

Estimating Changes in Temperature Distributions in a Large Ensemble of Climate Simulations Using Quantile Regression[✉]

MATZ A. HAUGEN, MICHAEL L. STEIN, AND ELISABETH J. MOYER

University of Chicago, Chicago, Illinois

RYAN L. SRIVER

University of Illinois at Urbana–Champaign, Urbana, Illinois

(Manuscript received 15 November 2017, in final form 2 August 2018)

ABSTRACT

Understanding future changes in extreme temperature events in a transient climate is inherently challenging. A single model simulation is generally insufficient to characterize the statistical properties of the evolving climate, but ensembles of repeated simulations with different initial conditions greatly expand the amount of data available. We present here a new approach for using ensembles to characterize changes in temperature distributions based on quantile regression that more flexibly characterizes seasonal changes. Specifically, our approach uses a continuous representation of seasonality rather than breaking the dataset into seasonal blocks; that is, we assume that temperature distributions evolve smoothly both day to day over an annual cycle and year to year over longer secular trends. To demonstrate our method's utility, we analyze an ensemble of 50 simulations of the Community Earth System Model (CESM) under a scenario of increasing radiative forcing to 2100, focusing on North America. As previous studies have found, we see that daily temperature bulk variability generally decreases in wintertime in the continental mid- and high latitudes ($>40^\circ$). A more subtle result that our approach uncovers is that differences in two low quantiles of wintertime temperatures do not shrink as much as the rest of the temperature distribution, producing a more negative skew in the overall distribution. Although the examples above concern temperature only, the technique is sufficiently general that it can be used to generate precise estimates of distributional changes in a broad range of climate variables by exploiting the power of ensembles.

1. Introduction

Time series of climate variables have generally been assumed to be separable into two components: randomness inherent in the underlying physical processes, which we call natural variability, and climatic trends, which we take to include seasonality and forced secular trends that follow from increasing concentrations of greenhouse gases. Recently, the degree to which natural variability may itself be changing has received significant scientific interest (e.g., Trenberth 2011; Donat and Alexander 2012; Deser et al. 2012a; Thompson et al. 2015; Kay et al. 2015). Potential changes in climate

extremes, because of their heightened societal impacts, are of special concern (e.g., Davison and Smith 1990; Stott et al. 2004; Chavez-Demoulin and Davison 2005; Eastoe and Tawn 2009; Otto et al. 2012; Swain et al. 2014; Singh et al. 2014; Trenberth et al. 2015; Diffenbaugh et al. 2015; Huang et al. 2016; Jalbert et al. 2017). However, fully characterizing the evolving natural variability of rare events is intrinsically challenging due to limited available observations or simulation data. Studies of future climate extremes therefore often employ statistical extreme value theory to make inferences about rare events with modest amounts of data (Swain et al. 2014).

In this work, we study the entire distribution of temperatures in a transient climate, including rare events, by employing quantile regression on an ensemble of simulations of an identical forcing scenario from a single climate model. An ensemble with sufficient sampling of the initial conditions' uncertainty will reflect the natural

[✉] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-17-0782.s1>.

Corresponding author: Matz A. Haugen, matz.haugen@orbitalinsight.com

variability of the climate system, since each simulation can be treated as statistically independent in terms of its natural variability. The increased data provided by multiple simulations enable more confident statements about changes in the statistical behavior of the system than can be made with a single simulation. The use of initial conditions for characterizing internal variability is growing rapidly (e.g., Deser et al. 2012a,b, 2014; Fischer and Knutti 2014; Kay et al. 2015; Striver et al. 2015; Rodgers et al. 2015; Hagos et al. 2016). Deser et al. (2012a,b) and Fischer and Knutti (2014) in particular discuss how ensembles help distinguish internal climate variability from anthropogenic effects on temperature changes and allow more comprehensive estimates of the model's temperature response to radiative forcing.

Large ensembles offer at least three advantages over a single simulation of a climate model. The most obvious advantage is that the increased data volume allows closer examination of the entire distribution of a climate variable. Studies of climate variability to date address both the center of the distribution (e.g., Semenov and Bengtsson 2002; Räisänen 2002; Kitoh and Mukano 2009; Screen 2014; Schneider et al. 2015) and its tails (e.g., Katz and Brown 1992; Meehl et al. 2009; Northrop and Jonathan 2011; Davison et al. 2012; Huser and Davison 2014; Trenberth et al. 2015; Huang et al. 2016; Jalbert et al. 2017), where the latter is often studied via extreme value theory. Fischer and Schär (2009) use an ensemble to study seasonal variability in a way that accounts for intraseasonal changes in mean temperature. A more limited body of studies address overall distributional changes in climate variables, but these generally focus on observations or observation-based data products, which are necessarily limited in terms of data volume and therefore require spatial or temporal aggregation (Donat and Alexander 2012; Stainforth et al. 2013; Chapman et al. 2013; Huybers et al. 2014; McKinnon et al. 2016; Rhines et al. 2017). Aggregating data spatiotemporally requires stationarity assumptions of the signal or explicitly modeling the spatiotemporal structure. When studying model projections using ensembles, the large amounts of data at each location allow us to estimate changes in the distribution of climate variables (e.g., temperature) without spatial aggregation. For example, Ylhäisi and Räisänen (2014) use CMIP3 data to remove a trend and a seasonal cycle and then take differences of quantiles of the residual temperature to measure variability and skewness. In their work, the seasonal cycle is not allowed to evolve over time, and the quantiles of the residual temperature are assumed constant within seasons and over 20-yr blocks.

A second advantage provided by large single model ensembles is that complex trends in both means and

variability can be stably estimated, in contrast to estimating quantiles using limited stretches of observational data and assuming linear trends in time (see, e.g., Franzke 2015; Gao and Franzke 2017). Forcings are not linear over centennial time scales, and a linear approximation can be misleading (Poppick et al. 2017). As we will show, daily temperature distributions have complex evolution in time, following different trajectories for different quantiles (i.e., different parts of the distribution). Analysis methods should therefore be able to take into account these complexities both in time and across quantiles.

Finally, a third advantage of ensembles is that they allow a more natural treatment of seasonal variation in climate variables. In situations of limited data, it is common to treat seasons separately, assuming that each season has a temporally constant average and stationary statistical properties discontinuous from neighboring seasons. With ensembles of simulations, we can allow for a smooth change in the underlying trend from day to day, using a parsimonious set of parameters. By modeling the entire year on a continuum, we can explore how each season transitions to the next and how seasonal patterns change over time, features that may be highly dependent on both geographic location and quantile.

We describe here a methodology for exploiting ensembles to study changing climate variability that captures these advantages: we model the complete distribution of daily temperatures as a continuous function of both seasonality and secular climate change over time. That is, we use quantile regression with separate basis functions for seasonal effects, for long-term trends, and for their interaction (i.e., changes in seasonality over the long term). Such an ensemble-based approach captures subtle features of seasonal changes and is well positioned for the purposes of uncertainty quantification. Because each simulation is treated as an independent sample drawn from the same stochastic process, we can obtain uncertainty quantifications for all estimates by resampling complete simulations from the ensemble without having to model temporal dependence within each simulation.

In the sections that follow, we describe estimated changes in both bulk and tail variability as differences in two quantiles; a large quantile difference implies more variability in a given part of the distribution. The difference between two quantiles in the low tail, for example, is a measure of the spread or thickness of the lower tail. Figure 1 gives a pictorial explanation of how quantile differences reflect bulk and tail variability.

2. Data

We apply our algorithm to an ensemble of 50 historical/future simulations of the Community Earth System Model

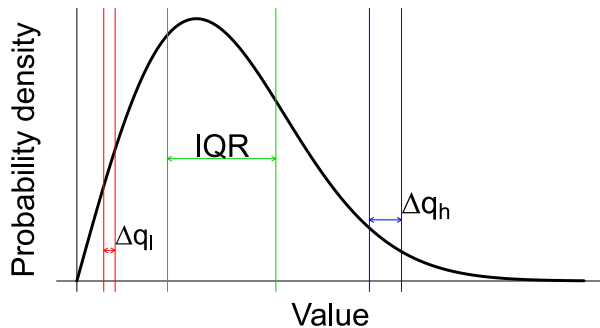


FIG. 1. An illustration of concepts and values related to distributions used in this paper. The cartoon shows a positively skewed (right skewed) probability distribution and the three quantile differences discussed in this paper, in the low tail, high tail, and middle of the distribution. The q quantile in a distribution is the value such that the probability of being below it has probability q . Here, Δq_l is the difference between the 0.05 and 0.025 quantiles; the IQR between the 0.75 and 0.25 quantiles; and Δq_h between the 0.975 and 0.95 quantiles. The values Δq_l and Δq_h quantify variability in extreme values, while IQR quantifies variability in the bulk of the distribution.

(CESM; [Srивer et al. 2015](#)). The atmospheric component is the low-resolution Community Atmosphere Model version 4, with T31 spectral resolution ($\sim 3.75^\circ \times 3.75^\circ$) and 26 vertical levels. The model ocean component is the low-resolution version of the Parallel Ocean Program version 2 ([Smith et al. 2010](#)), with a nominal horizontal grid resolution augmented to approximately 1° at the equator. The ocean model contains 60 vertical levels, down to a maximum depth of 5500 m.

Each simulation in the ensemble is initialized with ~ 4000 years of spinup using constant preindustrial conditions, after which the 50 historical hindcasts (1850–2005) are initialized from snapshots of the coupled model state taken every 100 years, so that the last hindcast is initialized after approximately 9000 years of the control simulation. Each hindcast is then extended to 2100 using the representative concentration pathway (RCP) 8.5 scenario. The 100-yr gap between each new initialization ensures nearly independent ensemble members that fully capture internal variability within the coupled system. RCP8.5 corresponds to anthropogenic radiative forcing of roughly 8.5 W m^{-2} by 2100 ([Moss et al. 2010](#)). More information about the model and ensemble design can be found in [Srивer et al. \(2015\)](#), [Hogan and Srивer \(2017\)](#), and [Vega-Westhoff and Srивer \(2017\)](#).

