Title
The changing shape of Northern Hemisphere summer temperature distributions

Journal
Journal of Geophysical Research: Atmospheres, 121(15)

ISSN
2169-897X

Authors
McKinnon, Karen A
Rhines, Andrew
Tingley, Martin P
et al.

Publication Date
2016-08-16

DOI
10.1002/2016jd025292

Peer reviewed

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD025292

The changing shape of Northern Hemisphere summer temperature distributions

Karen A. McKinnon[1], Andrew Rhines[2], Martin P. Tingley[3], and Peter Huybers[4]

[1]National Center for Atmospheric Research, Boulder, Colorado, USA, [2]Department of Atmospheric Sciences, University of Washington, Seattle, Washington, USA, [3]Department of Meteorology, Pennsylvania State University, University Park, Pennsylvania, USA, [4]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA

Key Points:
- Linear trends in summer temperature distributions from 1980 to 2015 are estimated using quantile regression
- Changes in the shape of the distributions can be summarized using four orthogonal basis functions
- Distributional changes are foremost explained by a positive shift, but significant regional changes in shape are also identified

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4

Correspondence to:
K. A. McKinnon,
mckinnon@ucar.edu

Citation:
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, doi:10.1002/2016JD025292.

Received 28 APR 2016
Accepted 14 JUL 2016
Accepted article online 18 JUL 2016
Published online 9 AUG 2016

Abstract The occurrence of recent summer temperature extremes in the midlatitudes has raised questions about whether and how the distributions of summer temperature are changing. While it is clear that in most regions the average temperature is increasing, there is less consensus regarding the presence or nature of changes in the shape of the distributions, which can influence the probability of extreme events. Using data from over 4000 weather stations in the Global Historical Climatology Network-Daily database, we quantify the changes in daily maximum and minimum temperature distributions for peak summer in the Northern Hemisphere midlatitudes during 1980–2015 using quantile regression. A large majority (87–88%) of the trends across percentiles and stations can be explained by a shift of the distributions with no change in shape. The remaining variability is summarized through projections onto orthogonal basis functions that are closely related to changes in variance, skewness, and kurtosis. North America and Eurasia show significant shifts in the estimated distributions of daily maximum and minimum temperatures. Although no general change in summer variance is found, variance has regionally increased in Eurasia and decreased in most of North America. Changes in shape that project onto the skewness and kurtosis basis functions have a much smaller spatial scale and are generally insignificant.

1. Introduction

The past 15 years have featured a number of high-impact summer temperature extremes in the Northern Hemisphere, including the 2003 and 2010 heat waves in Europe [e.g., *Schär et al.*, 2004; *Barriopedro et al.*, 2011] and the 2012 heat wave across the midwestern United States [e.g., *Karl et al.*, 2012]. In light of these events, there is increasing interest in determining whether and how the regional distributions of summer temperatures are changing. In particular, because changes in the shape of a distribution have a large influence on the probability of extreme events [*Mearns et al.*, 1984], it is important to identify whether the observed changes in distributions can be explained by a simple shift across all percentiles with no shape changes, or if the observations provide evidence for shape changes as well. While much of the trend in global land summer temperature variability appears well explained by a simple shift in the distribution [*Rhines and Huybers*, 2013], the changes in temperature distributions on a regional level can be more complex [*Huntingford et al.*, 2013].

A substantial body of work has focused on changes in the distribution of temperatures averaged over the summer season in Europe. Shortly after the 2003 heat wave, *Schär et al.* [2004] fit a normal distribution to the average June, July, and August temperatures in Switzerland, near the center of the heat wave, and determined that the event required an increase in both the variance and the mean. A similar conclusion was drawn by *Barriopedro et al.* [2011] after the hot European summer of 2010. Both studies, however, focused on the distribution of temperature averaged across the season and/or continent, which obscures distributional changes on smaller spatial and temporal scales relevant for many extreme events [*Alexander and Perkins*, 2013].

Somewhat different results have been found by researchers examining daily station data within Europe. Across studies and regions, there is generally an identifiable and significant increase in mean summer temperatures, but results regarding changes in the higher-order moments of the distributions are less consistent. Based on fitting extreme value distributions, *Della-Marta et al.* [2007] found an increase in central-western European summer temperature variance but no significant changes in Sweden or the Iberian Peninsula.

Acero et al. [2014] also failed to find significant changes in variance in the Iberian Peninsula. Trends in Italian summer temperatures also appear consistent with a simple shift across all percentiles [*Simolo et al.*, 2010].

Across the Northern Hemisphere midlatitudes as a whole, there is some evidence of a decrease in variance of all-season daily temperatures, as well as an increase in skewness [*Donat and Alexander*, 2012]. Within in the summer season, however, there is substantial regional variability with respect to trends in the moments of daily summer temperature over the past 60 years, including decreases in variance across much of North America and increases in central Europe [*Cavanaugh and Shen*, 2014], although the interannual variability of summer season temperatures across the hemisphere as a whole does not appear to be increasing [*Huntingford et al.*, 2013].

Due to the wide variety of methods applied to analyzing changes in temperature distributions, as well as the relatively small number of studies using daily data, it becomes difficult to draw general conclusions about the spatial structure of changes in daily temperature distributions. While some studies use gridded or otherwise spatially averaged data [*Schär et al.*, 2004; *Barriopedro et al.*, 2011; *Huntingford et al.*, 2013; *Brown et al.*, 2008; *Donat and Alexander*, 2012], others use point data from individual stations [*Della-Marta et al.*, 2007; *Acero et al.*, 2014; *Simolo et al.*, 2010; *Cavanaugh and Shen*, 2014]. Spatially averaging point data alters the higher-order moments of the distribution [*Haylock et al.*, 2008; *Director and Bornn*, 2015] and can induce spurious trends if station density is allowed to change [*Brown et al.*, 2008; *Rhines and Huybers*, 2013]. Furthermore, different studies make various assumptions about both the underlying distribution of temperature and how the distribution is changing that can influence inferred changes in shape. In particular, the presence of trends increases the sample variance over a given period such that differing trends in the mean could appear as changes in variance if they are not properly accounted for, as has been demonstrated for temperature data [e.g., *Rhines and Huybers*, 2013; *Huntingford et al.*, 2013].

Here we analyze daily summer temperature data using a single method across the Northern Hemisphere midlatitudes in order to gain a more complete picture of the spatial structure of changes in summer temperature distributions. Data are retained at the station level to avoid artifacts from interpolation or gridding. The changes in the distribution are cleanly partitioned into a shift across all percentiles, as well as three additional basis functions that are closely associated with changes in variance, skewness, and kurtosis. A consistent analysis of changes in daily temperature distributions should assist in the identification of physical processes underlying the observed trends.

2. Daily Temperature Data

Daily temperature data are from the Global Historical Climatology Network-Daily (GHCND) database, which contains weather station-based measurements of daily maximum (Tx) and minimum (Tn) temperature over land [*Menne et al.*, 2012]. Temperature measurements in the database go through multiple steps of quality assurance discussed in detail in *Durre et al.* [2010] and are assigned flags if they do not pass one or more of the quality assurance tests.

The spatial coverage of the GHCND database is highly variable, with the highest concentration of stations in the United States and Japan and very limited coverage in the Arctic (Figure 1). The focus of the study is the Northern Hemisphere midlatitudes (30–72°N), a region chosen to balance concerns about data availability with a desire to engage in cross-region comparisons of temperature trends. Based on this choice of spatial domain, we define summer as July and August (JA) rather than the more typical June, July, and August, because a substantial component of the temperature variability in June is due to the climatological seasonal cycle as compared to random weather variations (Figure 2). During July and August, the contribution to the variance of daily temperature from weather variations is over 23 times the contribution from the climatological seasonal cycle at the majority (> 50%) of stations, compared to only seven times when June is also included.

