Representation of U.S. Warm Temperature Extremes in Global Climate Model Ensembles

EMILY HOGAN

Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, Urbana, Illinois

ROBERT E. NICHOLAS

Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, Pennsylvania

KLAUS KELLER

Earth and Environmental Systems Institute, and Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania

STEPHANIE EILTS AND RYAN L. SRIVER

Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, Urbana, Illinois

(Manuscript received 12 February 2018, in final form 9 February 2019)

ABSTRACT

Extreme temperature events can have considerable negative impacts on sectors such as health, agriculture, and transportation. Observational evidence indicates the severity and frequency of warm extremes are increasing over much of the United States, but there are sizeable challenges both in estimating extreme temperature changes and in quantifying the relevant associated uncertainties. This study provides a simple statistical framework using a block maxima approach to analyze the representation of warm temperature extremes in several recent global climate model ensembles. Uncertainties due to structural model differences, grid resolution, and internal variability are characterized and discussed. Results show that models and ensembles differ greatly in the representation of extreme temperature over the United States, and variability in tail events is dependent on time and anthropogenic warming, which can influence estimates of return periods and distribution parameter estimates using generalized extreme value (GEV) distributions. These effects can considerably influence the uncertainty of model hindcasts and projections of extremes. Several idealized regional applications are highlighted for evaluating ensemble skill and trends, based on quantile analysis and root-mean-square errors in the overall sample and the upper tail. The results are relevant to regional climate assessments that use global model outputs and that are sensitive to extreme warm temperature. Accompanying this manuscript is a simple toolkit using the R statistical programming language for characterizing extreme events in gridded datasets.

1. Introduction

Many economic sectors such as health, agriculture, forestry, energy, and transportation are vulnerable to changes in global and regional climate (Handmer et al. 2012). Extreme weather events are a key driver

of damages in these sectors (e.g., Karl and Easterling 1999; Melillo et al. 2014; Kunkel et al. 1999). In particular, temperature extremes such as heat waves can negatively impact human health and disrupt electricity supplies (Wilbanks et al. 2008), reduce air quality (Wang and Angell 1999), lower the yields of grain crops (Schauberger et al. 2017; Schlenker and Roberts 2009), and increase stress on livestock (Karl et al. 2009). These events often lead to drought and low soil moisture and can also contribute to wildfires (Trenberth 2011).

Observational evidence suggests that the frequency and severity of these upper-tail-area temperature events

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-18-0075.s1>.

Corresponding author: Emily Hogan, eehogan2@illinois.edu

are increasing with global warming (Kunkel et al. 2008; Peterson et al. 2008; DeGaetano 2009; Diffenbaugh and Ashfaq 2010; Duffy and Tebaldi 2012; Sillmann et al. 2013a,b; Sriver et al. 2015), and there is strong evidence that this is due to planetary warming (Hartmann et al. 2013) and increasing greenhouse gases (Handmer et al. 2012). For example, the observed frequencies of cold days and nights around the world have been decreasing, while the frequencies of warm days and nights are increasing (Field et al. 2012; Dulière et al. 2012; Walsh et al. 2014). Over the continental United States (CONUS), record highs have outpaced record lows in recent decades (Walsh et al. 2014). Accurate estimates of temperature extremes are key to better understanding impacts associated with these instances (Field et al. 2012; Stein 2017). However, there is a lack of high-spatial- and high-temporal-resolution climate data that can be used to evaluate global changes in extreme events over the twentieth century due to limited spatial and temporal resolution over different areas of the planet (Diffenbaugh et al. 2016; Easterling et al. 2000; Easterling 2008). Additionally, complex interactions exist between anthropogenic global warming and large-scale circulation patterns (e.g., El Niño–Southern Oscillation, Pacific decadal and interdecadal variability, the North Atlantic Oscillation, and Atlantic Ocean variability; Hartmann et al. 2013), which pose major challenges for understanding how the frequency and severity of extreme temperature events may be changing (Diffenbaugh et al. 2016; Swain et al. 2014, 2016; Singh et al. 2014; Horton et al. 2015). These issues make estimating distributions of extreme temperatures difficult for both current-day climate (Easterling 2008) and future conditions (Huang et al. 2016).

Climate models have been able to provide valuable insights into past and potential future changes and variability in the climate system (Taylor et al. 2012). These models can often reproduce robust features of different climate parameters, particularly near-surface air temperature (Flato et al. 2013). However, these global climate model ensembles exhibit large spread in extreme regional temperature, particularly without any bias correction for the historical period (Sriver et al. 2015; Yao et al. 2013; Min et al. 2013). Coupled global climate model ensemble averages from phase 5 of the Coupled Model Intercomparison Project (CMIP5) models used by the Intergovernmental Panel on Climate Change (IPCC) are competent (Flato et al. 2013) at capturing key statistical characteristics of regional temperature extremes based on spatially and temporally limited observational records. However, ensemble variability is large and individual models can lack skill (Sriver et al. 2015; Yao et al. 2013; Min et al. 2013; Kharin et al. 2013). In addition, the ensemble variability

and uncertainty increase with decreasing spatial scale (Yao et al. 2013; Sriver et al. 2015). Modeled biases in extreme temperature are often assumed to follow mean temperature biases (Wuebbles et al. 2014) and changes in standard deviations (Huang et al. 2016; de Vries et al. 2012; Parey et al. 2013). Understanding the relationship between the biases and temperature extremes can help identify and prevent under- and overestimations of observed extremes by different individual models (Zahid et al. 2017; Janssen et al. 2016; Sriver et al. 2015).