CESM does show some known biases that affect primarily temperature means (and possibly trends in means), but also to some extent the higher-order moments of the temperature distribution (e.g., variance and skewness). Known model biases include reduced ocean

heat transport, low North Atlantic sea surface temperature, and excessive Northern Hemisphere sea ice ([Shields et al. 2012](#)). The model generally underestimates both temperature and precipitation extremes, compared with observations; that is, the mean of the extreme value distributions is biased, but the scale and shape are consistent with observations for the continental United States ([Srивer et al. 2015](#)).

In assessing the practical relevance of our results on the shapes of temperature distributions, it is perhaps helpful to compare CESM temperatures with those from the ERA-Interim data product ([Dee et al. 2011](#)). [Figure 2](#) shows the model/reanalysis comparison for winter; for summer, see Fig. S1 of the online supplemental material. The model underestimates variability in some places and produces excessively cold winter temperatures in the Arctic. The resulting temperature gradients contribute to excess variability and negative skew in the northern midlatitudes. Skewness is a measure of asymmetry in a distribution, with negative skewness roughly corresponding to a distribution with a fatter lower tail than upper tail; see [appendix A](#) for its definition. We see that despite some discrepancies, [Fig. 2](#) shows that the model exhibits skill in capturing where wintertime skewness is positive and negative. Throughout this work, we will show in-depth analysis from three locations with distinct temperature distributions to highlight our proposed method (locations a, b, and c shown in [Fig. 2](#)). See Fig. S2 for comparison of model and reanalysis temperature distributions in both summer and winter for these locations.

3. Methods

In the methodology presented here, we model temperature at each location as a function of both seasonality and long-term change of the annual temperature distribution. We use two independent variables: seasonality is represented by a variable d (the day of the year; spanning values 1–365), and change in annual temperature is represented by a variable t (years elapsed since 1850; spanning 0–250 for these scenarios). We thus assume that each temperature quantile can be described by two sets of basis functions that represent the two variables' separate relationships with temperature (called here $\{f_i(d)\}$ and $\{g_j(t)\}$) and interaction terms $h_i(d)s_j(t)$, where f_i , g_j , h_i , and s_j are all smooth functions of the appropriate variable. The interaction terms are required to capture effects in which long-term temperature evolution differs between seasons (e.g., the robust projection that winter temperatures warm more than summer temperatures). To impose our smoothness condition, we assume that

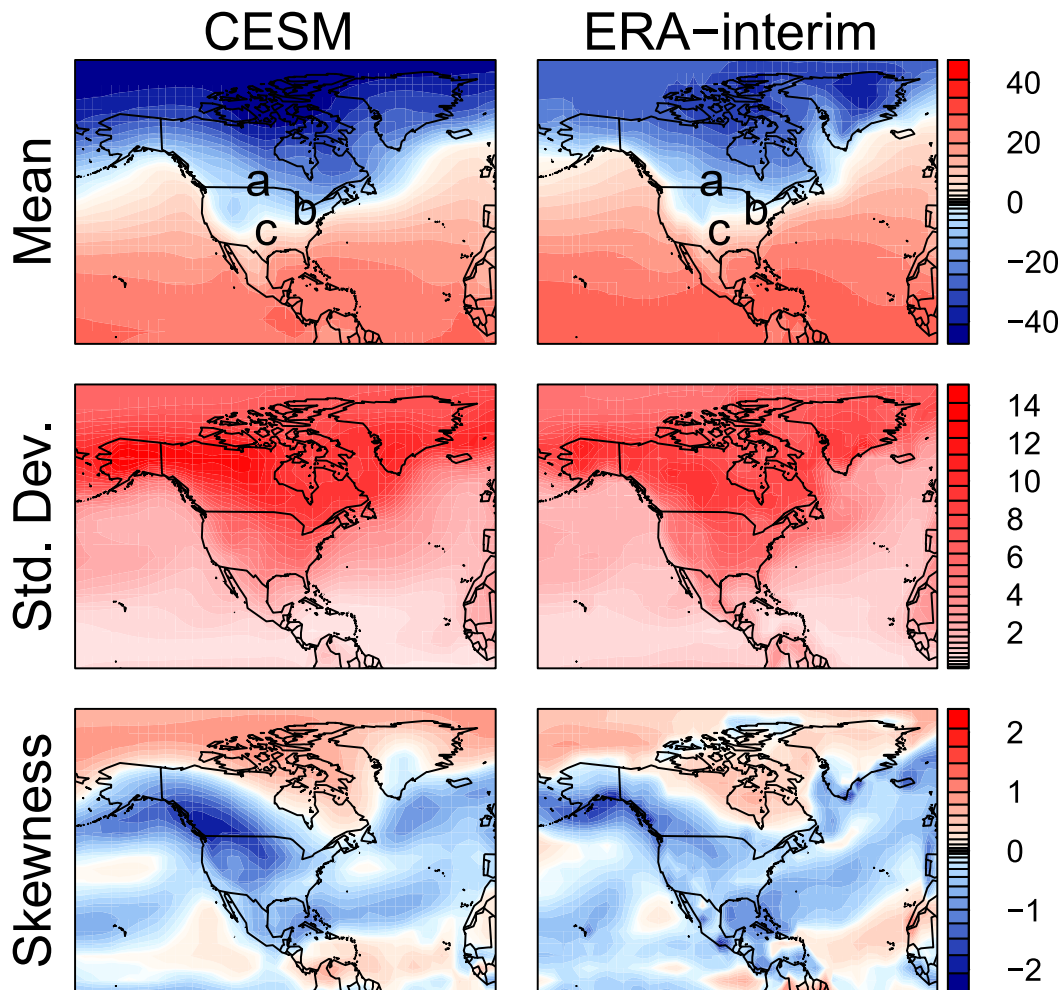


FIG. 2. Comparison of daily temperature distribution properties [(top) mean ($^{\circ}\text{C}$), (middle) standard deviation ($^{\circ}\text{C}$), and (bottom) skewness (dimensionless)] between the CESM ensemble and ERA-Interim for winter (DJF; aggregating all daily temperatures without deseasonalizing). We compare the years 1979–94, the first available 15 years of the ERA-Interim dataset, and upscale the reanalysis from 0.75° to 3.75° resolution to match CESM. Winter skewness over the continental United States is negative in both model and reanalysis, implying a thicker lower tail; see Fig. 1 and appendix A for example and definitions. Overall, large-scale geospatial patterns are similar in both datasets, though some discrepancies are present (e.g., abnormally cold model Arctic winters). Locations that will be used in examples throughout the paper are marked with letters a–c; these are ordered from north to south, with latitudes and longitudes of $(50.1^{\circ}, -101^{\circ})$ for location a, $(42.7^{\circ}, -82^{\circ})$ for location b, and $(35.3^{\circ}, -98^{\circ})$ for location c.

f_i , g_j , h_i , and s_j are cubic splines, which are piecewise cubic polynomials with a continuous second derivative [for a review of cubic polynomial basis functions, see Hastie et al. (2009), chapter 5]. Because the seasonality variable d is periodic, its basis functions are also assumed periodic. For more details, see appendix B, section a.

We choose the number of basis functions by evaluating a metric representing model adequacy. Our model sufficiency criterion is aimed at capturing the long-term underlying signal. We do not require estimated quantile functions to capture transient events

during the historical period like volcanic eruptions. Details on how we select the number of basis functions are given in appendix B, section b. In our climate simulation output, the intraseasonal effect requires more detailed modeling than the interseasonal effect. In the results shown here, we fit the model with 15 terms (i.e., basis functions) for the main seasonal effect $\{f_i\}$, but the interaction terms require less seasonal complexity, so we use only three terms for $\{h_i\}$. We use four terms for both the temporal change $\{g_j\}$ and the interaction terms $\{s_j\}$. That is, modeling long-term change generally requires fewer terms than modeling seasonality. In summary, we

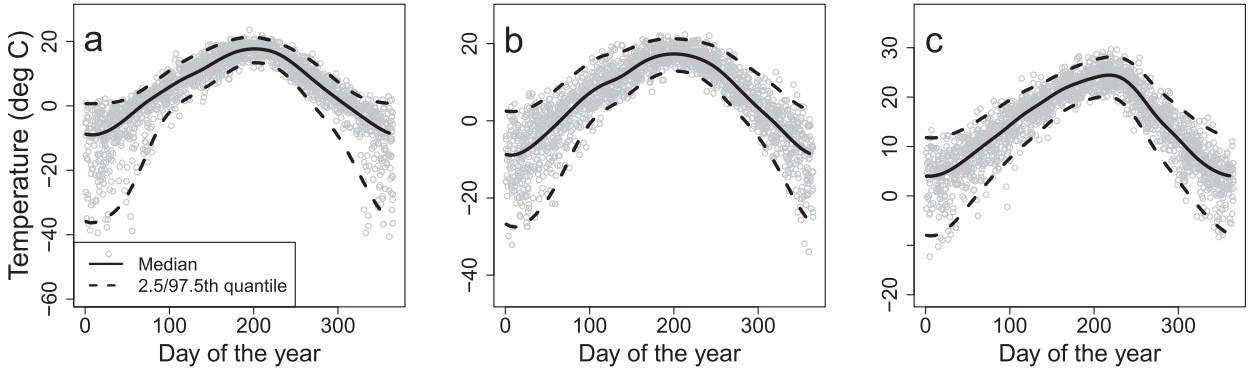


FIG. 3. Illustration of results of our quantile estimation procedure using the 50-member CESM ensemble. The figure shows ensemble daily mean temperatures for the year 1850 for the three representative locations a–c plotted in Fig. 2. The ensemble provides 50 points per day, but for clarity, we show only 10% of the data. Solid lines show the median daily temperature, and dashed lines show the 0.025 and 0.975 quantiles estimated by our procedure. The locations of the points outside of the 0.025 and 0.975 quantile curves are fairly evenly spread across day of year (notwithstanding the sizes of the exceedances), suggesting that these estimated quantile curves capture the seasonally changing patterns in the tails of the distributions reasonably well.

use 32 basis functions in total, including an intercept term. We then fit each q quantile of temperature

$$T_q(d, t) = \alpha + \sum_i a_i f_i(d) + \sum_j b_j g_j(t) + \sum_{i,j} c_{i,j} h_i(d) s_j(t), \quad (1)$$

where all of the coefficients depend on q , but we suppress the dependence for convenience. This fit determines coefficients α , a_i , b_j , and $c_{i,j}$ for each quantile at each location.

