percentile and ΔT is the projected change in annual global mean temperature during a future period relative to the historical period.

A number of factors can affect the success of this approach in practical applications. First, as noted above, the scaling relation may be less useful in regions where the response of extreme precipitation to external forcings is strongly affected by dynamics, such as the southernmost subtropics of North America, where strong temporally nonuniform precipitation change may occur as the atmosphere warms. Likewise, the scaling relation can be of limited use if it is sensitive to the choice of emissions scenario. Pendergrass et al. (2015) have reported that increases in extreme precipitation depend on the magnitude of warming and not on the emissions scenario in most CMIP5 global climate models. This finding, however, is subject to some uncertainty due to the coarse resolution of global models. High-resolution models are needed to further investigate the role of various drivers of climate change for extreme precipitation changes. Second, the observed extreme precipitation percentile y should be estimated reliably. Indeed, estimation of high percentiles of local precipitation extremes from the limited historical observations is also affected by the sampling issues that challenge the estimation of the temperature scaling rate. This issue can be addressed, at least in part, by regional frequency analysis approaches such as the index flood method. Further, the challenge is somewhat mitigated by the fact that simpler extreme value models that do not use covariates and thus have fewer parameters can be used when estimating percentiles of historical extreme precipitation even though large-scale evidence demonstrates that stationarity is indeed dead (e.g., Zhang et al., 2013). Third, our findings indicate that at present we have no choice but to rely on climate models to estimate the temperature scaling rate. This therefore requires that the chosen climate model (or models) should be able to resolve the rate of long-term change of local extreme precipitation with global warming. Unfortunately, we have as yet, very limited observational studies that assess the temperature scaling rate simulated by climate models, apart from large scale detection and attribution studies such as Zhang et al. (2013); see also Min et al. (2011). This represents a major source of uncertainty in estimating future extreme precipitation by means of temperature scaling even if the temperature scaling relation itself can be robustly determined. Last, the choice of statistical methods for temperature scaling estimation is also critically important, as will be discussed separately in the following subsection.

4.2. The Importance of Statistical Methods for Estimating the Temperature Scaling

We have emphasized that the statistical methods used for temperature scaling estimation should be able to represent the characteristics of changes across the full range of precipitation extremes. Studies have used a GEV distribution with a linear link between temperature and the GEV location parameter to estimate temperature scaling rate for daily extreme precipitation (e.g., Westra et al., 2013) and for modeling the nonstationarity of extreme precipitation changes (Cheng & AghaKouchak, 2014; Sun & Lall, 2015). Also, there are extreme precipitation event attribution studies that have assumed an exponential link to temperature that directly affects the GEV location parameter and proportionally affects the scale parameter with a constant proportional rate (or dispersion coefficient; e.g., van Oldenborgh et al., 2015; Eden et al., 2016). Under these above parametrizations, the estimated temperature scaling will be independent of the percentiles and thus cannot represent possible differences between changes in low-to-moderate and rare extremes. For example, if estimating the relative change in the 99th percentile of annual maximum 12-hr precipitation using a temperature scaling relation estimated with a regional GEV model with temperature-dependent location parameter and constant dispersion coefficient, the RMSE_S will significantly exceed RMSE_D over 21% of the continent by the end of the century (not shown). This is substantially more widespread than if assuming a dispersion coefficient that also varies with temperature (8.7%; not shown) and if further assuming a temperature-dependent shape parameter (7.3%; Figure 5d). Moreover, it is possible that in some regions and for extreme precipitation of short durations, the most complex GEV model with all parameters depending on temperature is still too rigid to represent the characteristics of changes across the full range of precipitation extremes (e.g., regions in middle-to-high latitudes with cross-hatching in Figures 4 and 5 and S8–S10). In this situation, quantile regression that estimates the scaling rate for each percentile separately may help but at the expense of having to estimate many more parameters.