Here we focus on characterizing the upper tails of daily temperature distributions from three different climate model ensembles sampling different model structures, resolutions and on initial-condition uncertainty. We evaluate yearly temperature extremes in global climate model (GCM) ensembles on scales where global models can skillfully represent key climate variables (Sriver et al. 2015). In using these large global climate model ensembles, we make no assumptions about spatial dependence or aggregation. We utilize simple statistical techniques to evaluate upper tails of observed, modeled, and projected temperature distributions. We present several “proof of concept” applications of the methods for regional climate change assessments under both stationary and nonstationary assumptions. Specifically, we analyze 1) the effects of internal variability and model structural uncertainties on estimates of simulated extreme temperature distributions for the historical period, 2) the influence of uncertainties in the upper tails on interpretations about how extremes may be changing under anthropogenic global warming, and 3) the relationship between simulated bulk temperature statistics and extreme warm events. Our results are relevant to regional climate impacts analysis sensitive to extreme temperature (e.g., Welch et al. 2010; Schlenker and Roberts 2009; Deschenes and Greenstone 2007).

2. Methods

We use extreme value theory (Coles 2001) to evaluate temperature extremes in three different GCM ensembles. We highlight results from a 50-member ensemble of the low-resolution version of the Community Earth System Model (CESM) (Sriver et al. 2015; Hogan and Sriver 2017; Vega-Westhoff and Sriver 2017), the NCAR 40-member Large Ensemble (LENS; Kay et al. 2015) and a 29-member subset of the CMIP5 ensemble (Taylor et al. 2012; see the list of models used in appendix A). These three ensembles allow us to assess how initial-condition uncertainty and structural model differences affect the spread in model temperature extremes. The low-resolution CESM ensemble [Sriver Low-Resolution

Ensemble (SLRE); [Sriner et al. 2015](#)] samples the initial conditions of the atmosphere and ocean, by using unique snapshots of a $\sim 10\,000$ -yr preindustrial control simulation as branch points for individual simulations. LENS ([Kay et al. 2015](#)) samples the initial conditions of the atmosphere by initializing simulations with perturbed atmospheric temperature at each grid point at the roundoff level (fourteenth decimal place). The three ensembles use historical forcing from 1850 to 2005 and are forced from 2006 to 2100 with projections from the representative concentration pathway (RCP) 8.5 ([Moss et al. 2010](#)). We use a high-resolution observation-based temperature product for the period 1956–2005 from [Maurer et al. \(2002\)](#), as the basis for evaluating model results during the historical period. Observations are used here primarily as a reference for interpreting model differences, given limitations in spatial and temporal coverage. All observations and model outputs are regridded to the low-resolution SLRE during post-processing and before evaluation, in order to control for resolution effects and achieve the most direct comparison between multiple temperature products and model outputs.

We use a block maxima approach ([Coles 2001](#)), in which we analyze distributions of annual maximum daily temperature, along with generalized extreme value (GEV) distributions. GEV distributions can be useful to assess climate extremes, though there are inherent trade-offs between the block size and total length of the record ([Stein 2017](#); [Huang et al. 2016](#)). These statistical methods have been used extensively in climate and hydrological studies ([Sriner et al. 2015](#); [Huang et al. 2016](#); [Collins et al. 2013](#); [Hasan et al. 2013](#); [Cooley 2009](#); [DeGaetano 2009](#); [Koutsoyiannis 2004](#); [Coles 2001](#); [Katz 1999](#)). We adopt the GEV distribution function to analyze extreme warm temperature:

$$F(x; \mu, \sigma, \zeta) = \exp\left\{-\left[1 + \zeta\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\zeta}\right\}, \quad \zeta \neq 0, \quad (1)$$

$$F(x; \mu, \sigma, \zeta) = \exp\left\{-\exp\left[\frac{-(x - \mu)}{\sigma}\right]\right\}, \quad \zeta = 0, \quad (2)$$

with parameters μ , σ , and ζ controlling the location, scale, and shape of the distribution, respectively ([Coles 2001](#)). The location parameter describes the mean of the extreme value distribution. The scale parameter describes the width of the distribution around the mean value, and the shape parameter characterizes asymmetry or “heaviness” in the tails of the distribution. A warming climate suggests an increase in the location parameter for variables such as temperature.

We define our block maxima temperatures as the maximum daily temperature value occurring between 1 June and 31 August of each year, or the maximum temperature of each summer. After evaluating different block lengths ([appendix B, Table B1](#)), we show results based on 50-yr intervals, sampling 1956–2005 for the historical period and 2050–99 for the projected period. This time frame provides a logical compromise given the length of the reliable instrumental record, changing anthropogenic forcings, and relevant time scales for regional decision planning. A block maxima approach may lead to underrepresenting extremes found within a given “block”; however, this method avoids selecting extreme events that are temporally dependent such as a lengthy heat wave, a caveat of the peak-over-threshold approach ([Huang et al. 2016](#)). Each of the three model ensembles evaluated in this study have different numbers of simulations, so in order to compare temperature variability in our ensembles, we evaluate the 95% range in spread (difference between the 0.025 and 0.975 quantiles) of the various outputs and parameters analyzed.

Temperature extremes are evaluated on three different spatial scales: (i) average values for the CONUS, (ii) spatially averaged scales typically used for regional climate assessments ([Wuebbles et al. 2017](#); [Melillo et al. 2014](#)), and (iii) local scales represented by individual grid points, such as over central Illinois (38°N, 90°W). We use root-mean-square error (RMSE) regression to assess model skill in simulating both the tails and central portions of the ensemble temperature distributions. The source code for our analyses and daily datasets for the regional extreme-value analysis are provided in the online repository (see supplemental information).