To simplify notation, consider a matrix \mathbf{X} where each column corresponds to a basis function, and each row refers to a unique value of d , t , and ensemble member. Using this matrix, we construct our temperature model in vectorized form:

$$\mathbf{T}_q = \mathbf{X}\boldsymbol{\beta}_q, \quad (2)$$

where $\boldsymbol{\beta}_q$ contains the 32 basis coefficients α , a_i , b_j , $c_{i,j}$. The predictor matrix \mathbf{X} has 32 columns, each corresponding to one basis function, and $365 \times 250 \times 50$ rows. To get a confidence interval for each entry of \mathbf{T}_q , we reestimate the coefficients $\boldsymbol{\beta}_q$ 100 times using resampled datasets. We resample the data by randomly drawing with replacement 50 whole simulations from our ensemble of 50 simulations. By resampling complete realizations, the dependency structure within realizations is maintained in the resampled data. Appendix B, section c provides further details about uncertainty quantification.

As an example of a typical model fit, we show in Fig. 3 the seasonal cycle in CESM daily temperatures for three locations, along with estimates of low, median, and high quantiles. We show here data from 1850 to demonstrate the seasonal fit rather than that of the long-term trend.

All locations show strong seasonal differences in variance that are well represented by our smooth estimates. Relevant features that are captured include an asymmetrical seasonal cycle in all locations; a clear left skewness in wintertime in all three locations (although most pronounced in the higher-latitude locations a and b); and a distinct springtime shoulder in the higher-latitude locations. These examples show the benefit of explicitly modeling the seasonal cycle in variability through smoothly varying quantile functions. The more standard practice of treating all days within a season as statistically identical would tend to obscure nuances evident in Fig. 3, such as the decrease in winter temperature spread (variability) from early to late winter.

4. Results and discussion

To facilitate comparison with previous studies, we first perform a preliminary analysis where we replicate more standard methods. That is, we examine changes in the aggregate distribution of temperatures over multiweek and multimonth intervals before we show results from our new approach that calculates responses for individual days. The standard analysis readily shows that temperature distributions in the CESM ensemble change over the RCP8.5 scenario (Figs. 4–6, which compare the “initial” and “final” time windows 1850–64 and 2086–2100). Means uniformly shift to warmer temperatures, but the distributions also change in terms of standard deviations and skewness. Figure 4 shows initial and final distributions in our example locations for aggregated 15-day periods in winter and summer.

Regarding the spatial characteristics of temperature distributions, we see the expected strong decrease in

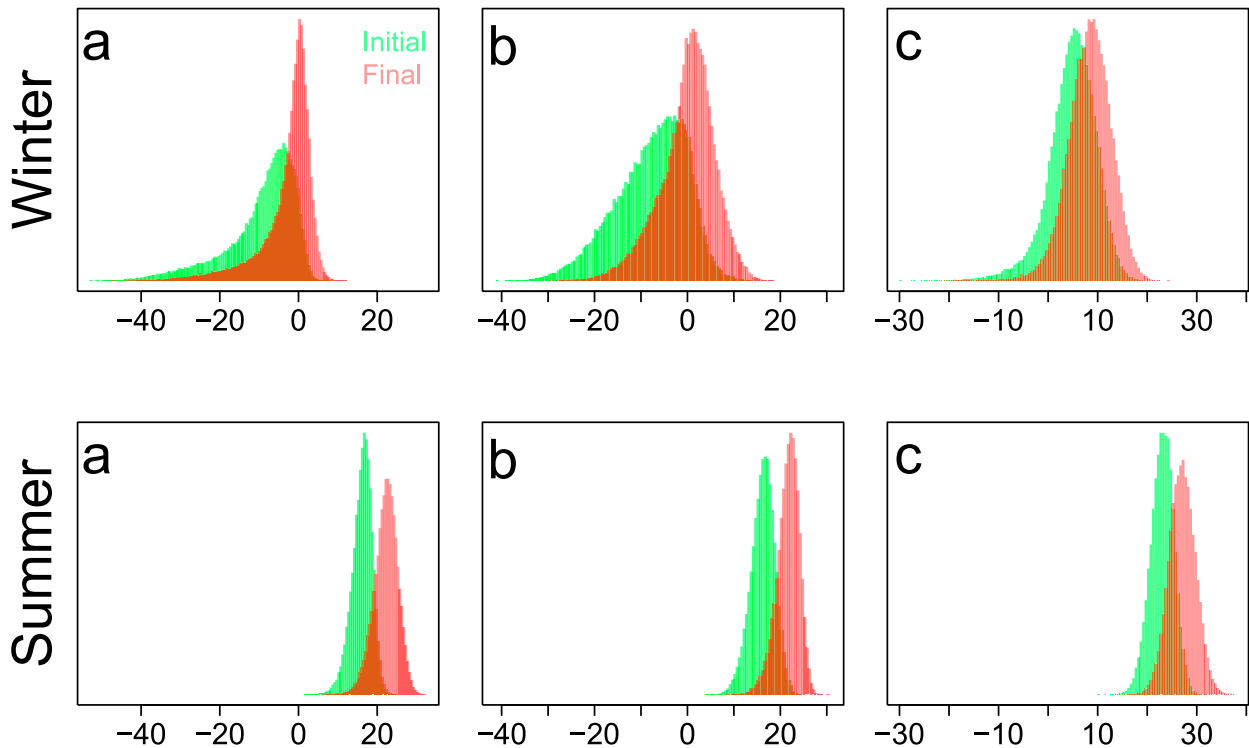


FIG. 4. Evolving distributions of daily mean temperature in the CESM ensemble RCP8.5 model runs at the locations a–c plotted in Fig. 2. Each distribution includes temperatures from a 15-day period over 15 model years for a total of 11 250 observations (15 days \times 15 years \times 50 ensemble members). Winter distributions are taken from 1 to 15 Jan and summer distributions from 5 to 19 Jul; “initial” distributions include years 1850–64 and “final” years 2096–2100. Changes in distributions are readily apparent, especially in winter at higher latitudes (locations a and b), but detailed quantification, especially of tail changes, requires more sophisticated techniques.

standard deviations in winter over land, especially in the northern midlatitudes (Figs. 5, 6). By contrast, summer standard deviation changes are much smaller and differ in sign in different locations. Temperature skewness (i.e., the asymmetry of the distribution) shows strong changes in winter over land in a dipole pattern. Winter temperature distributions are in all time periods negatively skewed throughout most of the domain, but in the north (including locations a and b), they become more negatively skewed in the future, while in the south (including location c), they become more symmetric. Summer skewness changes are again smaller and with less spatial coherence, other than the strong transitions in the southern Great Plains and in Mexico/Central America, where skewness in temperature distributions actually changes sign.

Our methodology for quantile estimation provides additional information that helps to quantify how temperature distributions are changing and to estimate the uncertainty associated with each change. We can evaluate not only bulk variability—the interquartile range (IQR), the difference between the 0.25 and 0.75 quantiles—but also differences between any two quantiles. In this

work, we typically consider the differences between the estimated 0.05 and 0.025 quantiles in the low tail and between the 0.975 and 0.95 quantiles in the high tail. These quantities measure tail variability in the same way that the IQR measures the variability of the bulk distribution.

If the skewness of a distribution changes over time, then future distributions are not simply scaled versions of present distributions. That is, their tail variabilities must change by a different factor than the IQR. In the case of the northern midlatitude winter temperatures shown in Fig. 5, where distributions become more negatively skewed as bulk variability decreases in the future, the effect could result from either a low tail contracting less than the bulk (or actually increasing) or a high tail contracting more than the bulk. Our methodology allows for differentiating these cases.