Data that fail any quality assurance check are assigned as missing values, as are data with a measurement flag indicating that the temperature values do not appear to align with the reported hour of observation. Across all stations with Tx (Tn) data in the Northern Hemisphere midlatitudes, we remove 11.6% (11.5%) of summer temperature values based on their flag. The fraction of missing data is not, however, distributed equally across stations. The majority of stations have no more than 2.8% (2.1%) of the observations flagged for removal, whereas the worst 10% of stations have at least 45.9% (44.4%) of the observations flagged for Tx (Tn). In general, stations with large amounts of flagged data are not concentrated in a single location with the exception

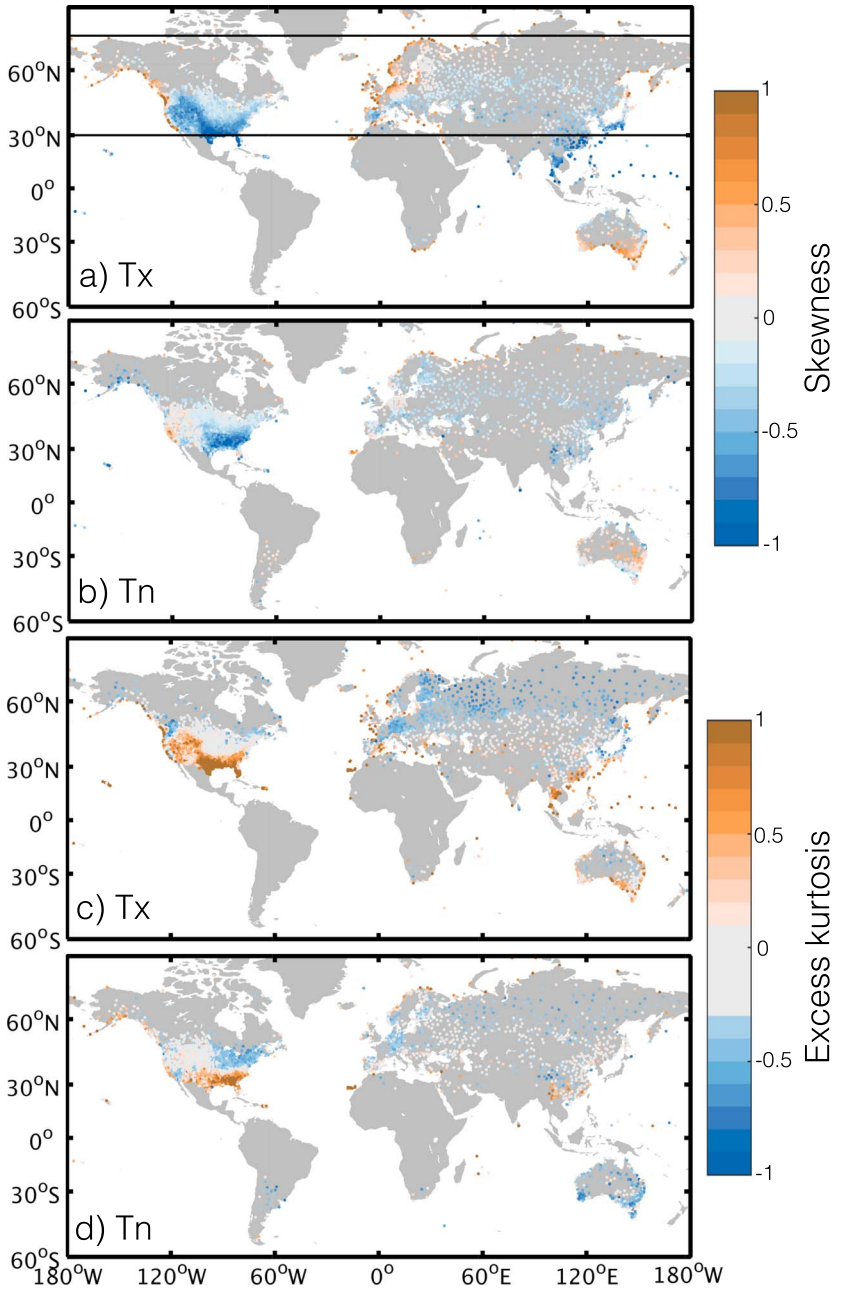

Figure 1. The distribution of Global Historical Climatology Network-Daily stations that have sufficient data for inclusion in the analysis. Daily summer temperature tends to be nonnormal, as shown by the nonzero (a, b) skewness and (c, d) excess kurtosis for (Figures 1a and 1c) Tx and (Figures 1b and 1d) Tn. Skewness and excess kurtosis values are calculated as the average across all summers between 1980 and 2015. Gray regions on both color bars indicate that the values of skewness and excess kurtosis are not inconsistent with a normal distribution at the 5% level based on 10,000 bootstrap samples. Due to the highly variable station density, points in densely sampled regions such as the United States are plotted on top of each other. In Figure 1a, the horizontal lines indicate the region of study (30–72°N).

of western Russia, where many stations have more than half of their data flagged. In order to be included in the analysis, stations are required to have at least 80% coverage during summer (JA) for at least 80% of the years considered in the analysis (1980–2015), yielding a minimum number of daily observations of 1428.

Some weather stations in the GHCND database suffer from changes in measurement practices that can influence the temperature observations termed inhomogeneities. For example, the switch from liquid in glass to

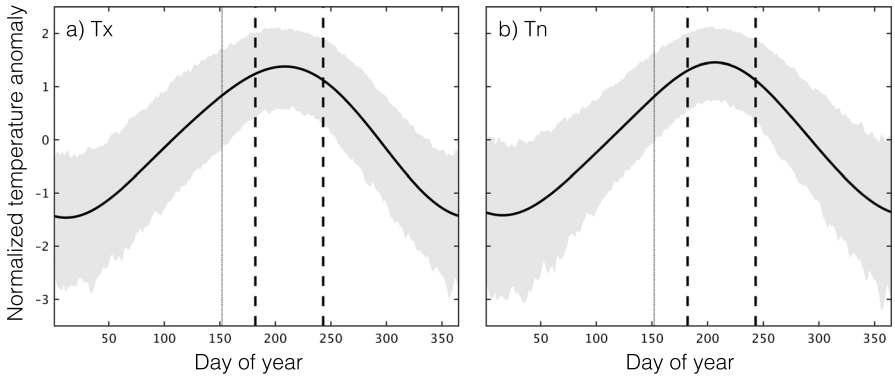

Figure 2. Weather variability during July and August is generally larger than changes in the climatological seasonal cycle. The climatological seasonal cycle (thick black line) and the interannual spread of temperature for each day (gray shading) averaged across all stations in the Northern Hemisphere midlatitudes (30–72°N). Before averaging across stations, both metrics are normalized by the standard deviation of the climatological seasonal cycle for each station. Spread is defined as the envelope that includes 95% of the interannual variability. The vertical thick dashed lines bracket July and August, whereas the thin dash-dotted line indicates 1 June.

minimum-maximum temperature system thermometers in the United States that occurred beginning in the mid-1980s induces a cold bias at the affected stations because of differences in calibration and positioning [*Doesken*, 2005]. More subtly, there has been a shift in the United States from recording the maximum and minimum temperature values from the previous 24 h in the afternoon to recording them in the morning [*Menne et al.*, 2009]. Reading and recording the maximum temperature over the past 24 h in the afternoon will, on average, cause a warm bias in the measurement compared to calculating it over a midnight-to-midnight period, so a switch from afternoon to morning observations can induce a spurious negative trend in Tx. Conversely, there is minimal bias in reading and recording the minimum temperature over the past 24 h in the afternoon, but doing so in the morning will, on average, cause a cold bias (Figure S1 in the supporting information). Thus, a switch from afternoon to morning observation times will also induce a spurious negative trend in Tn. Removing these nonclimatic trends from temperature records is termed homogenization, and various algorithms have been introduced to homogenize monthly average data [*Venema et al.*, 2012]. Although there exist daily homogenized data for individual countries including China [*Xu et al.*, 2013], Australia [*Trewin*, 2013], and Canada [*Vincent et al.*, 2012], different methods have been applied across data sets, and we are not aware of a hemispheric- or global-scale homogenized daily temperature data set.