The observational dataset ([Maurer et al. 2002](#)) is used to help guide first-order evaluations of model skill and uncertainty. The [Maurer et al. \(2002\)](#) temperature dataset is a reanalysis dataset of temperature, gridded from observations, with output at 3-h time steps and a spatial resolution of $1/8^\circ$. We chose this dataset for this analysis because it provides high-spatial-resolution daily temperature data over the continental United States for the last 50 years. It is well suited for examining regional extremes and has been used widely in hydrological studies ([Wood et al. 2004](#); [Christensen et al. 2004](#); [Hayhoe et al. 2007](#)). We use the portion of the data record corresponding to the 50-yr model analysis period (1956–2005).

3. Results

We present several proof-of-concept examples focusing on the representation of temperature extremes in

TABLE 1. Observed and model ensemble CONUS-averaged GEV parameters (location, scale, and shape) for the historical period (1956–2005) and the projected period (2050–99).

	Historical (1956–2005)				RCP8.5 (2050–99)		
	Observations	SLRE	LENS	CMIP5	SLRE	LENS	CMIP5
Location mean (°C)	24.9	21.2	23.4	24.0	24.6	27.8	28.2
Location spread (°C)	—	0.24	0.23	4.60	0.26	0.41	5.80
Scale mean (°C)	0.41	0.48	0.43	0.56	0.82	1.06	1.04
Scale spread (°C)	—	0.20	0.21	0.31	0.25	0.31	0.55
Shape mean (°C)	−0.11	−0.25	−0.22	−0.25	−0.35	−0.38	−0.31
Shape spread (°C)	—	0.45	0.40	0.29	0.42	0.35	0.35

the three GCM ensembles sampling internal variability and different model structures. The applications feature both spatial and temporal analysis of temperature distributions. We analyze first-order effects on ensemble spread and uncertainty in both temperature tail statistics and bulk statistics, and we investigate potential relationships between biases in the mean state and temperature extremes. Our results are relevant to impacts analyses, particularly those that use gridded climate information as inputs and are sensitive to extreme temperature.

GEV parameters are estimated for each individual model simulation and for observations (Maurer et al. 2002), using a maximum likelihood method (MLE; Wuertz et al. 2017). We chose this method over probability weighted moment (PWM), because it provides more stable estimates for distributions with large outliers. Sensitivity results to the PWM method are shown in appendix B (Table B2). Ensemble average and spread in the location, scale, and shape parameters calculated for the CONUS are displayed in Table 1. Each parameter describes the distribution of block maxima summer daily temperature over the observed period of 1956–2005, and then the projected 50-yr period of 2050–99. Results show the ensemble-average location and scale parameters increase under future conditions and we see a small decrease in each ensemble-average shape parameter. The PWM method (appendix B, Table B2) exhibits similar results, with largest differences in the shape parameter and in CMIP5. We also provide additional analysis that focuses on evaluating different block length sizes in SLRE (appendix B, Table B1; Felici et al. 2007a,b). The spread in parameters is larger for the ensemble sampling structural model differences (CMIP5) than for the ensemble sampling initial-condition uncertainty (SLRE and LENS), with the exception of the shape parameter, where the spread is smaller for CMIP5 than in both SLRE and LENS. It is important to note that the distribution parameters are estimated jointly, and there may be some compensation effects between the scale and shape parameters.

We next divide the CONUS into six regions, in order to perform regional analysis on scales similar to recent assessments, such as the National Climate Assessment reports (Wuebbles et al. 2017; Melillo et al. 2014). To illustrate the spatial differences, regional skill, and future projected changes within each ensemble, we focus on the ensemble-average location parameter for annual maximum daily summer temperature over the aforementioned two time periods (contours in Fig. 1). Values overplotted are the 95% ensemble range of the location parameter. Contours of the mean location for each region in SLRE show a previously documented cold bias (Shields et al. 2012). We evaluate patterns in SLRE and CMIP5 in Fig. 2, showing daily temperature over the course of the year, averaged over the historical 50-yr period for each model simulation and Maurer et al. (2002). We see strong seasonal dependence in the model–data differences, with the SLRE cold biases appearing most prominently in the Northern Hemisphere summer months. Additionally seen in Fig. 1, we find large regional variability in the location parameter within each of the model ensembles. Interestingly, for each region and in each 50-yr time window, ensembles exhibit similar 95% ranges in temperature distributions when sampling internal variability (SLRE and LENS) and in the ensemble sampling both structural model differences and internal variability (CMIP5). This is consistent with the hypothesis that internal variability is responsible for a large proportion of the uncertainty in extreme temperature events on decadal time scales, for the models and spatial scales considered here. For the projected time period, we find the 95% range in ensemble distribution increases from historical to projected periods, implying variability in tail properties is time dependent and also dependent on forcing. We find that the overall percentage change in the location parameter (contoured) is similar for all three ensembles. We also find effects of nonstationarity in the ensembles sampling specifically internal variability (SLRE and LENS), suggesting internal variability in extremes is time dependent (increases with time) in these models.

Location (μ) for Temperature ($^{\circ}\text{C}$)

FIG. 1. Regional contour plots depicting the location parameter μ of yearly maximum summer temperature ($^{\circ}\text{C}$). (a) The observed location parameter (1956–2005). (b),(d),(f) The corresponding historical model ensemble-mean location parameters. (c),(e),(g) Ensemble-mean location parameter in 2050–99. Overplotted values are the 95% range in ensemble spread for each region.