To assess whether the high tail and/or the low tail is driving changes in skewness, we consider the fractional changes in low, high, and bulk variabilities. For a given day of the year, we denote the initial and final quantile differences as $\Delta q_{r,i}$ and $\Delta q_{r,f}$, respectively, where $r = l, m$, or h indicates the low tail, middle of the

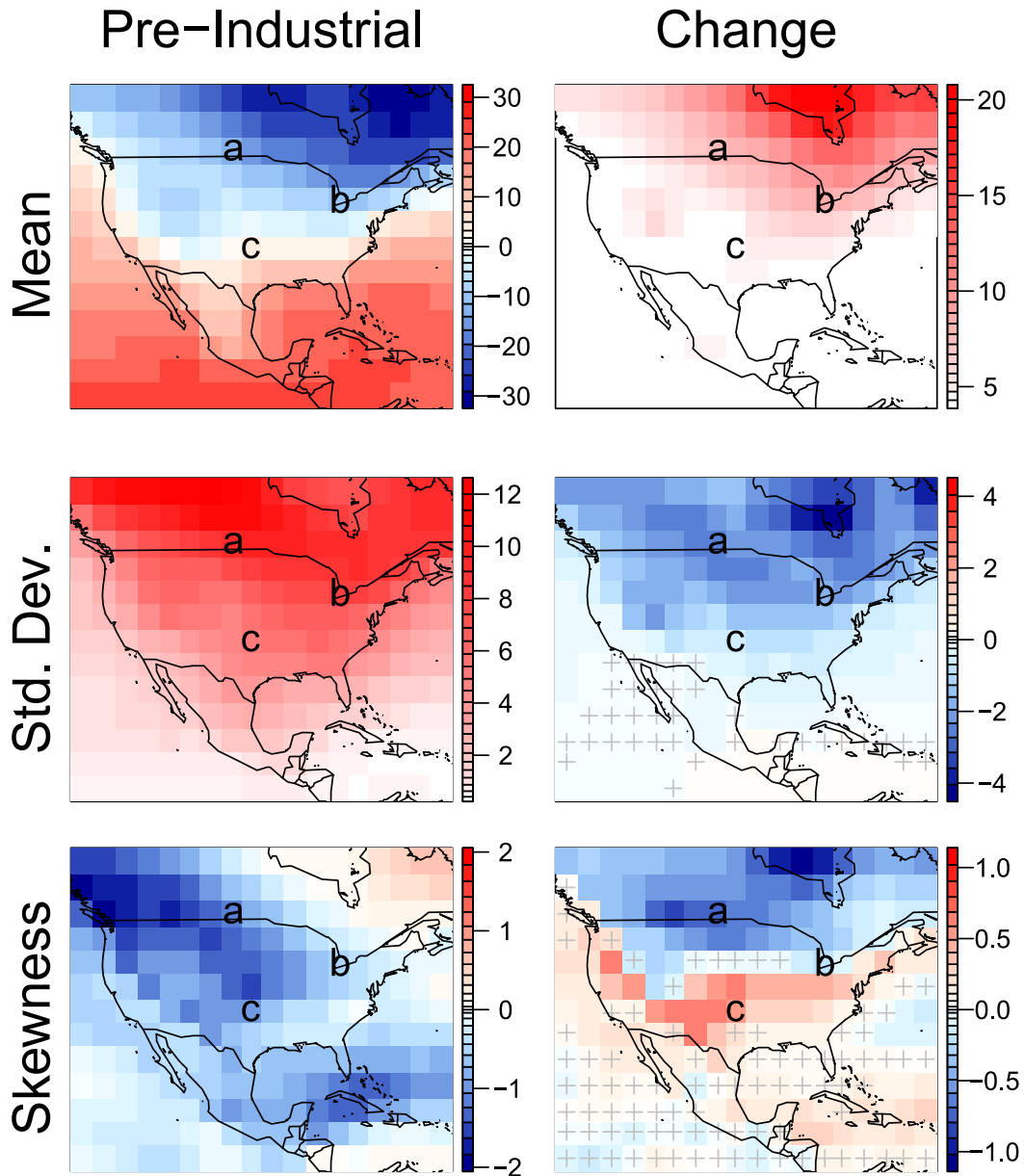


FIG. 5. (left) Initial temperature distribution properties and (right) their changes over time in the CESM ensemble RCP8.5 model runs for aggregate wintertime (DJF) daily temperature. Initial (preindustrial) and final periods are defined as in Fig. 4, as 15-yr periods 1850–64 and 2086–2100. Distributional moments (mean, standard deviation, and skewness) are defined as in Fig. 2. Gray crosses mark locations where the changes are not significant at the 0.05 level, obtained by resampling the set of 50 simulations (with replacement) and recalculating the sample moments. (top right) Mean temperature universally increases. Extreme warming in the Hudson Bay region occurs where the model is biased low in present-day simulations. (middle right) As expected, standard deviation decreases strongly at higher latitudes. (bottom right) Changes in winter skewness show a dipole pattern, which enhances negative skew above $\sim 40^\circ$ but reduces it at lower latitudes.

distribution (IQR), or high tail. The fractional variability change is then

$$\rho = \frac{\Delta q_{r,f} - \Delta q_{r,i}}{\Delta q_{r,i}}. \quad (3)$$

Because we model the complete temperature distribution for each day of the year for all years, we choose representative days and years to understand changes in winter and summer temperature distributions from beginning to end of the climate scenario: 1 January and

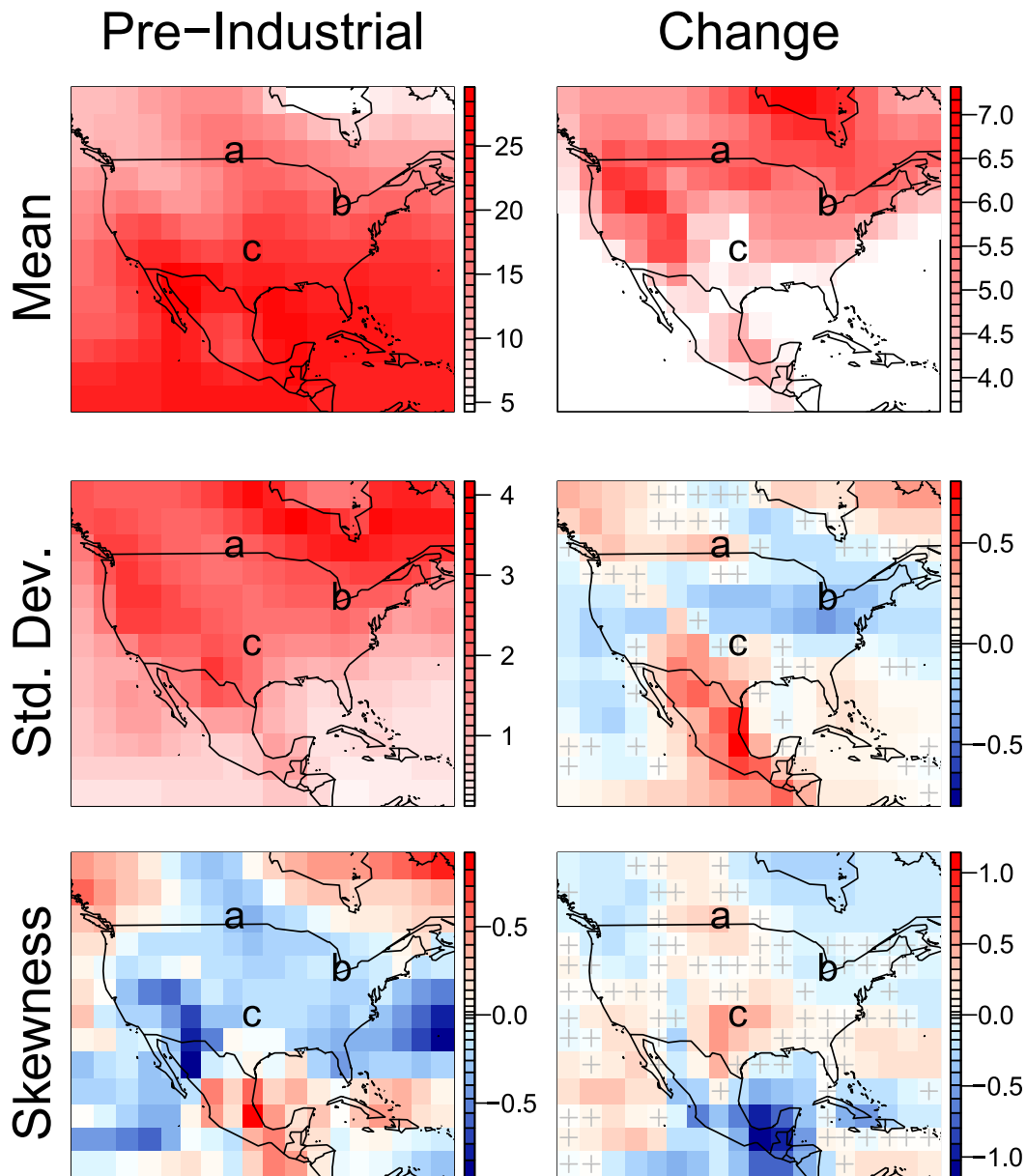


FIG. 6. As in Fig. 5, but for aggregate summer (JJA) temperatures; note that scales differ from those in Fig. 5. Except in the desert Southwest and Mexico, changes in (middle right) standard deviation and (bottom right) skewness are generally smaller in summer than in winter and often not significant at the 0.05 level.

5 July, evaluated in the years 1850 and 2100. For these representative days, we show in Fig. 7 the fractional variability change ρ for low and high tails as well as for the IQR.

Results show that tail changes can indeed differ strongly from changes in the bulk of the distribution. In wintertime (Fig. 7, top row), in much of the northern midlatitudes (including locations a and b), low tails change in a way that contributes to a more negative skew. Low-tail variability contracts less than does the

IQR, while high-tail variability contracts more strongly. High tails would contribute to more negative winter skew predominantly in the Hudson Bay region, where the model shows distinct bias. In summertime (Fig. 7, bottom row), the high tail dominates the transition to positive skew in the southern Great Plains region (including location c).