To assess the impact of using the nonhomogenized GHCND data in our analysis, we perform the full analysis on a separate data set, the Integrated Surface Database (ISD) [*Smith et al.*, 2011], a global network of stations that contains hourly measurements. We identify a subset of the ISD stations that are nearby GHCND stations using a cutoff of 100 km in the horizontal and 100 m in the vertical. ISD stations are only used if they pass the same data completion criterion as was used for GHCND. There are 1238 (832) pairs of stations with Tx (Tn) data in the Northern Hemisphere midlatitudes (Figures S2a and S3a). We directly address the "time of observation" inhomogeneity in GHCND through calculating Tx and Tn in the ISD data via consistent time of day sampling. The comparison with ISD also addresses other inhomogeneities related to changes in measurement practices under the assumption that two nearby stations would not change their measurement practices in tandem. The two data sets are not, however, entirely independent, with approximately 25% of the GHCND measurements in our comparisons originally coming from the Automatic Surface Observing System or Global Summary of the Day that are also incorporated into ISD. Nevertheless, this low fraction indicates that there is a large degree of independence between the two sources of temperature data.

The overall results using ISD data are very similar to those using GHCND data (Figures S2 and S3). Although any discrepancies between the results from the two data sets merit further examination, even the largest differences are much smaller than the trends presented in section 5. Thus, we proceed with the nonhomogenized data set assuming that the inhomogeneities do not obscure our findings.

3. Quantile Regression Trends

To assess changes in the shape of summer temperature distributions, we employ quantile regression, a nonparametric technique that allows for estimation of trends in any percentile of a distribution [*Koenker and Bassett*, 1978; *Cade and Noon*, 2003]. We make use of the MATLAB implementations of the quantile regression algorithms provided by Roger Koenker (www.econ.uiuc.edu/~roger/research/rq/rq.html).

Quantile regression results can be biased when applied to noncontinuous data [*Machado and Silva*, 2005], as is true of the temperature values in the GHCND database that are provided at a precision of 0.1°C. We correct for this bias through adding random noise realizations ("jittering") with an amplitude appropriate for the level of rounding [*Koenker and Bassett*, 1978; *Machado and Silva*, 2005]. Since the original precision and units of the temperature data are not necessarily recorded in the metadata, we employ the "precision decoding" algorithm discussed in detail in *Rhines et al.* [2015]. The algorithm uses fingerprints of the rounding and unit conversion process to identify the most likely state of the observing system. Based on this methodology, it is possible to determine the range of temperature values that are consistent with the recorded value. For example, if the original precision of the data was 1°C, and the recorded value is 17°C, the actual temperature could have been any value within the range of [16.5°C, 17.5°C) under the assumption of conventional rounding. These uncertainties are incorporated into our subsequent estimation of significance of trends, as discussed below.

Linear trends in quantiles from the 5th to the 95th percentiles in steps of 5% are calculated for each station and percentile independently. Linearity is assumed on the basis of the relatively short time period of the analysis, and the influence of this assumption is tested in section 5.3. Trends are calculated separately for daily minima and maxima; we do not calculate trends in average daily temperatures due to challenges in estimating an average temperature from the maximum and minimum [e.g., *Wang*, 2014]. The 62 days of summer during July and August for a given year are assumed to be exchangeable, consistent with the minimal curvature of the seasonal cycle during peak summer as compared to day-to-day weather variability (Figure 2).

We account for two sources of uncertainty in our estimation of trends. The first results from our use of the rounded data provided in the GHCND database, and the second is due to interannual variability in the temperature time series. To calculate the uncertainty related to rounding for a given station, percentile, and type of temperature measurement (Tx or Tn), we jitter the time series of temperature measurements 100 times and calculate the quantile trend for each jittered measurement. To calculate the uncertainty in linear trends related to the magnitude of interannual variability, we use a block bootstrap. After a given time series is jittered once to produce continuous values, the best fit trend is estimated and removed. We resample the residuals with replacement using a block size of one summer season, add them back to the originally estimated best fit trend line, and re-estimate the trend. The block size is chosen under the assumption that temperatures can be correlated within a summer but that summer temperatures are independent across years. The bootstrap process is repeated 1000 times. After assessing each source of uncertainty independently, we find that the uncertainty related to interannual variability is typically 2 orders of magnitude greater than that from rounding. Nevertheless, in order to account for both sources of uncertainty, we modify the bootstrap process such that the temperature time series are rejittered at each bootstrap iteration. Trends are deemed significant if 95% of the distribution of bootstrapped trends are of a consistent sign. Significance in this context is not a statement of anthropogenic attribution but rather that there is a nonzero linear trend that can be detected amidst interannual variability.

To demonstrate the quantile regression method, we first present examples for Tx from weather station USC00027281 in Roosevelt, Arizona, and USW00013782 in Charleston City, South Carolina (Figure 3). The underlying distribution of Tx in Roosevelt is negatively skewed, which is representative of most inland weather stations during the summer months in the Northern Hemisphere midlatitudes (Figure 1). Trends in all percentiles are positive but not uniform, implying a change in the shape of the distribution. For example, trends in the 5th percentile are greater than trends in the 50th, whereas trends in the 95th are greater than those in the 50th and 5th. Although there is not a direct mapping from these three percentile trends to changes in moments, these trends would typically be associated with an increase in the variance and skewness of the distribution. The underlying distribution of Tx in Charleston City is instead positively skewed, similar to other stations that have a strong maritime influence. The trends are of mixed sign: while the 5th percentile is warming, both the 50th and 95th percentiles are cooling, suggesting a decrease in variance.

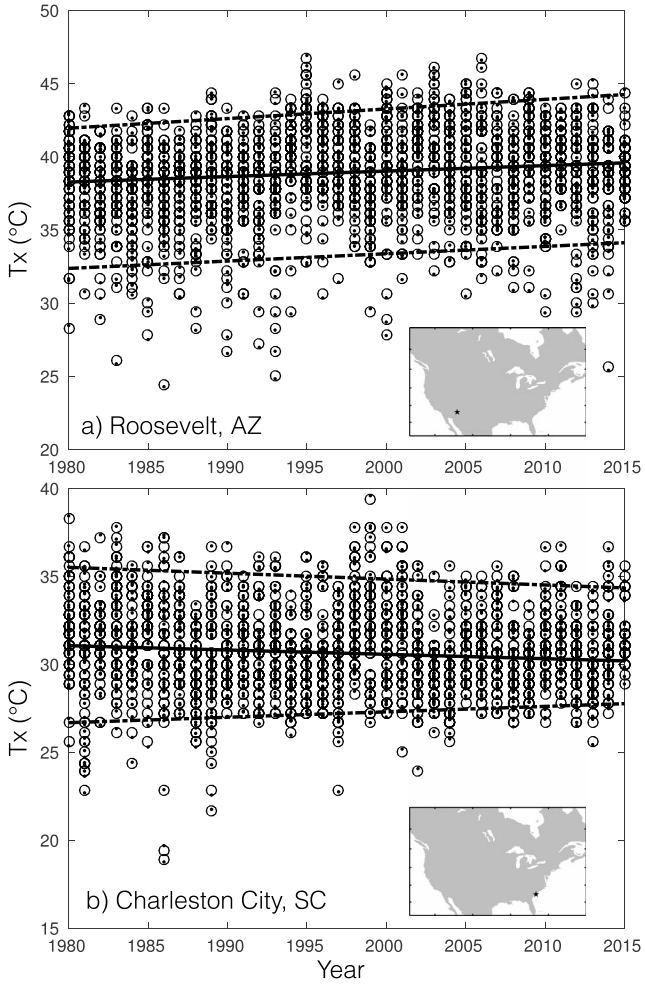

Figure 3. Quantile regression trend examples for daily maximum temperature at station (a) USC00027281 in Roosevelt, Arizona, and (b) USW00013782 in Charleston City, South Carolina. The black circles show the original, unjittered temperature data as a function of year, and the black dots show an example output from one iteration of jittering (see section 3). The thick black line is the trend in the 50th percentile, and the dash-dotted black lines are the trends in the 5th and 95th percentiles. The inset maps indicate the location of each station.