The increase in spread of the ensembles with climate change shows that we may not be able to assume stationarity over these time periods. The projected location parameters are plotted as the percentage change from their respective historical parameters in Fig. S1 in the supplemental material. Here, we see the overall percentage change in the location parameter (contoured) is similar for all three ensembles. As the ensembles sample different sources of variability, a similar magnitude of change emphasizes the importance of internal variability in tail events. We refer the reader to Figs. S2 and S3 for contour plots of the shape and scale parameters for these regions. The scale parameter (Fig. S2) is replicated reasonably well by SLRE and LENS, while we find larger values in CMIP5. The 95% range of the GEV parameters across each ensemble still increases from historical years to future projections, but the magnitudes are smaller than for the location parameter. The trend in the shape parameter (Fig. S3) is not as obvious. The average shape across the ensemble remains fairly constant in SLRE, with similar 95% ranges from historical to future periods. We see a decrease in the shape parameter from

historical to projected periods in LENS to the west (meaning the largest negative values of the tail will be mostly affected), and an increase in the east of CONUS (meaning here, the smallest negative values of the tail will be most affected), with some regions showing a fair increase in the 95% range. This is reversed in CMIP5, where we see an increase in the shape parameter from historical to projected periods in the west, and a decrease in the east. Unlike the other two ensembles, the 95% range in ensemble parameters decreases a fair amount from historical to projected periods, with the exception of the north-central area.

For the remainder of the study, we evaluate individual model grid points, focusing on central Illinois in the main text as a proof of concept and example of the regional applications. We present results from other locations across various climates in the supplemental information (Denver, Colorado: 39 $^{\circ}\text{N}$, 105 $^{\circ}\text{W}$; Houston, Texas: 31.5 $^{\circ}\text{N}$, 94 $^{\circ}\text{W}$; Los Angeles, California: 25 $^{\circ}\text{N}$, 120 $^{\circ}\text{W}$; and New York City, New York: 42 $^{\circ}\text{N}$, 75 $^{\circ}\text{W}$). We highlight the effects of internal variability and model structure on the representation of local-scale extreme temperature.

FIG. 2. Climatology of daily average temperature ($^{\circ}\text{C}$) averages over the period 1956–2005 for the CMIP5 ensemble and SLRE for each of the 6 CONUS locations (see Fig. 1).

First, we produce return periods for grid cells representing annual maximum daily summer temperature for historical (1956–2005) and projected (2050–99) time periods in each of the three model ensembles, and compare these to observed maximum temperature (Fig. 3). The cold bias in SLRE (Shields et al. 2012) is visible in the raw output (Fig. 3a). The spread in return periods depicts uncertainty in the model ensemble extremes. We see a large amount of uncertainty in CMIP5 (Fig. 3c), which samples both structural model differences and internal variability and this variability increases for future projections. We provide similar plots of return periods where model runs have been bias corrected to the observations (Figs. 3d–f). We correct each ensemble member such that it has the same mean location as the observed distribution of temperature block maxima.

The simple adjustment of the location parameter greatly improves the model fit, but also masks key model differences. We perform simple statistical significance tests of these return periods by calculating the mean

RMSE of each model ensemble (overplot on the respective panels). RMSE is the square root of the variance of residual values. The RMSE of each ensemble run is found with respect to the observed extremes and then the mean of these values is calculated. Differences in the mean RMSE are attributed to differences in the internal variability of each ensemble as bias correction minimizes structural model differences for the historical period. To assess changes in the variability of the temperature extremes within each ensemble, we calculate the 95% ensemble range of the 0.5 quantile, and the 0.95 quantile of the block maxima distribution (Table 2). We also display percentage changes in quantiles from historical to future periods for the ensembles. For example, for the 50th percentile of SLRE, we show that this historical 95% spread of 0.369°C is 39.7% of the future 95% spread of 0.929°C .

Spread increases as the models simulate larger temperature extremes. This is portrayed through the increase in ensemble spread from the 0.5 quantile to the 0.95 quantile seen in all three ensembles.

FIG. 3. Return periods depicting yearly maximum summer temperature ($^{\circ}\text{C}$) over central Illinois (38°N , 90°W) for the three ensembles. (a)–(c) Raw output from the model ensembles and (d)–(f) data with location parameter adjustment (the difference between the respective ensemble mean and observed mean is subtracted from each line).

Uncertainty also increases in both quantiles from the historical to the future 50-yr period, implying the modeled ranges in extreme temperature increase with climate change for all three ensembles. In particular, SLRE and LENS show very similar percentage increases in variability from historical to future time periods.

We present return periods for grid cells representing Denver, Houston, New York City, and Los Angeles in SLRE (Fig. 4). We see the location of the return periods adjusting with climate; the desert and southern regions of Los Angeles and Houston are warmer, while the mountain and coastal regions of Denver and New York City are cooler. At first glance, the general shape of the return periods is similar between each of the four locations and central Illinois, but we do see a larger underestimation of the observations in central Illinois, when compared to the other regions. The changes in the 1-in-50-yr event are particularly striking. The magnitude of the 1-in-50-yr

event increases in all locations by approximately 2°C in all locations, with the exception of Denver, where the 1-in-50-yr event increases by almost 5°C . We also see a slight increase in the ensemble range from historical to future time periods. To complement the return-period plots, we also show the temperature survival functions (1 cumulative frequency) for all five locations in Fig. S4, highlighting the regional dependence of extreme temperature and effect of internal variability.

We illustrate the temporal changes in the three GEV parameters, displaying the 50-yr moving time window of location, scale, and shape parameters (assuming stationarity within each window; Fig. 5). Each time series is offset by its mean value for the displayed time period. Insets in each panel are time series of the 95% ensemble range. The contribution of internal variability to overall ensemble spread remains roughly constant in the location parameter (Fig. 5a) over these time periods for SLRE and

TABLE 2. The 95% ensemble ranges of the 0.5 and 0.95 quantiles of block maxima over central Illinois once each ensemble member has adjusted to the mean of the location parameter. Historical as a percentage of RCP8.5 is shown for each quantile and ensemble.