To clarify the relative contributions of high and low tails to skewness changes, we also examine evolving temperature variability in the bulk and tails as a function

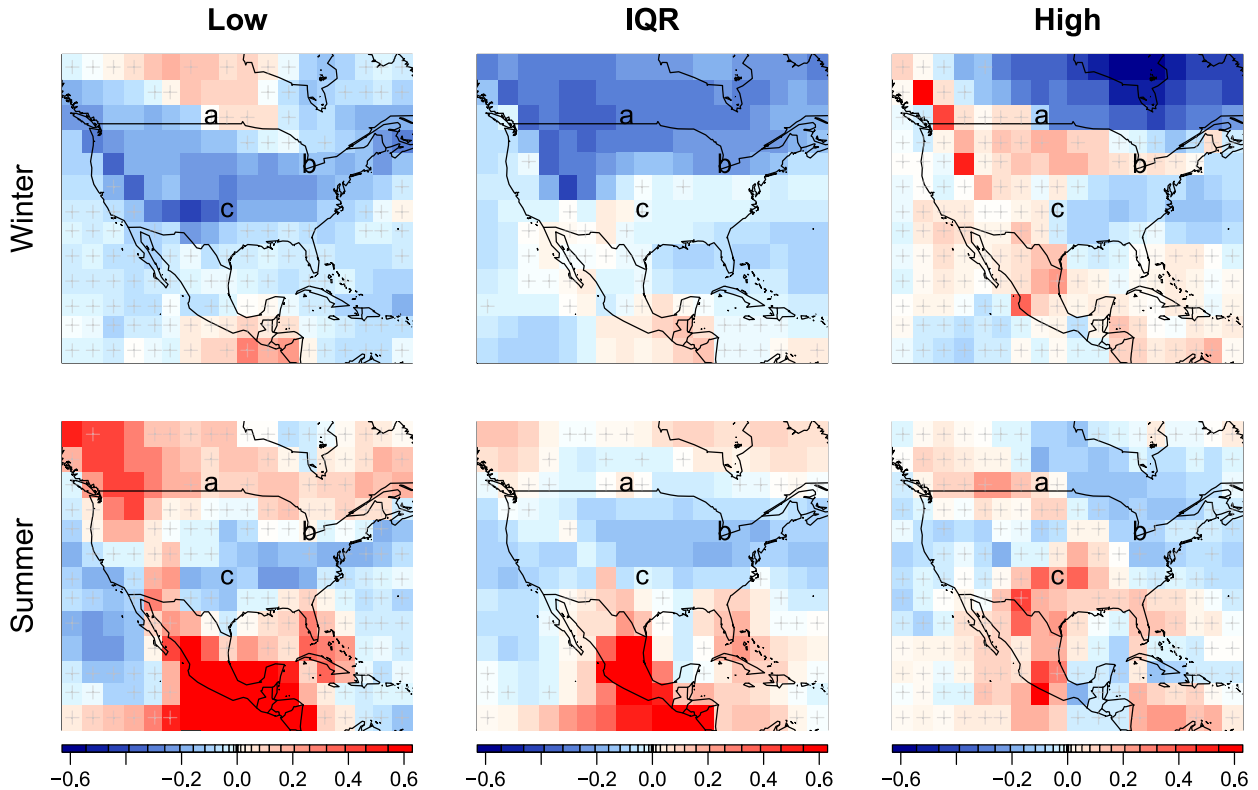


FIG. 7. Changes in daily temperature variability (quantile differences) over time in CESM ensemble RCP8.5 runs estimated using our statistical approach. Because our approach removes the need to aggregate over time when presenting changes, we show here differences in distributions for a single day and year: (top) 1 Jan for winter and (bottom) 5 Jul for summer, with differences evaluated between the years 1850 and 2100. Changes are expressed as fractions of initial variability [see Eq. (3) for definition], so that the value 0 indicates no change with respect to the initial year. (left) Changes in low-tail variability, (center) IQR, and (right) high-tail variability, as previously defined. Gray crosses mark grid points where the change is fewer than three standard deviations from the original estimate. As expected, estimated changes in IQR are similar to changes in standard deviation seen in Figs. 5 and 6. Changes in tail variability are clearly different from those in IQR, meaning that future distributions are not simply a rescaling of the present-day distributions.

of seasonality as well as long-term change. Figure 8 shows absolute variability changes for the three example locations a–c estimated using our quantile model; for fractional changes, see Fig. S6. The uncertainty around our estimates is quantified using the resampling approach described in the previous section. In all locations, wintertime skewness changes are driven by the relative changes in IQR and low tails. In the higher-latitude locations a and b, more negative winter skew results because the IQR contracts even more strongly than does the low-tail variability. In the lower-latitude location c, more positive winter skew results because the IQR changes slightly while the low-tail variability contracts strongly.

The complexity of the relationships in Fig. 8 also shows how misleading it may be to use a 3-month block to represent a season. While all three locations show larger IQR in winter than summer, the transition from winter to summer occurs more quickly at some locations than at others: abruptly in the northernmost location a, but more gradually in the southernmost location c. Similarly,

seasonal transitions in low-tail variability occur even more abruptly than those in IQR in locations a and b, but more gradually in location c. The trends over time show substantial decreases in wintertime IQR at locations a and b but not location c, whereas location c shows the largest decreases in wintertime low-tail variability. The transition around day 280 at location c between times of year with almost no change in low-tail variability and substantial decreases in low-tail variability is quite dramatic. In contrast, high-tail variability is relatively constant over both season and year at all sites.

While we show only three locations in the text here, an online interactive application allows similar in-depth examination of changes in model temperature distributions at all locations within North America (<http://statvm1.cs.uchicago.edu/mahaugen/extremes/>). The application allows the user to browse through any desired location to see how the variability changes as a function of season, year, and quantile difference. We include temperature histograms of the first and last simulation

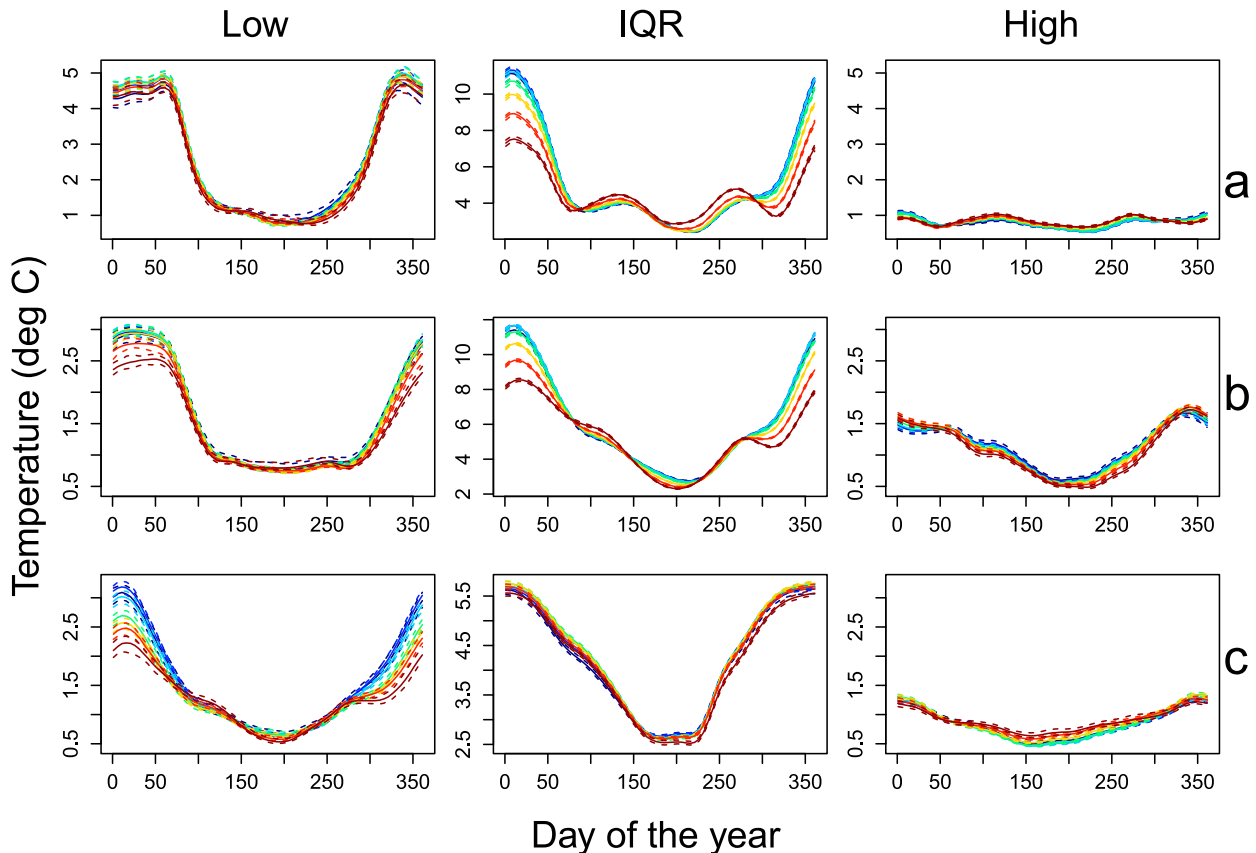


FIG. 8. Evolving daily temperature variability (quantile differences) over time in CESM ensemble RCP8.5 runs estimated using our statistical approach for locations a–c. Using the analysis described in Fig. 7, we show absolute IQR and tail variability as a function of seasonality, with different years (at 40-yr intervals) shown as different colored lines from 1850 (dark blue) to 2090 (dark red). Dashed lines represent pointwise 90% confidence intervals. Note the complexity of seasonal cycles in variability at different locations. These results show that the dipole pattern of changes in wintertime skewness changes seen in Fig. 5 is driven by low- rather than high-tail behavior. In wintertime, in the more northern locations (a and b), IQR reduces more strongly than does low-tail variability, making skew more negative. In the more southern location (location c), IQR change is negligible while low-tail variability reduces strongly, making skew more positive. In all locations, absolute changes in wintertime low-tail variability are larger than changes in high tails. For fractional changes, see Fig. S6.

years for the designated location, as well as maps that show the variability change spatially.