Extending the analysis across the Northern Hemisphere midlatitudes demonstrates that the changes in summer temperature distributions are not qualitatively following a simple shift behavior, which would result in the same pattern and magnitude of trends across all percentiles (Figures 4–7). We briefly highlight some of the dominant features in the maps.

The 95th percentile of Tx in the United States shows cooling along the West Coast, warming throughout the interior West, cooling in the Midwest and Southeast, and a slight warming signal in the Northeast. The 95th percentile of Tn in the United States shows a more homogeneous warming signal, although there is significant cooling in parts of the Midwest. Similar patterns to the 95th percentile in Tx and Tn are evident in the 50th percentile. In contrast, changes in the 5th percentile have a somewhat different spatial structure. The West and East Coasts tend to show strong warming trends in Tx and Tn, whereas the Southeast and Midwest have more mixed signals. Western coastal stations show a cooling in Tx but a warming in Tn.

Most stations in Europe show large and significant warming across the 5th, 50th, and 95th percentiles in Tx. Trends in European Tn for the 5th, 50th, and 95th percentiles tend to be smaller than those for Tx but still significantly positive. In contrast to Europe, many Asian stations exhibit insignificant or slightly negative trends in both Tx and Tn with the exception of Tx in central-eastern Eurasia and Japan and Tn in northeastern Russia. The large positive trends in western Russia are identified as significant within the bootstrapping framework,

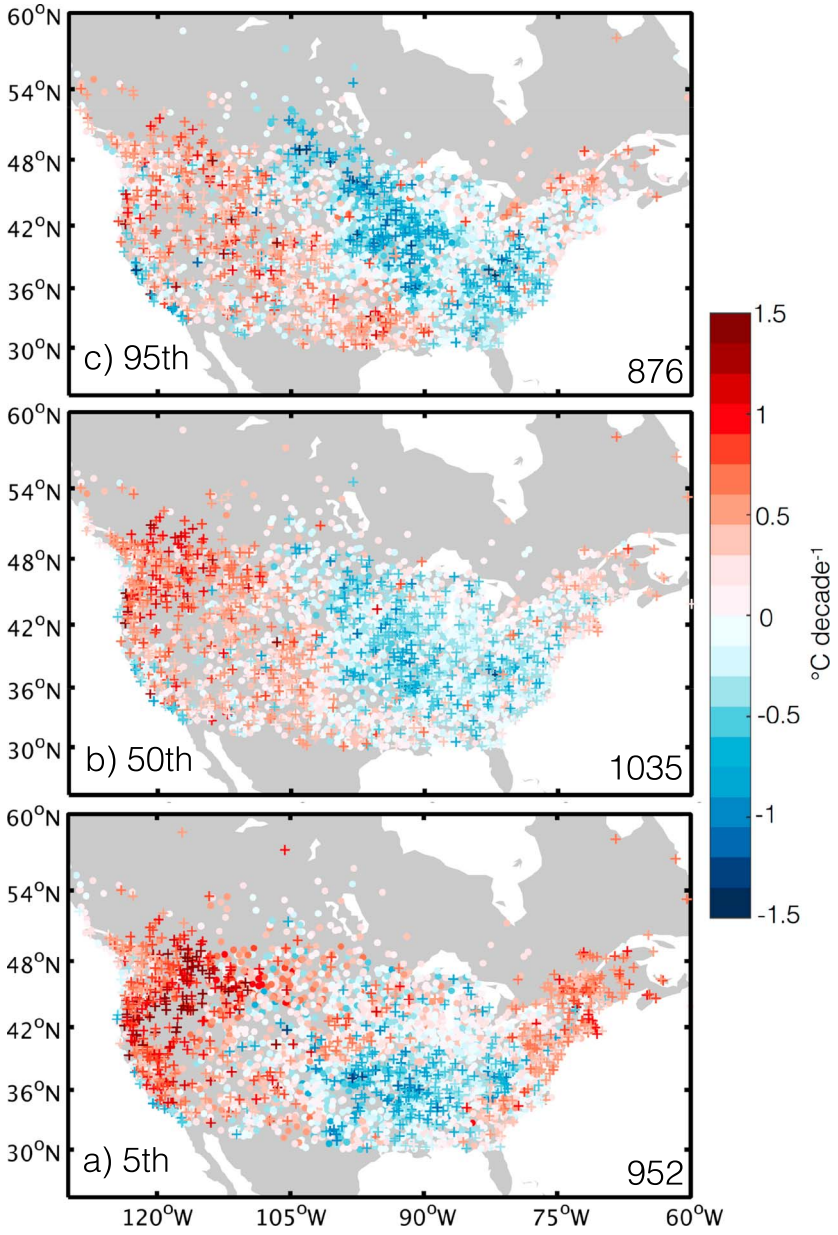

Figure 4. Quantile regression trends for the 95th, 50th, and 5th percentiles of Tx in North America. The plots are bounded at 60°N due to the extreme sparsity of station data at higher latitudes in this longitude band. Crosses show stations where trends are found to be significant at the 95% level based on a bootstrap analysis, whereas dots show stations with insignificant trends. There are 3239 total stations shown in each panel, and the number of stations that are found to be significant is in the lower right-hand corner.

suggesting that they do not result from the occurrence of the 2010 heat wave alone. Indeed, the exclusion of 2010 from the analysis yields smaller but still positive trends in the region.

4. Dominant Modes of Distributional Changes

Rather than focusing on specific percentiles, we would like to be able to quantify changes in the distribution of temperature as a whole. Two methods for doing so include extending our analysis of trends in individual percentiles to span the distribution and calculating the change in statistical moments over time. The two approaches, however, can be combined by identifying the trends in percentile space that are implied by changes in statistical moments. In this section, we first discuss the percentile trends associated with changes

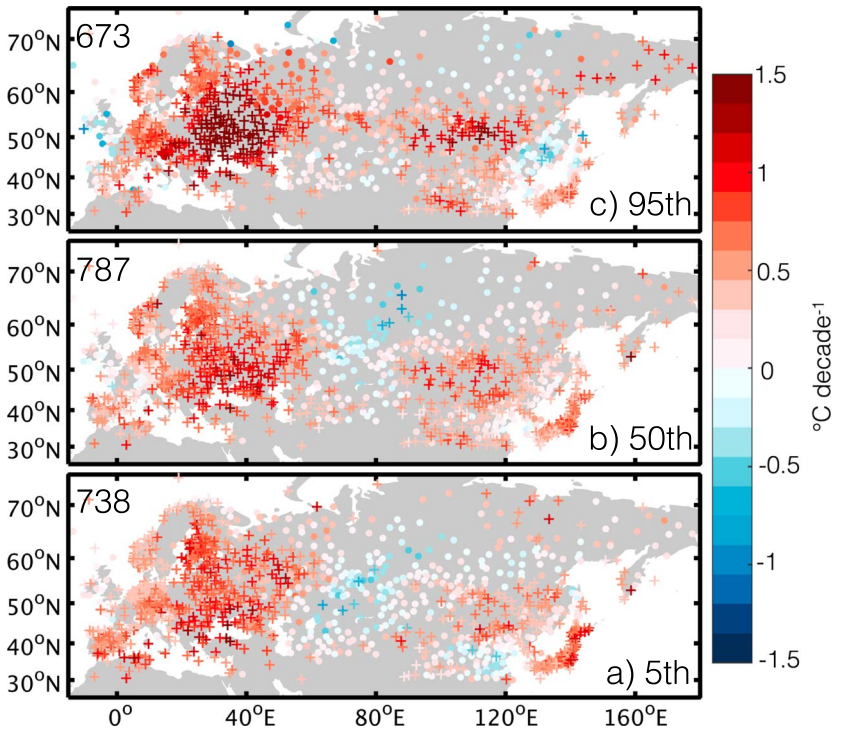

Figure 5. As in Figure 4 but for Eurasia. There are 1243 total stations shown in each panel, and the number of stations that are found to be significant is in the upper left-hand corner.

in the first four statistical moments and subsequently compute a similar but somewhat more optimal set of basis functions that summarizes the observed changes in shape of the temperature distributions.