	SLRE		LENS		CMIP5	
	Historical	RCP8.5	Historical	RCP8.5	Historical	RCP8.5
Ensemble range of 50th percentile ($^{\circ}\text{C}$)	0.37	0.93	0.71	1.75	0.85	5.80
Historical as a percentage of RCP8.5: 50th percentile (%)		39.7		40.6		14.7
Ensemble range of 95th percentile ($^{\circ}\text{C}$)	0.99	2.10	1.91	3.34	3.47	7.81
Historical as a percentage of RCP8.5: 95th percentile (%)		47.7		57.2		44.7

FIG. 4. Return periods depicting yearly maximum summer temperature ($^{\circ}\text{C}$) over grid cells representing (top left) Denver (39°N , 105°W), (top right) Houston (31.5°N , 94°W), (bottom left) Los Angeles (25°N , 120°W), and (bottom right) New York City (42°N , 75°W) for SLRE.

LENS, but the contribution of both internal variability and structural model differences generally increases variability over the time period (CMIP5 ensemble), excluding the dramatic decrease in variability from 2000 to 2020. We attribute this dip in CMIP5 spread over this time period to outlier models that display a sudden increase in the location parameter over this time, or potentially to the setup of climate simulations with a transition between historical to RCP forcing, though accurate diagnosis is beyond the scope of this paper. We highlight these models in Fig. S5, noting that they are included in all analyses. We find an increase in the scale parameter over time for all three model ensembles (Fig. 5b), which represents a larger spread in temperature extremes in the future projected period. In addition, the 95% range generally increases with time in each ensemble, implying increased uncertainty in projected ranges. Last, we plot the time variations in the shape parameter (Fig. 5c). We find few robust changes overall in the shape parameter under anthropogenic warming. However, for LENS, the shape parameter does decrease slightly at the end of the

projected period. The 95% range in variability for the shape parameter remains fairly constant over the time period.

We assess ensemble skill in capturing the observed temperature distributions over central Illinois (Fig. 6) by comparing the RMSE of the main portion (95% range) of the distribution of summer daily average temperature with the RMSE of annual summer maximum temperature over the historical period (1956–2005). We calculate RMSE to evaluate our model ensembles against observed temperature data. Note that we base the calculation of RMSE on differences in the quantiles of the distributions. We use this approach to evaluate how accurately the model ensembles are able to replicate observed temperatures and can therefore be used to infer model skill.

We find that the model runs that generally demonstrate skill in replicating the central portion of the distribution show similar skill in replicating the tail events for temperature in these models (Fig. 6a; Wuebbles et al. 2014). We see a larger spread in temperature values for the ensemble sampling structural model differences

FIG. 5. Time series plots depicting the (a) location, (b) scale, and (c) shape parameters for yearly maximum summer temperature over central Illinois. Each point of the time series is calculated over a 50-yr period and is displayed above the middle of the 50-yr period. For example, the location parameter calculated from 1950 to 1999 is displayed above the year 1974. Bars plotted to the right of each panel show the 95% ensemble ranges for the last-calculated GEV parameter time step (2050–99, displayed above 1974). Each time series is normalized to its average calculated across the displayed time period and is offset by its respective ensemble-mean GEV parameter calculated at the first time point (1921–70). Inset in each panel is the time series of 95% ensemble ranges for the respective GEV parameter shown.

and internal variability (CMIP5) than for the ensembles sampling only internal variability (SLRE and LENS). In addition, we assess changes in model skill and variability after simple bias correction (Fig. 6b; note changes in x - and y -axis limits between Figs. 6a and 6b). Specifically, we subtract the mean temperature of each dataset from each dataset before estimating the 95% quantile range and the maximum yearly temperatures. This is performed to each individual model run and also to the observations (Maurer et al. 2002). We then calculate the RMSE between the location-adjusted observations and each model run. Location parameter adjustment significantly improves the skill of all model ensembles in simulating temperature in both the bulk distribution and the block maxima, along with reducing ensemble spread, but the correlation between maximum temperature and bulk distribution skill decreases. We suggest that this is most likely due to the mean location affecting the bulk and the tails of each distribution differently, due to skewness/asymmetry in the temperature distributions. For this grid point and model variable (central Illinois; temperature), model resolution and improvements increase skill in both average temperature and extremes for different versions of CESM. Adjusting the location parameter allows the low-resolution SLRE to exhibit greater skill in replicating observed extremes than higher-resolution versions of CESM within the CMIP5 ensemble [CCSM4, CESM1(BGC), and CESM1(CAM5)]; Fig. 6b).

Results from alternative grid points across CONUS show that models vary regionally in terms of their RMSE (Fig. 7). We evaluated the same RMSE technique for our additional four locations (Los Angeles, Houston, New York City, and Denver) and found each location replicates both maximum summer temperature and the bulk of the distribution with greater skill than central Illinois based on the Maurer et al. (2002) temperature dataset.

4. Conclusions

We examine temperature extremes in three large GCM ensembles and quantify model skill and uncertainty using 95% range and GEV theory. We evaluate uncertainties due to internal model variability (SLRE and LENS) and the combined effect of structural model differences and internal variability (CMIP5). We see larger uncertainty due to both structural model differences and internal variability in CMIP5 than in LENS and SLRE. We assess changes in temperature extremes on different

FIG. 6. Scatterplot depicting the RMSE of yearly summer maximum temperature in the model against the observations (x axis) compared to RMSE of daily average summer temperature (95% of the range) in the model against the observations (y axis). Data correspond to central Illinois. (a) Raw model output and (b) each data point adjusted before the RMSE is calculated (see text for details). Note the change in x - and y -axis limits for (b). The green, cyan, and yellow points are the three CMIP5 models that use different model versions of CESM than our low-resolution ensemble.

temporal and spatial scales over the CONUS and see increased uncertainty as we move from historical time spans to future projections and decreased spatial scales. Results using GEV methods suggest that both the scale and location parameters generally increase with warming expected with climate change. An increase in the scale parameter implies a widening of the GEV distribution. In contrast, the shape parameter is negative and in general decreases with climate change, implying a less heavy tail of the GEV in the future. The different sign of the changes in the scale and shape parameter, along with a severe lack a reliable observational record, creates challenges in diagnosing the overall changes in the upper tail of the GEV; thus, we focus primarily on analyzing return periods and survival functions of extreme warm temperature.