The seasonal evolution of temperature distributions is of strong interest in biology, because this evolution plays a large role in driving the seasonal responses of plant and animal species (their “phenology”). Models of phenologic responses are complex, but often involve threshold crossings such as the last freeze date of spring (e.g., Schwartz et al. 2006; Basler 2016, and references therein). Observational studies suggest that phenological transitions are already occurring at earlier dates and may be becoming more variable year to year (Pearse et al. 2017). The quantile estimation method described here allows evaluating a related index, the day of year on which a given quantile crosses some threshold temperature. Figure 9 shows the evolution over time of the first day of the year on which the 0.05, 0.25, and 0.5 quantiles cross

above -2.2° and 0°C in location b (near Detroit). That is, we are estimating the first day on which only 5%, 25%, or 50% of days would be below each near-freezing threshold temperature. Over time, all threshold days advance earlier in the season (for the 0.05 quantile, ~ 1.5 days decade $^{-1}$ on average between 2000 and 2100), with accelerating rates of advance. The higher quantiles advance more quickly. In particular, by the year 2080, at this location there are no more winter days with over 50% probability of temperature below -2.2°C . Determining these seasonal advances requires an approach like that presented here that allows quantile functions to vary smoothly day to day.

5. Conclusions

We present a method to quantify evolving temperature distributions in ensembles of climate model simulations

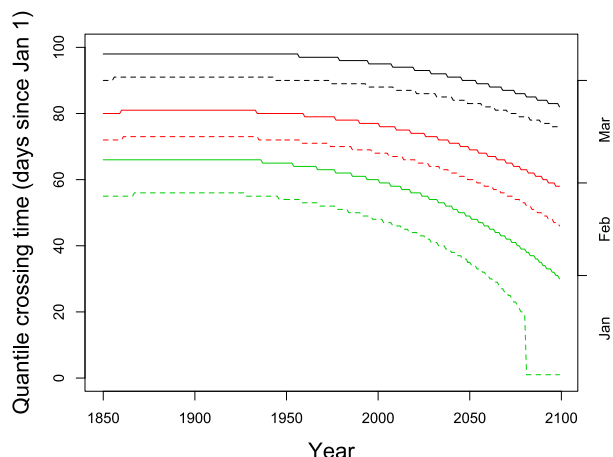


FIG. 9. First day of year three fitted quantiles of average daily temperature are above freezing (solid lines) and -2.2°C (dashed lines) for each year from 1850 to 2100 at location b. The quantiles pictured are 0.5 (green), 0.25 (red), and 0.05 (black); that is, the days shown are the first for which the probability of exceeding a temperature threshold is at least 5%, 25%, and 50%, respectively. The -2.2°C threshold is chosen for consistency with Schwartz et al. (2006). When these curves are re-estimated using resampled simulations, results generally change by at most 1 day, so no uncertainties are shown.

that involves analyzing the whole year simultaneously, rather than binning entire seasons and treating all days within a season as identically distributed. We use quantile regression with separate basis functions for seasonal effects, long-term trends, and long-term changes in seasonality. Using a 50-member ensemble of the CESM climate model (Srивer et al. 2015; Hogan and Srивer 2017; Vega-Westhoff and Srивer 2017), we demonstrate that the method captures subtle details of changing seasonal transitions. For example, we show that in CESM, at a site in the southern plains, low-tail variability hardly changes between 1850 and 2090 during April–September but decreases quite substantially by late October and then throughout most of the winter.

The abundance of data available in large single model ensembles allows using quantile regression to estimate high quantiles accurately within a single model structure, avoiding assumptions about the shape of the tail of the distribution that are required to apply extreme value theory. The use of initial condition ensembles also allows estimation of uncertainties without strong assumptions. By resampling entire simulations from the ensemble of climate simulations and recalculating the quantiles, we obtain confidence bands that do not require any assumptions of independence within an ensemble member, only independence across ensemble members. That is, we do not need to model temporal dependence within model runs. We show that the

smooth quantile estimates are accurate even across small intervals of the domain of the predictors. The fidelity of these intervals serves as a criterion to determine the required complexity in the statistical model.

The techniques presented in this study replicate several prior conclusions made in the literature: for example, the well-known projected decrease in winter variability in the continental northern midlatitudes (e.g., Schneider et al. 2015), likely due to amplified warming in the Arctic (Screen 2014). The use of differences of quantiles to measure tail variability provides a tool for studying more nuanced changes in distributions. In the case study analyzed here, we relate the changes in tail variability to changes in skewness of the temperature distributions, and we find that in most of the domain analyzed, wintertime skewness changes are driven largely by the relative behavior of IQR and low tails. For example, in much of the continental northern United States and Canada, the low tail of temperature contracts substantially less than does the overall temperature variability.

While the use of quantile regression here is natural for an analysis that focuses on quantiles and quantile differences, it is possible to estimate quantiles using alternative methods. An alternative approach that loses within-season detail might be to evaluate empirical distributions of temperatures within seasonal blocks for successive time windows (e.g., decades) and use these results to study how quantile differences evolve over time. The quantile regression approach described here enables the study of seasonal transitions with a flexible framework that allows different combinations of basis functions for seasonality, long-term trends, and changes in seasonality as appropriate for different datasets. While we analyze only temperature here, our method is intended to be general enough to be applied to other climate variables such as precipitation or humidity. These detailed insights into climate variable distributions may be valuable for risk assessment studies that emphasize extreme events.

Acknowledgments. This work was supported in part by STATMOS, the Research Network for Statistical Methods for Atmospheric and Oceanic Sciences (NSF-DMS Awards 1106862, 1106974, and 1107046), and by RDCEP, the University of Chicago Center for Robust Decision-making on Climate and Energy Policy (NSF-SES Awards 0951576 and 146364). We acknowledge the University of Chicago Research Computing Center, whose resources were used in the completion of this work. Ryan L. Srивer acknowledges support from the Department of Energy sponsored Program on Integrated Assessment Model Development, Diagnostics

and Inter-Model Comparisons (PIAMDDI), and the Program on Coupled Human Earth Systems (PCHES).

APPENDIX A

Model and Reanalysis Comparisons

Following the discussion in the paper, we define sample mean, variance, and skewness as

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i, \\ s^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{and} \\ \gamma &= \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3.\end{aligned}\quad (\text{A1})$$

These definitions are used in Figs. 2, 5, and 6 in the main text and in Figs. S1 and S2 in the online supplemental material. We plot the standard deviation s rather than the variance s^2 .

APPENDIX B

Model Details

In the following, we first give details regarding the regression of temperature quantiles on a fixed set of basis functions. We then discuss how to select the number of basis functions through a “sufficiency criterion.” Last, we describe how we quantify uncertainty in the quantile estimates.

a. Model estimation

Given the number of basis functions in our model, represented by the columns in a matrix \mathbf{X} with number of rows equal to the number of observations in the dataset, we construct our temperature quantile estimate $\hat{\mathbf{T}}_q$ and corresponding coefficients $\hat{\boldsymbol{\beta}}_q$, namely,

$$\hat{\mathbf{T}}_q = \mathbf{X}\hat{\boldsymbol{\beta}}_q, \quad (\text{B1})$$

such that the q th fraction of residuals between the observations \mathbf{T} at a particular location and their estimates $\mathbf{T} - \hat{\mathbf{T}}_q$ is greater than zero, and a fraction $1 - q$ is less than zero. With the temperature model in Eq. (2), our coefficient vector estimate $\hat{\boldsymbol{\beta}}$ contains the estimates of α , a_i , b_j , c_{ij} . Note that the seasonal interaction terms corresponding to the coefficients c_{ij} are not necessarily the same as the main seasonal terms corresponding to a_i . In fact, we find that fewer seasonal interaction terms are needed to describe the interaction behavior.

Computationally, obtaining $\hat{\boldsymbol{\beta}}$ is equivalent to solving the following optimization problem (Koenker and Bassett 1978):

$$\min_{\boldsymbol{\beta}} \left[\sum_{d,t: \mathbf{T}(d,t) \geq \mathbf{X}(d,t)\boldsymbol{\beta}} q |\mathbf{T}(d,t) - \mathbf{X}^T(d,t)\boldsymbol{\beta}| + \sum_{d,t: \mathbf{T}(d,t) < \mathbf{X}(d,t)\boldsymbol{\beta}} (1-q) |\mathbf{T}(d,t) - \mathbf{X}^T(d,t)\boldsymbol{\beta}| \right], \quad (\text{B2})$$

and can be implemented in either R or MATLAB using existing libraries.¹ Because we have access to 50 simulations, each location provides us with $365 \times 250 \times 50$ or approximately 4.5 million observations. Consequently, even fairly high quantiles can be accurately estimated without borrowing data from neighboring locations through a spatial model (as done by, e.g., Reich et al. 2011). However, making inferences about more extreme quantiles, such as the 0.001 or 0.999 quantiles, cannot be guaranteed to work as well with our methods.

We do not experience issues with quantile estimates crossing in our study area even though the optimization framework above does not explicitly enforce monotonicity with increasing quantile estimates. The absence of crossing quantiles is likely also due to the large sample size. [For strict enforcement of monotonicity in the quantile curves, see, e.g., Bondell et al. (2010).]

b. Model selection

We describe our approach to selecting a modest set of basis functions that can accurately represent the temperature data. If the model chosen has too many basis functions, we run the risk of overfitting out-of-sample observations. To make sure this does not happen, we need a metric to quantify the goodness-of-fit of the model.