4.1. Percentile Trends Due To Changing Moments

The relationship between changes in the first four statistical moments of a distribution and trends for the percentiles of the distribution is demonstrated using the Pearson system [*Pearson*, 1895] (see also Figure S4). The function associated with an increase in the mean of a normal distribution of 1 maps to a horizontal line in percentile space (Figure 8a). The function associated with an increase in variance is calculated by taking the difference between the cumulative distribution function for a normal distribution with a variance of 1.5 and that for a normal distribution with a variance of 1. The resulting function is monotonically increasing in percentile space and includes positive trends for percentiles above the 50th percentile and negative trends for those below the 50th percentile. The mapping from changes in skewness and kurtosis to trends in percentile space has some dependence on the underlying distribution that is assumed, but the qualitative shape of the trends in percentile space is not sensitive to the choice. Here the functions are calculated by increasing the skewness from −0.5 to 0.5 in a distribution with zero mean, unit variance, and zero excess kurtosis and by increasing the excess kurtosis from −1 to 1 in a distribution with zero mean, unit variance, and zero skewness. The skewness function is associated with positive trends in both tails and negative trends in the middle of the distribution, whereas the kurtosis function has no impact on the 50th percentile and opposite impacts on either tail. Note that the choice of different increments of change for each moment is because the changes in percentile space associated with a unit shift in each moment can be quite disparate, with a unit change in variance having a much larger impact on the low and high percentiles than a unit change in skewness and kurtosis. This choice is primarily for display purposes and does not quantitatively influence our results.

We do not, however, proceed with using this set of functions to summarize the trends across percentiles because they are not orthogonal. In particular, the variance and kurtosis functions are correlated at $r = -0.81$, preventing a clean separation of the contribution of each function to the percentile trends. Instead, we first use principal component analysis to calculate a set of orthogonal basis functions and then demonstrate that they can be well approximated with the Legendre polynomials.

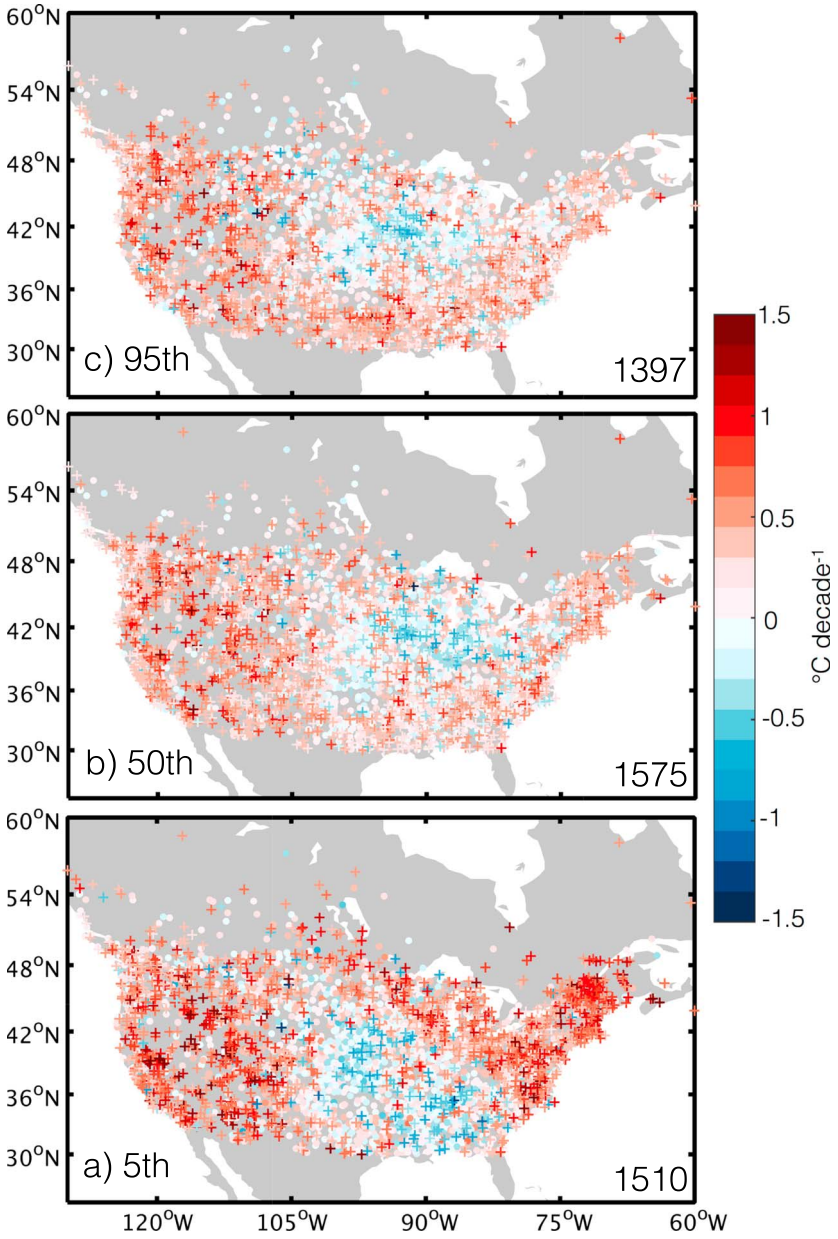

Figure 6. As in Figure 4 but for Tn. There are 3215 total stations shown in each panel.

4.2. Principal Component Analysis of Percentile Trends

To create an orthogonal set of basis functions that best explain the variance of the quantile trends across stations, we employ principal component analysis. The first basis function, B0, is defined as an equal shift across all percentiles and so is identical to the basis function that results from an increase in the mean of a normal distribution. The next three basis functions, B1–B3, have similar structures to the functions associated with changes in variance, skewness, and kurtosis (Figures 8b and 8c). The four basis functions collectively explain 98.8% of the variance of quantile trends across stations for Tx and 98.7% for Tn. Although stations are distributed unequally across the Northern Hemisphere, we find that the basis functions calculated using only the Western Hemisphere, only the western half of Eurasia, and only the eastern half of Eurasia are all very similar and explain comparable amounts of variance and so do not further account for the spatial distribution of stations. The basis functions produced through principal component analysis have the advantage of being orthogonal in addition to resembling the functions associated with changes in the first four statistical moments; however, they have a dependence on the underlying data that can be seen when comparing the