This work emphasizes the potential varying levels of climate information that can be necessary for regional impacts assessments, in particular related to extreme temperature characteristics based on limited observational data. We also highlight potential compensations and trade-offs between model skill in capturing tail versus bulk statistics. Both raw and location-parameter-adjusted data are plotted in several figures to emphasize these trade-offs in how to present and evaluate climate model output. We acknowledge the limitations involved due to the lack of data required for GEV analysis and

potential nonstationary effects. We only use one observational dataset (Maurer et al. 2002), which limits our conclusions, but as noted previously serves as a useful reference for interpreting model results. Additionally, interpolation effects due to regridding all

FIG. 7. RMSE of yearly summer maximum temperature in SLRE against observations (x axis) vs RMSE of daily average summer temperature (95% of the range; y axis) for central Illinois, Denver, Houston, Los Angeles, and New York City.

TABLE B1. SLRE CONUS average and spread of historical GEV parameters, isolating different block lengths: 40 (1961–2000), 50 (1956–2005), and 60 (1951–2010) years.

	SLRE		
	40-yr block	50-yr block	60-yr block
Location mean (°C)	21.1	21.2	21.2
Location spread (°C)	0.24	0.24	0.25
Scale mean (°C)	0.45	0.48	0.50
Scale spread (°C)	0.23	0.20	0.20
Shape mean (°C)	−0.31	−0.25	−0.22
Shape spread (°C)	0.50	0.45	0.35

data to the low-resolution SLRE may affect the magnitude of our temperature extremes.

Results can be useful for informing broader initiatives in impacts-related fields that use gridded climate information as inputs: for example, regional model assessments, economic and crop yield models or for comparing different downscaling techniques.

Acknowledgments. This study was cosupported by the U.S. Department of Energy (DOE), Office of Science, Biological and Environmental Research Program, Integrated Assessment Research Program, Grant DE-SC0005171, the DOE Program on Coupled Human Earth Systems (PCHES) under DOE Cooperative Agreement DE-SC0016162, and the Penn State Center for Climate Risk Management. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding entities. Any errors and opinions are, of course, those of the authors. All results, model codes, analysis codes, data, and model outputs used for analysis are freely available from the corresponding author and are distributed under the GNU general public license. The datasets, software tools, and other resources are provided as is without warranty of any kind, express or implied. In no event shall the authors

or copyright holders be liable for any claim, damages, or other liability in connection with the use of these resources. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in [appendix A](#) of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We provide all source code (written in R) in the supplemental information.

APPENDIX A

CMIP5 Models Used

This appendix provides a list of CMIP5 models used in this study. More information on these models can be found on the CMIP5 website (<http://cmip-pcmdi.llnl.gov/cmip5/index.html>): ACCESS1.0, ACCESS1.3, BCC_CSM1.1, BCC_CSM1.1(m), BNU-ESM, CanESM2, CCSM4, CESM1(BGC), CESM1(CAM5), CMCC-CESM, CMCC-CMS, CNRM-CM5, CSIRO Mk3.6.0, EC-EARTH, GFDL-ESM2G, GFDL-ESM2M, INM-CM4, IPSL-SM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MIROC5, MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM3, MRI-ESM1, and NorESM1-M.

APPENDIX B

GEV Parameter Sensitivity Tests

[Tables B1](#) and [B2](#) provide additional analysis on [Table 1](#).

TABLE B2. Observed and model ensemble CONUS average and spread of GEV parameters (location, scale, and shape) for the historical period (1956–2005) and the projected period (2050–99). Values are calculated using the probability weighted moment method.

	Historical (1956–2005)				RCP8.5 (2050–99)		
	Observations	SLRE	LENS	CMIP5	SLRE	LENS	CMIP5
Location mean (°C)	24.9	21.2	23.4	24.0	24.6	27.8	28.2
Location spread (°C)	—	0.25	0.24	4.60	0.27	0.39	5.75
Scale mean (°C)	0.40	0.48	0.43	0.56	0.82	1.08	1.06
Scale spread (°C)	—	0.19	0.23	0.35	0.25	0.30	0.54
Shape mean (°C)	−0.07	−0.22	−0.21	−0.23	−0.29	−0.33	−0.26
Shape spread (°C)	—	0.40	0.41	0.28	0.29	0.33	0.29