5. Concluding Remarks

We have presented examples illustrating that the estimation of local-scale short-duration (i.e., hourly to daily) extreme precipitation response to warming is highly uncertain based on model simulations with a length comparable to that of typical long-term instrumental records, due to the inability to adequately sample internal variability in precipitation extremes over short records. This means that precipitation trends estimated with available observations cannot provide reliable guidance for future planning at local spatial scales. Climate models coupled with statistical tools such as temperature scaling remain indispensable for understanding extreme precipitation in a warmer climate. High-resolution convection-permitting regional model experiments have been found to be able to more realistically simulate small-scale short-duration extreme precipitation (e.g., Prein et al., 2015; Westra et al., 2014). Coordinated multimodel multimember convection-permitting regional model experiments, though computationally expensive, will be needed to better assess their potential to better project future short-duration precipitation extremes. We have demonstrated the potential of regional analysis by spatial pooling with the index flood method in better sampling internal variability, suggesting that the use of regional analysis can help reduce the cost of regional climate change simulation experiments.

Our results emphasize that temperature scaling relationships estimated from the limited historical observations are unlikely to be able to provide reliable guidance for future adaptation planning at local spatial scales. In the climate simulated by CanRCM4, well-constrained temperature scaling provides a feasible basis for projecting future precipitation extremes over >90% of North America, but data that are equivalent in length of many multiples of historical observational records are required to robustly identify the scaling relationships. Compared with estimating future precipitation extremes directly from climate model projections, estimating changes in precipitation extremes with a well-constrained temperature scaling relation can provide extreme precipitation estimates with small bias and much lower uncertainty, and thus statistically significantly smaller RMSE, particularly in regions where the response of extreme precipitation to external forcings is not strongly affected by dynamics.

Acknowledgments
We thank Alex Cannon and Jana Sillmann for their comments that helped to improve this paper. We thank two anonymous reviewers for their comments and feedback on this paper. We acknowledge the Canadian Center for Climate Modeling and Analysis of Environment and Climate Change Canada for executing and making available the CanRCM4 large ensemble simulations. Chao Li was supported in part by the National R&D Program of China (2018YFC1507702). Data associated with this paper are available from http://data.ec.gc.ca/data/climate/scientificknowledge/the-canadian-regional-climate-model-large-ensemble/?lang=en.

References

Allen, M. R., & Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, *419*, 224–232. https://doi.org/10.1038/nature01092

Cheng, L., & AghaKouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, *4*. https://doi.org/10.1038/srep07093

Coles, S. G. (2001). *An introduction to statistical modeling of extreme values*. New York: Springer.

Deser, C., Knutti, R., Solomon, S., & Phillips, A. S. (2012). Communication of the role of natural variability in future North American climate. *Nature Climate Change*, *2*, 775–779. https://doi.org/10.1038/nclimate1562

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2014). Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, *27*(6), 2271–2296. https://doi.org/10.1175/JCLI-D-13-00451.1

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change*, *6*, 508–513. https://doi.org/10.1038/nclimate2941

Eden, J. M., Wolter, K., Otto, F. E. L., & van Oldenborgh, G. J. (2016). Multi-method attribution analysis of extreme precipitation in Boulder, Colorado. *Environmental Research Letters*, *11*, 124009. https://doi.org/10.1088/1748-9326/11/12/124009

Fischer, E. M., Beyerle, U., & Knutti, R. (2013). Robust spatially aggregated projections of climate extremes. *Nature Climate Change*, *3*, 1033–1038. https://doi.org/10.1038/nclimate2051

Fischer, E. M., & Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nature Climate Change*, *5*, 560–564. https://doi.org/10.1038/nclimate2617

Fischer, E. M., Sedláček, J., Hawkins, E., & Knutti, R. (2014). Models agree on forced response pattern of precipitation and temperature extremes. *Geophysical Research Letters*, *41*, 8554–8562. https://doi.org/10.1002/2014GL062018

Hanel, M., Buishand, T. A., & Ferro, C. A. T. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *Journal of Geophysical Research*, *114*, D15107. https://doi.org/10.1029/2009JD011712

Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, *37*, 407–418. https://doi.org/10.1007/s00382-010-0810-6

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis*. New York: Cambridge University Press.