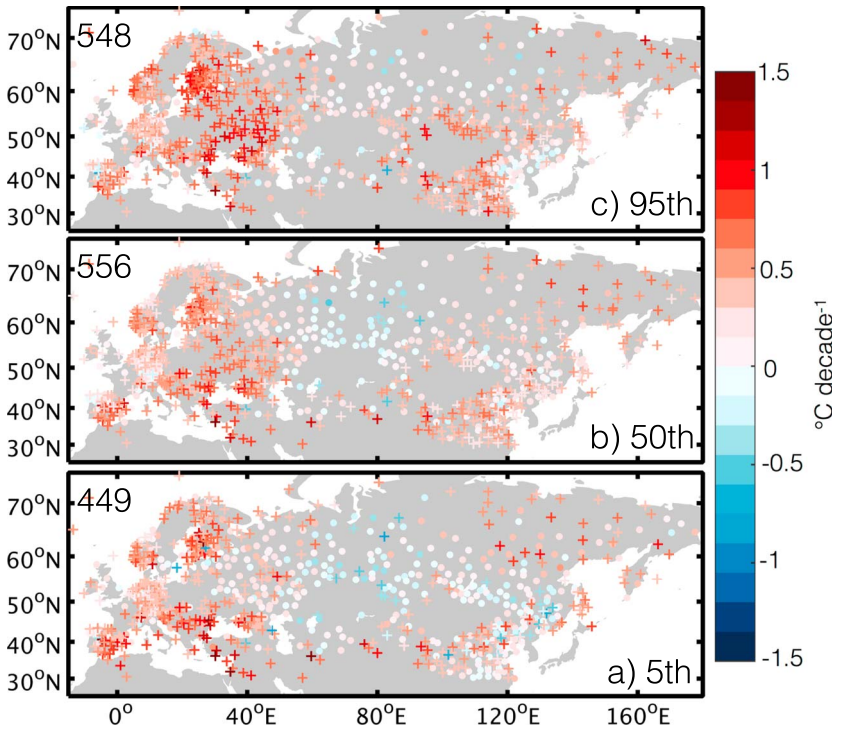

Figure 7. As in Figure 5 but for Tn. There are 850 total stations shown in each panel.

basis function for Tx and Tn (Figures 8b and 8c). To improve interpretability of our results, we now seek a more universal set of basis functions.

4.3. Approximation With Legendre Polynomials

The form of the basis functions calculated using principal component analysis emerges solely from the covariance structure of the observed trends in Tx and Tn across percentiles, but they are closely related in shape to Legendre polynomials (Figure 8d). The similarity is not incidental: the results from principal component analysis can be closely approximated by Legendre polynomials under a specific set of conditions [*Gibson et al.*, 1992]. In this case, the relevant condition is sufficiently high autocorrelation of trends across percentiles, which holds true in the data (Figure 9). The correlation between trends in different percentiles of Tx tends to be higher than that for Tn, consistent with the greater similarity between the Tx basis functions and the Legendre polynomials than for Tn. Nevertheless, for both Tx and Tn, a nearly equivalent amount of variance in the trends across percentiles and stations can be explained by the Legendre polynomials (98.8% for Tx and 98.6% for Tn) as for the optimal set calculated from principal component analysis (98.8% for Tx and 98.7% for Tn). The small amount of the remaining variance (1.2% for Tx and 1.4% for Tn) is associated with changes in higher moments of the distribution that are not considered in this analysis. We choose to employ the Legendre polynomials as our set of basis functions throughout the remainder of the analysis because their form is similar to the results from the principal component analysis yet is independent of the underlying data. This independence allows for a more clear comparison of the results using Tx and Tn, as well as between this and future analyses of other data sets with similar properties.

Like the basis functions from the principal component analysis, Legendre polynomials are orthogonal and have strong similarities to the percentile trends that result from shifts in the first four statistical moments. Due to the similarity, we will refer to the first four Legendre polynomials as the shift, variance, skewness, and kurtosis basis functions, but it is important to recall that they do not exactly represent changes in the moments of the same name. A positive projection onto each basis function is associated with an increase in the associated moment. The Legendre polynomials are rescaled to have a standard deviation identical to the aforementioned prescribed shifts in the associated statistical moments (Figures 8a and S4). Positive projections onto the variance, skewness, and kurtosis basis functions are all associated with an amplification of hot extremes.

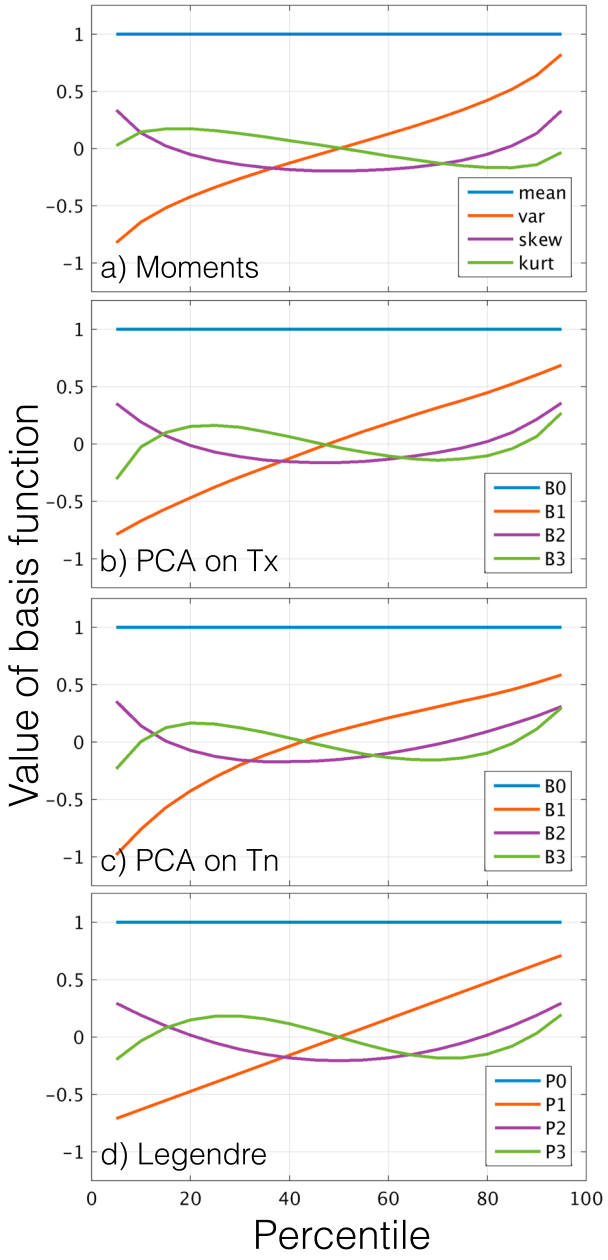

Figure 8. Basis functions for summarizing trends across percentiles. (a) The trends across percentiles associated with a mean shift of 1 (blue), half a unit increase in variance (red), a unit increase in skewness (purple), and a two-unit increase in kurtosis (green). Basis functions for trends across percentiles calculated via principal component analysis for (b) Tx and (c) Tn. (d) The first four Legendre polynomials that are used as the orthogonal bases for the analysis. The basis functions in Figures 8b–8d are orthogonal. The variance and kurtosis functions in Figure 8a are correlated at $r = -0.81$.

5. Observed Distributional Changes

Using the Legendre polynomials, it is now possible to cleanly attribute changes in summer temperature distributions to four basis functions that directly map to trends in each percentile of the distribution. The shift basis function alone explains 88% of the variance in the quantile trends for Tx and 87% for Tn. Thus, the vast majority of the observed changes in summer temperature distributions between 1980 and 2015 can be understood as a simple shift. This shift behavior, however, cannot explain any amplification or damping of extremes with respect to the mean. We therefore turn our focus to the 12–13% of variance that is unexplained by a simple shift. For Tx, 69% of the remaining variance is explained by the variance basis function, 15% is explained by the skewness basis function, and 5% is explained by the kurtosis basis function. For Tn, 62% of the remaining

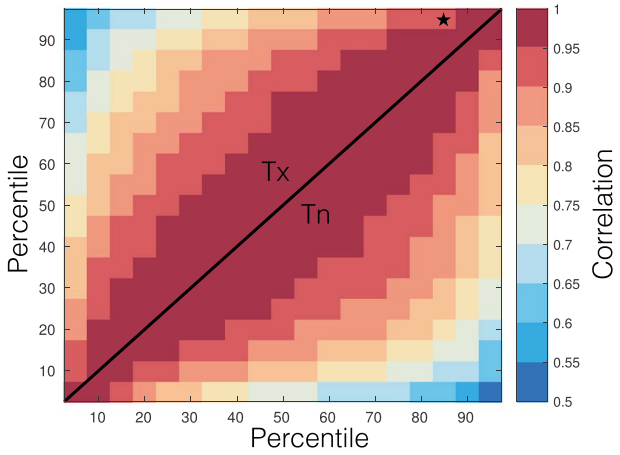

Figure 9. Correlation across stations between trends in each percentile of (top left) Tx and (bottom right) Tn. For example, the correlation between the trends in the 85th percentile and 95th percentile of Tx across all stations is 0.93 and is indicated by a black star in the figure. All values along the diagonal have a correlation of unity, because they measure the correlation of the trend of each percentile with itself.

variance is explained by the variance basis function, 20% is explained by the skewness basis function, and 4% is explained by the kurtosis basis function. The partitioning of variance is similar if any of the other basis functions in Figure 8 are used, although—consistent with the covariance between the percentile trends from increases in variance and kurtosis—the use of the bases in Figure 8a leads to a decrease in the variance explained by the variance basis function, with some of the explained variance shifted to the kurtosis basis function.