REFERENCES

- Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, **62**, 337–363, <https://doi.org/10.1023/B:CLIM.0000013684.13621.1f>.
- Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer, 208 pp.
- Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1029–1136.
- Cooley, D., 2009: Extreme value analysis and the study of climate change. *Climatic Change*, **97**, 77–83, <https://doi.org/10.1007/s10584-009-9627-x>.
- DeGaetano, A. T., 2009: Time-dependent changes in extreme-precipitation return-period amounts in the continental United States. *J. Appl. Meteor. Climatol.*, **48**, 2086–2099, <https://doi.org/10.1175/2009JAMC2179.1>.
- Deschenes, O., and M. Greenstone, 2007: The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *Amer. Econ. Rev.*, **97**, 354–385, <https://doi.org/10.1257/aer.97.1.354>.
- de Vries, H., R. J. Haarsma, and W. Hazeleger, 2012: Western European cold spells in current and future climate. *Geophys. Res. Lett.*, **39**, L04706, <https://doi.org/10.1029/2011GL050665>.
- Diffenbaugh, N. S., and M. Ashfaq, 2010: Intensification of hot extremes in the United States. *Geophys. Res. Lett.*, **37**, L15702, <https://doi.org/10.1029/2010GL043888>.
- , and Coauthors, 2016: Quantifying the influence of global warming on unprecedented extreme climate events. *Proc. Natl. Acad. Sci. USA*, **114**, 4881–4886, <https://doi.org/10.1073/pnas.1618082114>.
- Duffy, P. B., and C. Tebaldi, 2012: Increasing prevalence of extreme summer temperatures in the U.S. *Climatic Change*, **111**, 487–495, <https://doi.org/10.1007/s10584-012-0396-6>.
- Dulière, V., Y. Zhang, and E. P. Salathé Jr., 2012: Changes in twentieth-century extreme temperature and precipitation over the western United States based on observations and regional climate model simulations. *J. Climate*, **26**, 8556–8575, <https://doi.org/10.1175/JCLI-D-12-00818.1>.
- Easterling, D. R., 2008: Observed changes in the global distribution of daily temperature and precipitation extremes. *Climate Extremes and Society*, H. F. Diaz and R. J. Murnane, Eds., Cambridge University Press, 24–34.
- , J. L. Evans, P. Ya. Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje, 2000: Observed variability and trends in extreme climatic events: A brief review. *Bull. Amer. Meteor. Soc.*, **81**, 417–426, [https://doi.org/10.1175/1520-0477\(2000\)081<0417:OVATIE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<0417:OVATIE>2.3.CO;2).
- Felici, M., V. Lucarini, A. Speranza, and R. Vitolo, 2007a: Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part I: Stationary case. *J. Atmos. Sci.*, **64**, 2137–2158, <https://doi.org/10.1175/JAS3895.1>.
- , —, —, and —, 2007b: Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part II: Trend detection and assessment. *J. Atmos. Sci.*, **64**, 2159–2175, <https://doi.org/10.1175/JAS4043.1>.
- Field, C. B., and Coauthors, 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*. Cambridge University Press, 582 pp.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.
- Handmer, J., and Coauthors, 2012: Changes in impacts of climate extremes: Human systems and ecosystems. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, Cambridge University Press, 231–290.
- Hartmann, D. L., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 159–254.
- Hasan, H., N. Salam, and S. Kassim, 2013: Modeling annual extreme temperature using generalized extreme value distribution: A case study in Malaysia. *AIP Conf. Proc.*, **1522**, 1195–1202, <https://doi.org/10.1063/1.4801267>.
- Hayhoe, K., and Coauthors, 2007: Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dyn.*, **28**, 381–407, <https://doi.org/10.1007/s00382-006-0187-8>.
- Hogan, E., and R. Sriver, 2017: Analyzing the effect of ocean internal variability on depth-integrated steric sea-level rise trends using a low-resolution CESM ensemble. *Water*, **9**, 483, <https://doi.org/10.3390/w9070483>.
- Horton, D. E., N. C. Johnson, D. Singh, D. L. Swain, B. Rajaratnam, and N. S. Diffenbaugh, 2015: Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature*, **522**, 465–469, <https://doi.org/10.1038/nature14550>.
- Huang, W. K., M. L. Stein, D. J. McInerney, S. Sun, and E. J. Moyer, 2016: Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **2**, 79–103, <https://doi.org/10.5194/ascmo-2-79-2016>.
- Janssen, E., R. L. Sriver, D. J. Wuebbles, and K. E. Kunkel, 2016: Seasonal and regional variations in extreme precipitation event frequency using CMIP5. *Geophys. Res. Lett.*, **43**, 5385–5393, <https://doi.org/10.1002/2016GL069151>.
- Karl, T. R., and D. R. Easterling, 1999: Climate extremes: Selected review and future research directions. *Climatic Change*, **42**, 309–325, <https://doi.org/10.1023/A:1005436904097>.
- , J. M. Melillo, and T. C. Peterson, Eds., 2009: *Global Climate Change Impacts in the United States*. Cambridge University Press, 188 pp., downloads.globalchange.gov/usimpacts/pdfs/climate-impacts-reports.pdf.
- Katz, R. W., 1999: Extreme value theory for precipitation: Sensitivity analysis for climate change. *Adv. Water Resour.*, **23**, 133–139, [https://doi.org/10.1016/S0309-1708\(99\)00017-2](https://doi.org/10.1016/S0309-1708(99)00017-2).
- Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, <https://doi.org/10.1175/BAMS-D-13-00255.1>.
- Kharin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, <https://doi.org/10.1007/s10584-013-0705-8>.
- Koutsoyiannis, D., 2004: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrol. Sci. J.*, **49**, 575–590, <https://doi.org/10.1623/hysj.49.4.575.54430>.
- Kunkel, K. E., R. A. Pielke, and S. A. Changnon, 1999: Temporal Fluctuations in weather and climate extremes that cause economic and human health impacts: A review. *Bull. Amer. Meteor. Soc.*, **80**, 1077–1098, [https://doi.org/10.1175/1520-0477\(1999\)080<1077:TFIWAC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<1077:TFIWAC>2.0.CO;2).
- , and Coauthors, 2008: Observed changes in weather and climate extremes. *Weather and Climate Extremes in a Changing*