Any reasonable temperature model we fit to the data will by definition contain the desired amount of positive and negative residuals *globally* according to the desired quantile q . A more stringent requirement would be that the smooth temperature estimate contains approximately an appropriate fraction of positive and negative residuals on a *daily* basis: for each d and t ,

$$S(d,t) = \frac{1}{n} \sum_{i=1}^n I[\hat{T}_i(d,t) - T_i(d,t) > 0] \approx q, \quad (\text{B3})$$

¹ We use the R library `rq` and the function `rq.fit.pfn`, developed by Portnoy and Koenker (1997). Basis functions are created using `pbs` for periodic spline basis functions and `ns` for nonperiodic splines. The nonperiodic splines are constrained to be linear beyond the domain, 1850–2100, and are called *natural splines*.

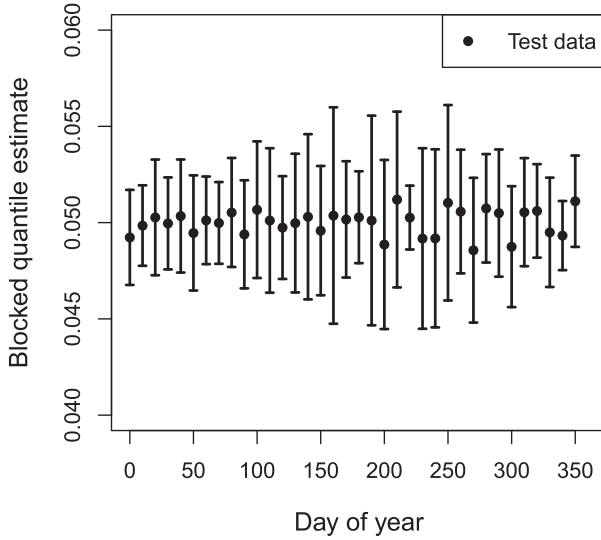


FIG. B1. Exceedance probability of temperature events above the 0.95 quantile estimate. The density is obtained by making 10-day bins and counting the number of observations that are above the quantile estimate and normalizing by the total number of exceedances aggregated across all model runs. Each bin is represented by the bin start day; that is, an x -axis value of 0 includes the interval $(0,10]$. We hold out 10 different sets of simulations to obtain 10 different estimates for each block of time, from which we calculate their mean shown as points and standard deviation shown as error bars around \hat{S}_{test} as defined in Eq. (B4).

where I is the indicator function, and n is the total number of samples (i.e., 50 for our CESM ensemble dataset). If $S(d, t)$ is close to the value q for each d and t , the model would accurately describe the data, and the number of basis functions is sufficient. In fact, as we will describe below, we average $S(d, t)$ over blocks of days or years to produce more stable values of this statistic.

To estimate the appropriate number of basis functions, we hold out five simulations from the fitting process and use these to calculate our exceedances, which we call $S_{\text{test}}(d, t)$. We repeat this 10 times so that all the simulations are eventually held out, giving 10 samples of $S_{\text{test}}(d, t)$. As we increase model complexity through the number of basis functions, the variability of S_{test} should reach a minimum when the necessary number of basis functions is reached and the quantile estimate is the same for each time point. If the number of basis functions is increased beyond this point, we start to overfit the data, and the out-of-sample variability of S_{test} will increase.

To estimate S_{test} , we block the variables in two ways, one for each variable. First, we divide each year in 10-day bins and calculate the average exceedance estimate \hat{S}_{test} in each bin. We sum over the whole domain of long-term change t and a subset of the seasonality variable d . Specifically, let A be a set of nonoverlapping contiguous

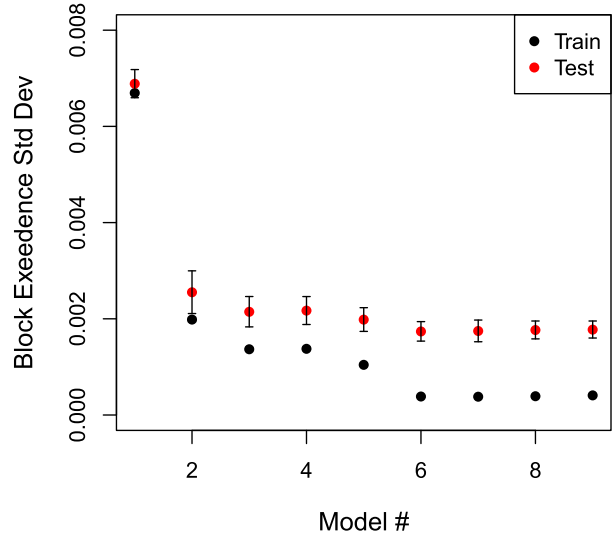


FIG. B2. Training and test exceedance standard deviation as a function of model number (see Table B1), where increasing model number signifies increasing degrees of freedom in the spline basis functions. The data were extracted from the grid box located at $(\text{lat}, \text{lon}) = (31.5^\circ, -93.8^\circ)$. The exceedance is calculated by binning seasonality in 10-day blocks and summing over the long-term change [see Eq. (B4)].

blocks of days that together cover the whole year, where $a_j, j = 1, \dots, m$ are the elements of the set. Also, let K be the index set for long-term change $K = [1850, 2100]$, measured in years. Then, for all $a_j \in A$,

$$\hat{S}_{\text{test}}(a_j) = \frac{1}{n} \sum_{i \in [1, n], d \in a_j, t \in K} I[\hat{T}_i(d, t) - T_i(d, t) > 0]. \quad (\text{B4})$$

To get an equal number of days in each bin, we use the first 360 days of the year only.

Second, we divide the long-term change variable t in bins and repeat the process by flipping the role of the variables in Eq. (B4) to get a set of $\hat{S}_{\text{test}}(b_j)$ with $b_j \in B$, a set of nonoverlapping contiguous blocks of long-term change indices in K . An example of the blocked exceedance estimate is shown in Fig. B1. Note that the pointwise quantile estimate is contained between the error bars, suggesting that the model is sufficiently complex. The standard deviation of these estimates of \hat{S}_{test} is our measure of exceedance variability.

We seek the simplest model that gives good calibration of the quantile estimates (so close to 0.05 in Fig. B1). At the same time, we have to watch out to not overfit the data, so we also want to minimize out-of-sample variability. We find that a model with 15 seasonal, three seasonal interaction, and four temporal degrees of freedom minimizes the variability of exceedances \hat{S}_{test} , shown in Fig. B2, where seasonality has

TABLE B1. Degrees of freedom in the spline basis for each independent variable, with the number of seasonal basis functions in the interaction term [the number of h_i terms in Eq. (1)] given by the “season interaction” column. The temporal polynomials are the same in both the main effect and interaction terms [the g_j and s_j terms in Eq. (1)].

Model	Seasonal	Season interaction	Temporal
1	5	3	3
2	7	3	3
3	10	3	3
4	10	3	4
5	12	3	4
6	15	3	4
7	15	3	5
8	15	5	5
9	18	5	5

been binned. The out-of-sample fit when binning by years is shown in Fig. S7. Here, models 4–6 have approximately equal test error, so since binning seasonality suggests the complexity of model 6, we chose model 6 as the overall model. Including the interaction terms, the full model has 32 free parameters to be fitted, or $\beta \in \mathbb{R}^{32}$. All model candidates are shown in Table B1. We reach the same conclusion when blocking the long-term change t and when analyzing different spatial locations (see Fig. S7).

c. Uncertainty estimation

With a reasonable model chosen through cross-validation, we present a way to quantify its uncertainty. Because we are using multiple simulations that are assumed independent, we resample entire simulations from the set of 50 simulations. Resampling 50 new simulations with replacement from the original set of simulations yields a new dataset. From the new dataset, we obtain another temperature estimate with the same model basis functions but different coefficients β^* . After repeating this resampling and re-estimating procedure 100 times, we generate pointwise confidence intervals for temperature quantiles. For example, in Fig. 8, we show the 90% confidence interval by selecting the pointwise 0.05 and 0.95 quantiles of temperature variability estimates. Because the confidence intervals are quite tight, we deem the 100 new estimates (or bootstraps) sufficient to indicate that the results we describe in section 4 are not due to random variation. Larger numbers of bootstrap replicates might give slightly more accurate intervals but would not change our conclusions. One might also consider fewer simulations as a compromise between computation time and quality of the estimates. For n simulations, we would expect the standard error to scale as $1/\sqrt{n}$. Thus, if one is willing to widen the confidence intervals by a factor of 2,

(approximately) only 12 simulations would suffice. However, one could compensate for this greater variability by using fewer basis functions at a cost, of course, of obtaining less-resolved estimates of seasonal patterns and long-term trends in the quantiles.