5.1. Consistency With Stationarity at the Station Level

In addition to identifying the sign and magnitude of the projections of the quantile trends onto the four basis functions, it is of interest to determine whether the trends are consistent with stationarity, defined here as a distribution whose moments do not change with time. To do so, we compare the projection of the quantile trends onto each basis function to a null hypothesis assuming interannual stationarity of summer temperature. To calculate the null distribution for changes in each basis, the temperature records at individual stations are reshuffled 1000 times using a block size of one summer season. Quantile regression is performed on each randomized time series for each station, and the resulting quantile trends are projected onto the Legendre basis functions. The distribution of the coefficients of the projection illustrates the types of shape changes that could occur even if the time series was stationary. The projections of the observed trends onto the basis functions are deemed significant if they are greater (less) than 97.5% of the values produced from the bootstrapped time series for positive (negative) values of the coefficient. This is equivalent to using a two-sided test at a 95% confidence level, because we do not have a priori knowledge about the sign of the projections onto each basis function. Due to the differences in the underlying temperature distributions at each station, the confidence intervals are station dependent. The results are not overly sensitive to the choice of a 95% confidence level compared to a 90% level. In particular, altering the level has no impact on the assessment of the significance of the projections onto the four basis functions for each region in Tables 1 and 2.

5.2. Spatial Correlation Across Stations

Daily temperature has a correlation length scale that is generally larger than the distance between stations such that each station should not be viewed as independent. The correlation across stations has two implications for our estimates of significance.

First, insomuch as stations are correlated there is a smaller number of effective degrees of freedom than the number of stations, and it is more likely that stations will exhibit similar trends than if they were independent. In the limit of perfect correlation, either all or no stations would exhibit trends deemed significant or insignificant on a station-by-station basis. Nevertheless, the fraction of stations in each region found to have significant projections onto the variance, skewness, and kurtosis basis functions is generally at least 0.02-0.03 at each tail for most regions, consistent with the expectations of our null distribution. An exception

Table 1. The Fraction of Stations That Have Positive or Negative Projections on the Mean, Variance, Skewness, and Kurtosis Basis Functions for Tx[a]

	Mean Basis	Variance Basis	Skewness Basis	Kurtosis Basis
North America (3239 Stations)				
Fraction of positive projections (significant)	0.52 **(0.16)**	0.31 (0.03)	0.60 **(0.04)**	0.46 (0.02)
Fraction of negative projections (significant)	0.48 **(0.07)**	0.69 **(0.09)**	0.40 (0.01)	0.54 (0.02)
Western Eurasia (720 Stations)				
Fraction of positive projections (significant)	0.92 **(0.61)**	0.61 **(0.05)**	0.63 **(0.05)**	0.49 **(0.04)**
Fraction of negative projections (significant)	0.08 (0.00)	0.39 (0.03)	0.37 (0.01)	0.51 (0.02)
Eastern Eurasia (387 Stations)				
Fraction of positive projections (significant)	0.89 **(0.41)**	0.66 **(0.11)**	0.38 (0.03)	0.54 **(0.04)**
Fraction of negative projections (significant)	0.11 (0.01)	0.34 **(0.05)**	0.62 **(0.06)**	0.46 (0.01)
Japan (130 Stations)				
Fraction of positive projections (significant)	1.00 **(0.75)**	0.35 (0.01)	0.31 (0.02)	0.36 (0.00)
Fraction of negative projections (significant)	0.00 (0.00)	0.65 **(0.09)**	0.69 (0.02)	0.64 (0.02)

[a]The fraction of stations that are found to be significant at the 95% level using a two-sided test is shown parenthetically. Fractions that exceed the expected value of 0.025 of significant stations for each tail are in bold. Eurasia is separated into western and eastern at 80°E. Regions are chosen such that they have relatively equal spatial distributions of stations in order to prevent a more densely sampled region from dominating the assessments of significance.

is Japan, where the fractions of trends found to be significant at each tail are generally lower than 0.025, possibly because of the small number of degrees of freedom in the region.

Second, our assessment of significance does not take advantage of spatial information, which reduces the noise in individual temperature time series and allows for the identification of more significant temporal trends [*Reich*, 2012; *Fischer and Knutti*, 2014]. For the purposes of our analysis, we suggest that regions that have large-scale trends of a consistent sign and include interspersed significant and insignificant stations can be interpreted as exhibiting significant trends across the region.

5.3. Effect of Assumption of Linear Trend

To assess whether our initial assumption that the quantile regression trends are linear has implications for our estimates of shape changes, we perform our full analysis on two types of 36 year synthetic time series: one whose seasonal mean increases linearly at a rate of 1°C per decade and another whose seasonal mean increases exponentially with an *e*-folding time of 23 years. The *e*-folding time is chosen such that the least squares linear approximation to the exponential function is similar to the linear time series. For both seasonal mean time series, the distribution of daily temperatures about the seasonal mean is chosen to be normal with a standard deviation of 5°C. We produce 1000 pairs of the linear and exponential synthetic time series, each with different daily realizations of anomalies about the trend, calculate their quantile regression slopes, and project them onto each basis function. The distributions for each of the four basis functions are not found to

Table 2. As in Table 1 but for Tn[a]

	Mean Basis	Variance Basis	Skewness Basis	Kurtosis Basis
North America (3215 Stations)				
Fraction of positive projections (significant)	0.76 **(0.39)**	0.42 **(0.04)**	0.60 **(0.08)**	0.41 (0.02)
Fraction of negative projections (significant)	0.24 (0.03)	0.58 **(0.08)**	0.40 (0.02)	0.59 **(0.06)**
Western Eurasia (544 Stations)				
Fraction of positive projections (significant)	0.90 **(0.66)**	0.60 **(0.06)**	0.63 **(0.05)**	0.49 (0.02)
Fraction of negative projections (significant)	0.10 (0.00)	0.40 (0.02)	0.37 (0.01)	0.51 (0.03)
Eastern Eurasia (301 Stations)				
Fraction of positive projections (significant)	0.88 **(0.54)**	0.68 **(0.13)**	0.43 (0.02)	0.58 **(0.05)**
Fraction of negative projections (significant)	0.12 (0.01)	0.32 (0.01)	0.57 **(0.06)**	0.42 **(0.04)**

[a]There are no stations in Japan that pass our data completion requirements for Tn.