- Climate—Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*, T. R. Karl et al., Eds., U.S. Climate Change Science Program, 35–80.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate*, **15**, 3237–3251, [https://doi.org/10.1175/1520-0442\(2002\)015<3237:ALTHBD>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2).
- Melillo, J. M., T. Richmond, and G. W. Yohe, Eds., 2014: *Climate Change Impacts in the United States: The Third National Climate Assessment*. U.S. Global Change Research Program, 841 pp., <https://doi.org/10.7930/J0Z31WJ2>.
- Min, E., W. Hazeleger, G. J. van Oldenborgh, and A. Sterl, 2013: Evaluation of trends in high temperature extremes in northwestern Europe in regional climate models. *Environ. Res. Lett.*, **8**, 014011, <https://doi.org/10.1088/1748-9326/8/1/014011>.
- Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, <https://doi.org/10.1038/nature08823>.
- Parey, S., T. T. H. Hoang, and D. Dacunha-Castelle, 2013: The importance of mean and variance in predicting changes in temperature extremes. *J. Geophys. Res. Atmos.*, **118**, 8285–8296, <https://doi.org/10.1002/jgrd.50629>.
- Peterson, T. C., and Coauthors, 2008: Why weather and climate extremes matter. *Weather and Climate Extremes in a Changing Climate—Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*, T. R. Karl et al., Eds., U.S. Climate Change Science Program, 11–34.
- Schauberger, B., and Coauthors, 2017: Consistent negative response of US crops to high temperatures in observations and crop models. *Nat. Commun.*, **8**, 13931, <https://doi.org/10.1038/ncomms13931>.
- Schlenker, W., and M. K. Roberts, 2009: Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci. USA*, **106**, 15 594–15 598, <https://doi.org/10.1073/pnas.0906865106>.
- Shields, C. A., D. A. Bailey, G. Danabasoglu, J. T. Kiehl, S. Levis, M. Jochum, and S. Park, 2012: The low-resolution CCSM4. *J. Climate*, **25**, 3993–4014, <https://doi.org/10.1175/JCLI-D-11-00260.1>.
- Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh, 2013a: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.*, **118**, 1716–1733, <https://doi.org/10.1002/jgrd.50203>.
- , —, —, —, and —, 2013b: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *J. Geophys. Res. Atmos.*, **118**, 2473–2493, <https://doi.org/10.1002/jgrd.50188>.
- Singh, D., M. Tsiang, B. Rajaratnam, and N. S. Diffenbaugh, 2014: Observed changes in extreme wet and dry spells during the South Asian summer monsoon season. *Nat. Climate Change*, **4**, 456–461, <https://doi.org/10.1038/nclimate2208>.
- Striver, R. L., C. E. Forest, and K. Keller, 2015: Effects of initial conditions uncertainty on regional climate variability: An analysis using a low-resolution CESM ensemble. *Geophys. Res. Lett.*, **42**, 5468–5476, <https://doi.org/10.1002/2015GL064546>.
- Stein, M. L., 2017: Should annual maximum temperatures follow a generalized extreme value distribution? *Biometrika*, **104**, 1–16, <https://doi.org/10.1093/biomet/asw070>.
- Swain, D. L., and Coauthors, 2014: The extraordinary California drought of 2013/14: Character, context, and the role of climate change [in “Explaining Extreme Events of 2013 from a Climate Perspective”]. *Bull. Amer. Meteor. Soc.*, **95** (9), S3–S7, <https://doi.org/10.1175/1520-0477-95.9.S1.1>.
- , D. E. Horton, D. Singh, and N. S. Diffenbaugh, 2016: Trends in atmospheric patterns conducive to seasonal precipitation and temperature extremes in California. *Sci. Adv.*, **2**, e1501344, <https://doi.org/10.1126/sciadv.1501344>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Trenberth, K. E., 2011: Changes in precipitation with climate change. *Climate Res.*, **47**, 123–138, <https://doi.org/10.3354/cr00953>.
- Vega-Westhoff, B., and R. Sriver, 2017: Analysis of ENSO’s response to unforced variability and anthropogenic forcing using CESM. *Sci. Rep.*, **7**, 18047, <https://doi.org/10.1038/s41598-017-18459-8>.
- Walsh, J. D., and Coauthors, 2014: Our changing climate. *Climate Change Impacts in the United States: The Third National Climate Assessment*, J. M. Melillo, T. C. Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 19–67, <https://doi.org/10.7930/J0Z31WJ2>.
- Wang, J. X. L., and J. K. Angell, 1999. *Air Stagnation Climatology for the United States (1948–1998)*. NOAA/Air Resources Laboratory ATLAS 1, 73 pp.
- Welch, J. R., J. R. Vincent, M. Auffhammer, P. F. Moya, A. Dobermann, and D. Dawe, 2010: Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures. *Proc. Natl. Acad. Sci. USA*, **107**, 14 562–14 567, <https://doi.org/10.1073/pnas.1001222107>.
- Wilbanks, T., and Coauthors, 2008: *Effects of Climate Change on Energy Production and Use in the United States*. U.S. Climate Change Science Program, 84 pp.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, <https://doi.org/10.1023/B:CLIM.0000013685.99609.9e>.
- Wuebbles, D., and Coauthors, 2014: CMIP5 climate model analyses: Climate extremes in the United States. *Bull. Amer. Meteor. Soc.*, **95**, 571–583, <https://doi.org/10.1175/BAMS-D-12-00172.1>.
- , and Coauthors, 2017: Executive summary. *Climate Science Special Report: Fourth National Climate Assessment*, Vol. 1, D. J. Wuebbles et al., Eds., U.S. Global Change Research Program, 12–34, <https://doi.org/10.7930/J0DJ5CTG>.
- Wuertz, D., T. Setz, and Y. Chalabi, 2017: Rmetrics—Modelling extreme events in finance. Package “fExtremes,” Comprehensive R Archive Network Rep., 41 pp., <https://cran.r-project.org/web/packages/fExtremes/fExtremes.pdf>.
- Yao, Y., Y. Luo, J. Huang, and Z. Zhao, 2013: Comparison of monthly temperature extremes simulated by CMIP3 and CMIP5 models. *J. Climate*, **26**, 7692–7707, <https://doi.org/10.1175/JCLI-D-12-00560.1>.
- Zahid, M., R. Blender, V. Lucarini, and M. C. Bramati, 2017: Return levels of temperature extremes in southern Pakistan. *Earth Syst. Dyn.*, **8**, 1263–1278, <https://doi.org/10.5194/esd-8-1263-2017>.