REFERENCES

- Basler, D., 2016: Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe. *Agric. For. Meteorol.*, **217**, 10–21, <https://doi.org/10.1016/j.agrformet.2015.11.007>.
- Bondell, H. D., B. J. Reich, and H. Wang, 2010: Noncrossing quantile regression curve estimation. *Biometrika*, **97**, 825–838, <https://doi.org/10.1093/biomet/asq048>.
- Chapman, S. C., D. A. Stainforth, and N. W. Watkins, 2013: On estimating local long-term climate trends. *Philos. Trans. Roy. Soc. A*, **371**, 20120287, <https://doi.org/10.1098/rsta.2012.0287>.
- Chavez-Demoulin, V., and A. C. Davison, 2005: Generalized additive modelling of sample extremes. *J. Roy. Stat. Soc. C*, **54**, 207–222, <https://doi.org/10.1111/j.1467-9876.2005.00479.x>.
- Davison, A. C., and R. L. Smith, 1990: Models for exceedances over high thresholds. *J. Roy. Stat. Soc. B*, **52**, 393–442.
- , S. A. Padoan, and M. Ribatet, 2012: Statistical modeling of spatial extremes. *Stat. Sci.*, **27**, 161–186, <https://doi.org/10.1214/11-STS376>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Deser, C., R. Knutti, S. Solomon, and A. S. Phillips, 2012a: Communication of the role of natural variability in future North American climate. *Nat. Climate Change*, **2**, 775–779, <https://doi.org/10.1038/nclimate1562>.
- , A. Phillips, V. Bourdette, and H. Teng, 2012b: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, <https://doi.org/10.1007/s00382-010-0977-x>.
- , —, M. A. Alexander, and B. V. Smoliak, 2014: Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *J. Climate*, **27**, 2271–2296, <https://doi.org/10.1175/JCLI-D-13-00451.1>.
- Diffenbaugh, N. S., D. L. Swain, and D. Touma, 2015: Anthropogenic warming has increased drought risk in California. *Proc. Natl. Acad. Sci. USA*, **112**, 3931–3936, <https://doi.org/10.1073/pnas.1422385112>.
- Donat, M. G., and L. V. Alexander, 2012: The shifting probability distribution of global daytime and night-time temperatures. *Geophys. Res. Lett.*, **39**, L14707, <https://doi.org/10.1029/2012GL052459>.
- Eastoe, E. F., and J. A. Tawn, 2009: Modelling non-stationary extremes with application to surface level ozone. *J. Roy. Stat. Soc. C*, **58**, 25–45, <https://doi.org/10.1111/j.1467-9876.2008.00638.x>.
- Fischer, E. M., and C. Schär, 2009: Future changes in daily summer temperature variability: Driving processes and role for temperature extremes. *Climate Dyn.*, **33**, 917, <https://doi.org/10.1007/s00382-008-0473-8>.
- , and R. Knutti, 2014: Detection of spatially aggregated changes in temperature and precipitation extremes. *Geophys. Res. Lett.*, **41**, 547–554, <https://doi.org/10.1002/2013GL058499>.

- Franzke, C. L., 2015: Local trend disparities of European minimum and maximum temperature extremes. *Geophys. Res. Lett.*, **42**, 6479–6484, <https://doi.org/10.1002/2015GL065011>.
- Gao, M., and C. L. Franzke, 2017: Quantile regression-based spatiotemporal analysis of extreme temperature change in China. *J. Climate*, **30**, 9897–9914, <https://doi.org/10.1175/JCLI-D-17-0356.1>.
- Hagos, S. M., L. R. Leung, J.-H. Yoon, J. Lu, and Y. Gao, 2016: A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the Large Ensemble CESM simulations. *Geophys. Res. Lett.*, **43**, 1357–1363, <https://doi.org/10.1002/2015GL067392>.
- Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning*. 2nd ed. Springer, 745 pp.
- 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>.
- 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>.
- Huser, R., and A. Davison, 2014: Space–time modelling of extreme events. *J. Roy. Stat. Soc. B*, **76**, 439–461, <https://doi.org/10.1111/rssb.12035>.
- Huybers, P., K. A. McKinnon, A. Rhines, and M. Tingley, 2014: U.S. daily temperatures: The meaning of extremes in the context of nonnormality. *J. Climate*, **27**, 7368–7384, <https://doi.org/10.1175/JCLI-D-14-00216.1>.
- Jalbert, J., A.-C. Favre, C. B  lisle, and J.-F. Angers, 2017: A spatiotemporal model for extreme precipitation simulated by a climate model, with an application to assessing changes in return levels over North America. *J. Roy. Stat. Soc. C*, **66**, 941–962, <https://doi.org/10.1111/rssc.12212>.
- Katz, R. W., and B. G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, **21**, 289–302, <https://doi.org/10.1007/BF00139728>.
- 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>.
- Kitoh, A., and T. Mukano, 2009: Changes in daily and monthly surface air temperature variability by multi-model global warming experiments. *J. Meteor. Soc. Japan Ser. II*, **87**, 513–524, <https://doi.org/10.2151/jmsj.87.513>.
- Koenker, R., and G. Bassett Jr., 1978: Regression quantiles. *Econometrica*, **46**, 33–50, <https://doi.org/10.2307/1913643>.
- McKinnon, K. A., A. Rhines, M. P. Tingley, and P. Huybers, 2016: The changing shape of Northern Hemisphere summer temperature distributions. *J. Geophys. Res. Atmos.*, **121**, 8849–8868, <https://doi.org/10.1002/2016JD025292>.
- Meehl, G. A., C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, 2009: Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophys. Res. Lett.*, **36**, L23701, <https://doi.org/10.1029/2009GL040736>.
- 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>.
- Northrop, P. J., and P. Jonathan, 2011: Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics*, **22**, 799–809, <https://doi.org/10.1002/env.1106>.
- Otto, F. E. L., N. Massey, G. J. Oldenborgh, R. G. Jones, and M. R. Allen, 2012: Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophys. Res. Lett.*, **39**, L04702, <https://doi.org/10.1029/2011GL050422>.
- Pearse, W. D., C. C. Davis, D. W. Inouye, R. B. Primack, and T. J. Davies, 2017: A statistical estimator for determining the limits of contemporary and historic phenology. *Nat. Ecol. Evol.*, **1**, 1876–1882, <https://doi.org/10.1038/s41559-017-0350-0>.
- Poppick, A., E. J. Moyer, and M. L. Stein, 2017: Estimating trends in the global mean temperature record. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **3**, 33–53, <https://doi.org/10.5194/ascmo-3-33-2017>.
- Portnoy, S., and R. Koenker, 1997: The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Stat. Sci.*, **12**, 279–300, <https://doi.org/10.1214/ss/1030037960>.
- R  is  nen, J., 2002: CO₂-induced changes in interannual temperature and precipitation variability in 19 CMIP2 experiments. *J. Climate*, **15**, 2395–2411, [https://doi.org/10.1175/1520-0442\(2002\)015<2395:CICIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2395:CICIT>2.0.CO;2).
- Reich, B. J., M. Fuentes, and D. B. Dunson, 2011: Bayesian spatial quantile regression. *J. Amer. Stat. Assoc.*, **106**, 6–20, <https://doi.org/10.1198/jasa.2010.ap09237>.
- Rhines, A., K. A. McKinnon, M. P. Tingley, and P. Huybers, 2017: Seasonally resolved distributional trends of North American temperatures show contraction of winter variability. *J. Climate*, **30**, 1139–1157, <https://doi.org/10.1175/JCLI-D-16-0363.1>.
- Rodgers, K. B., J. Lin, and T. L. Fr  licher, 2015: Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, **12**, 3301–3320, <https://doi.org/10.5194/bg-12-3301-2015>.
- Schneider, T., T. Bischoff, and H. Plotka, 2015: Physics of changes in synoptic midlatitude temperature variability. *J. Climate*, **28**, 2312–2331, <https://doi.org/10.1175/JCLI-D-14-00632.1>.
- Schwartz, M. D., R. Ahas, and A. Aasa, 2006: Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biol.*, **12**, 343–351, <https://doi.org/10.1111/j.1365-2486.2005.01097.x>.
- Screen, J. A., 2014: Arctic amplification decreases temperature variance in northern mid- to high-latitudes. *Nat. Climate Change*, **4**, 577–582, <https://doi.org/10.1038/nclimate2268>.
- Semenov, V., and L. Bengtsson, 2002: Secular trends in daily precipitation characteristics: Greenhouse gas simulation with a coupled AOGCM. *Climate Dyn.*, **19**, 123–140, <https://doi.org/10.1007/s00382-001-0218-4>.
- Shields, C. A., D. A. Bailey, G. Danabasoglu, M. Jochum, J. T. Kiehl, S. Levis, and S. Park, 2012: The low-resolution CCSM4. *J. Climate*, **25**, 3993–4014, <https://doi.org/10.1175/JCLI-D-11-00260.1>.
- Singh, D., and Coauthors, 2014: Severe precipitation in northern India in June 2013: Causes, historical context, and changes in probability [in “Explaining Extremes of 2013 from a Climate Perspective”]. *Bull. Amer. Meteor. Soc.*, **95** (9), S58–S61, <https://doi.org/10.1175/1520-0477-95.9.S1.1>.
- Smith, R., and Coauthors, 2010: The Parallel Ocean Program (POP) reference manual: Ocean component of the Community Climate System Model (CCSM) and Community Earth System Model (CESM). CESM Rep. LAUR-10-01853, 141 pp.

- Sriver, 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>.
- Stainforth, D. A., S. C. Chapman, and N. W. Watkins, 2013: Mapping climate change in European temperature distributions. *Environ. Res. Lett.*, **8**, 034031, <https://doi.org/10.1088/1748-9326/8/3/034031>.
- Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614, <https://doi.org/10.1038/nature03089>.
- Swain, D. L., M. Tsiang, M. Haugen, D. Singh, A. Charland, B. Rajaratnam, and N. S. Diffenbaugh, 2014: The extraordinary California drought of 2013/2014: Character, context, and the role of climate change [in “Explaining Extremes 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>.
- Thompson, D. W. J., E. A. Barnes, C. Deser, W. E. Foust, and A. S. Phillips, 2015: Quantifying the role of internal climate variability in future climate trends. *J. Climate*, **28**, 6443–6456, <https://doi.org/10.1175/JCLI-D-14-00830.1>.
- Trenberth, K. E., 2011: Attribution of climate variations and trends to human influences and natural variability. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 925–930, <https://doi.org/10.1002/wcc.142>.
- , J. T. Fasullo, and T. G. Shepherd, 2015: Attribution of climate extreme events. *Nat. Climate Change*, **5**, 725–730, <https://doi.org/10.1038/nclimate2657>.
- Vega-Westhoff, B., and R. L. 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>.
- Ylhäisi, J. S., and J. Räisänen, 2014: Twenty-first century changes in daily temperature variability in CMIP3 climate models. *Int. J. Climatol.*, **34**, 1414–1428, <https://doi.org/10.1002/joc.3773>.