Kendon, E. J., Powell, D. P., Jones, R. G., & Buonomo, E. (2008). Robustness of future changes in local precipitation extremes. *Journal of Climate*, *21*, 4280–4297. https://doi.org/10.1175/2008JCLI2082.1

Kharin, V. V., Flato, G. M., Zhang, X., Gillett, N. P., Zwiers, F., & Anderson, K. J. (2018). Risks from climate extreme change differently from 1.5°C to 2.0°C depending on rarity. *Earth's Future*, 6, 704–715. https://doi.org/10.1002/2018EF000813

Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, *119*(2), 345–357. https://doi.org/10.1007/s10584-013-0705-8

Lorenz, R., Argüeso, D., Donat, M. G., Pitman, A. J., van den Hurk, B., Berg, A., et al. (2016). Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble. *Journal of Geophysical Research: Atmospheres*, *121*, 607–623. https://doi.org/10.1002/2015JD024053

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management. *Science*, *319*, 573–574. https://doi.org/10.1126/science.1151915

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., et al. (2015). On critiques of "Stationary is dead: Whither water management". *Water Resources Research*, *51*, 7785–7789. https://doi.org/10.1002/2015WR017408

Min, S. K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature*, *470*, 378–381. https://doi.org/10.1038/nature09763

National Building Code of Canada (NCBC) (2005). https://www.nrc-cnrc.gc.ca/eng/publications/codes_centre/2015_national_building_code.html

Pall, P., Allen, M. R., & Stone, D. A. (2007). Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO_2 warming. *Climate Dynamics*, *28*(4), 351–363. https://doi.org/10.1007/s00382-006-0180-2

Pendergrass, A. G., Lehner, F., Sanderson, B. M., & Xu, Y. (2015). Does extreme precipitation intensity depend on the emissions scenario? *Geophysical Research Letters*, *42*, 8767–8774. https://doi.org/10.1002/2015GL065854

Pfahl, S., O'Gorman, P. A., & Fischer, E. M. (2017). Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, *7*, 423–427. https://doi.org/10.1038/CLIMATE3287

Prein, A. F., Langhans, W., Fosser, G., Andrew, F., 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*, 48–52. https://doi.org/10.1038/nclimate3168

Scinocca, J. F., Kharin, V. V., Jiao, Y., Qian, M. W., Lazare, M., Solheim, L., et al. (2016). Coordinated global and regional climate modeling. *Journal of Climate*, *29*(1), 17–35. https://doi.org/10.1175/JCLI-D-15-0161.1

Serinaldi, F., & Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources*, *77*, 17–36. https://doi.org/10.1016/j.advwatres.2014.12.013

Sherwood, S. C., Roca, R., Weckwerth, T. M., & Andronova, N. G. (2010). Tropospheric water vapor, convection, and climate. *Reviews of Geophysics*, *48*, RG2001. https://doi.org/10.1029/2009RG000301

Sun, X., & Lall, U. (2015). Spatially coherent trends of annual maximum daily precipitation in the United States. *Geophysical Research Letters*, *42*, 9781–9789. https://doi.org/10.1002/2015GL066483

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205–1217. https://doi.org/10.1175/BAMS-84-9-1205

van Oldenborgh, G. J., Otto, F. E. L., Haustein, K., & Cullen, H. (2015). Climate change increases the probability of heavy rains like those of storm Desmond in the UK—An event attribution study in near-real time. *Hydrology and Earth System Sciences Discussions*, *12*, 13,197–13,216. https://doi.org/10.5194/hessd-12-13197-2015

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., et al. (2011). The representative concentration pathways: An overview. *Climatic Change*, *109*, 5. https://doi.org/10.1007/s10584-011-0148-z

Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, *26*, 3904–3918. https://doi.org/10.1175/JCLI-D-12-00502.1

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., et al. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, *52*, 522–555. https://doi.org/10.1002/2014RG000464

Whan, K., & Zwiers, F. (2016a). Evaluation of extreme rainfall and temperature over North America in CanRCM4 and CRCM5. *Climatic Dynamics*, *46*, 3821–3843. https://doi.org/10.1007/s00382-015-2807-7

Whan, K., & Zwiers, F. (2016b). The impact of ENSO and the NAO on extreme winter precipitation in North America in observations and regional climate models. *Climatic Dynamics*, *48*, 1401–1411. https://doi.org/10.1007/s00382-016-3148-x

Zhang, X., Wan, H., Zwiers, F. W., Hegerl, G. C., & Min, S. K. (2013). Attributing intensification of precipitation extremes to human influence. *Geophysical Research Letters*, *40*, 5252–5257. https://doi.org/10.1002/grl.51010

Zhang, X., Zwiers, F. W., Li, G., Wan, H., & Cannon, A. J. (2017). Complexity in estimating past and future extreme short-duration rainfall. *Nature Geoscience*, *10*, 255–259. https://doi.org/10.1038/ngeo2911