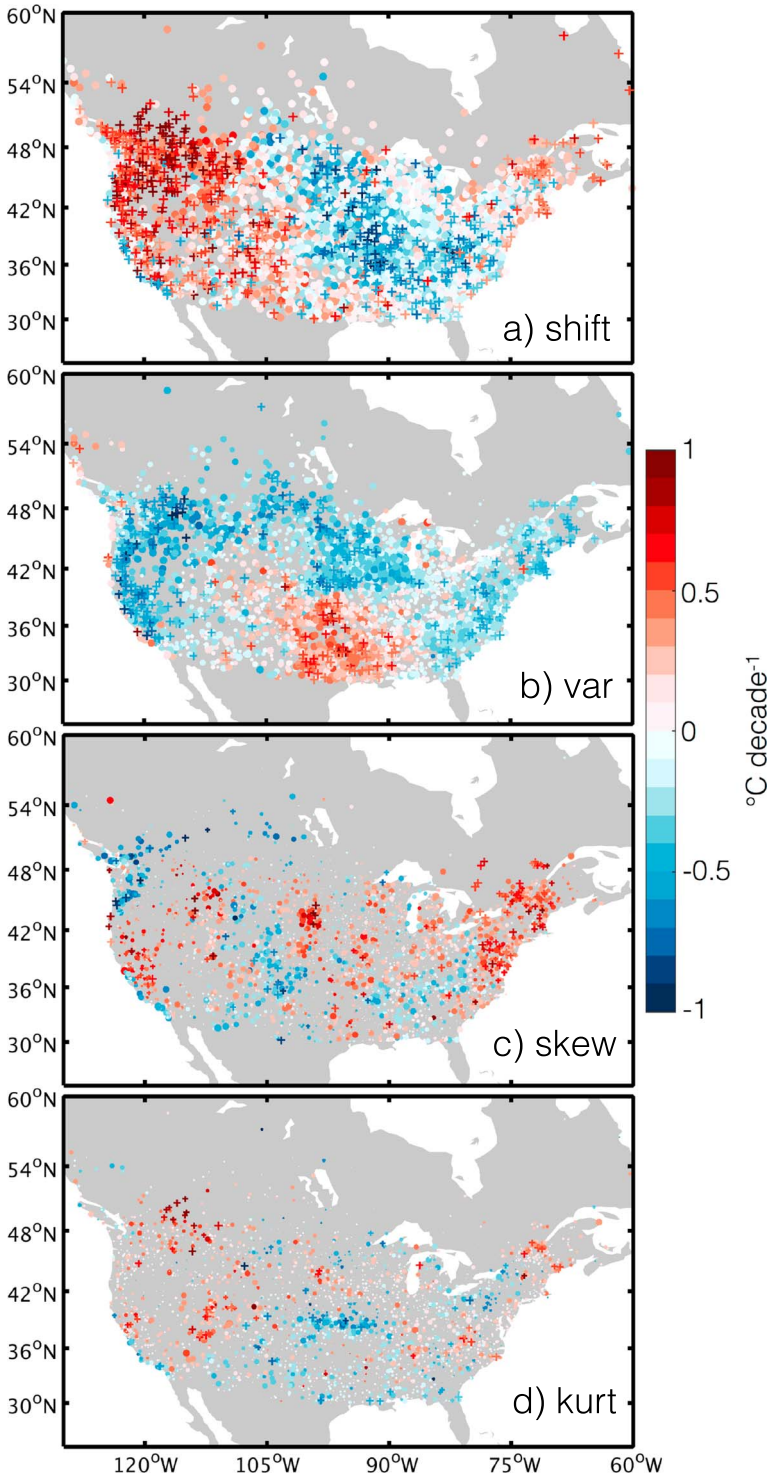

Figure 10. The projection of trends in Tx onto the first four unitless Legendre basis functions (Figure 8d) in North America. The first basis function (a) is a shift across all percentiles, whereas the latter three are closely associated with changes in (b) variance, (c) skewness, and (d) kurtosis. The same stations are plotted in each panel, but the size of the markers in Figures 10b–10d is scaled by the percent of variance in the quantile trends explained by each basis function. As such, stations that project weakly onto the higher-order basis functions appear small or absent. The dots and crosses in Figure 10a are all of the same size and are scaled such that they would represent 100% of the variance explained in Figures 10b–10d. Crosses indicate significance at the 95% level, where significance is assessed against the null hypothesis of stationarity (see main text). The units for each panel are identical because the values reflect the projection of the quantile trends onto unitless basis functions.

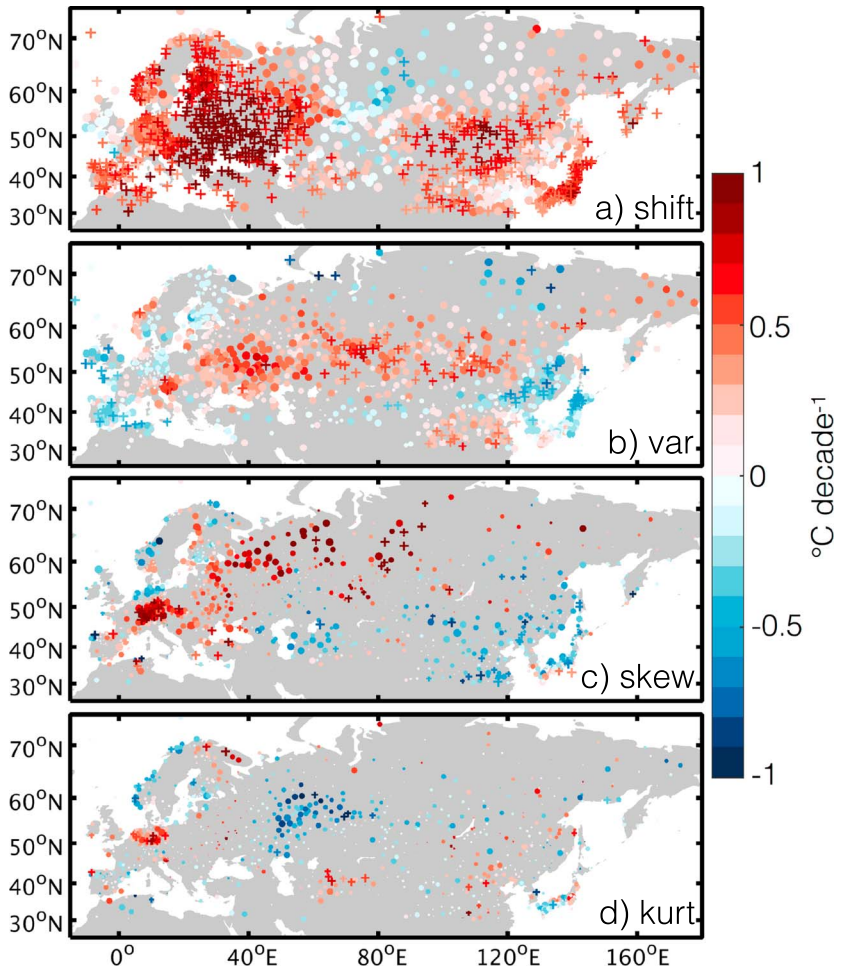

Figure 11. As in Figure 10 but for Eurasia.

be inconsistent with each other based on a two-sample Kolmogorov-Smirnov test ($\alpha = 0.01$), indicating that the results are not affected by reasonable nonlinearities.

5.4. Spatial Structure

The majority but not entirety of stations exhibit a positive shift across all percentiles for Tx and Tn (Figures 10a, 11a, 12a, and 13a and Tables 1 and 2), consistent with a global warming signal. North American stations show a more consistent warming in Tn than Tx, with significant negative trends in Tx over the midwestern and southeastern United States. Eurasia shows the most consistent and significant increases in both Tx and Tn. Almost none of the stations that exhibit cooling in Eurasia are found to be significant.

The sign of the projections of temperature trends onto the variance basis function is less consistent than that for the mean shift, but the fraction of stations that show significant projections onto the variance basis function remains greater than the 5% that would be expected by chance. The majority of North American stations show a negative projection onto the variance basis function in both Tx and Tn, with the exception of eastern Texas. A decrease in variance alone will decrease the probability of extreme events and so can help compensate for increases in the probability of hot temperatures associated with positive shifts in the distribution. The effect can be clearly seen through comparing the distribution of Tx in the western United States during the first and last 10 years of the period of study (Figure 14a). Note that trends that project onto the variance basis function capture trends in both the within-season and across-season temperature variability. For example, a time series that has the same day-to-day variability within each season but an increase in the year-to-year variability of average summer temperatures would project positively onto the variance basis function, as would one with an increase in the within-season variability with no change in the interannual variability.

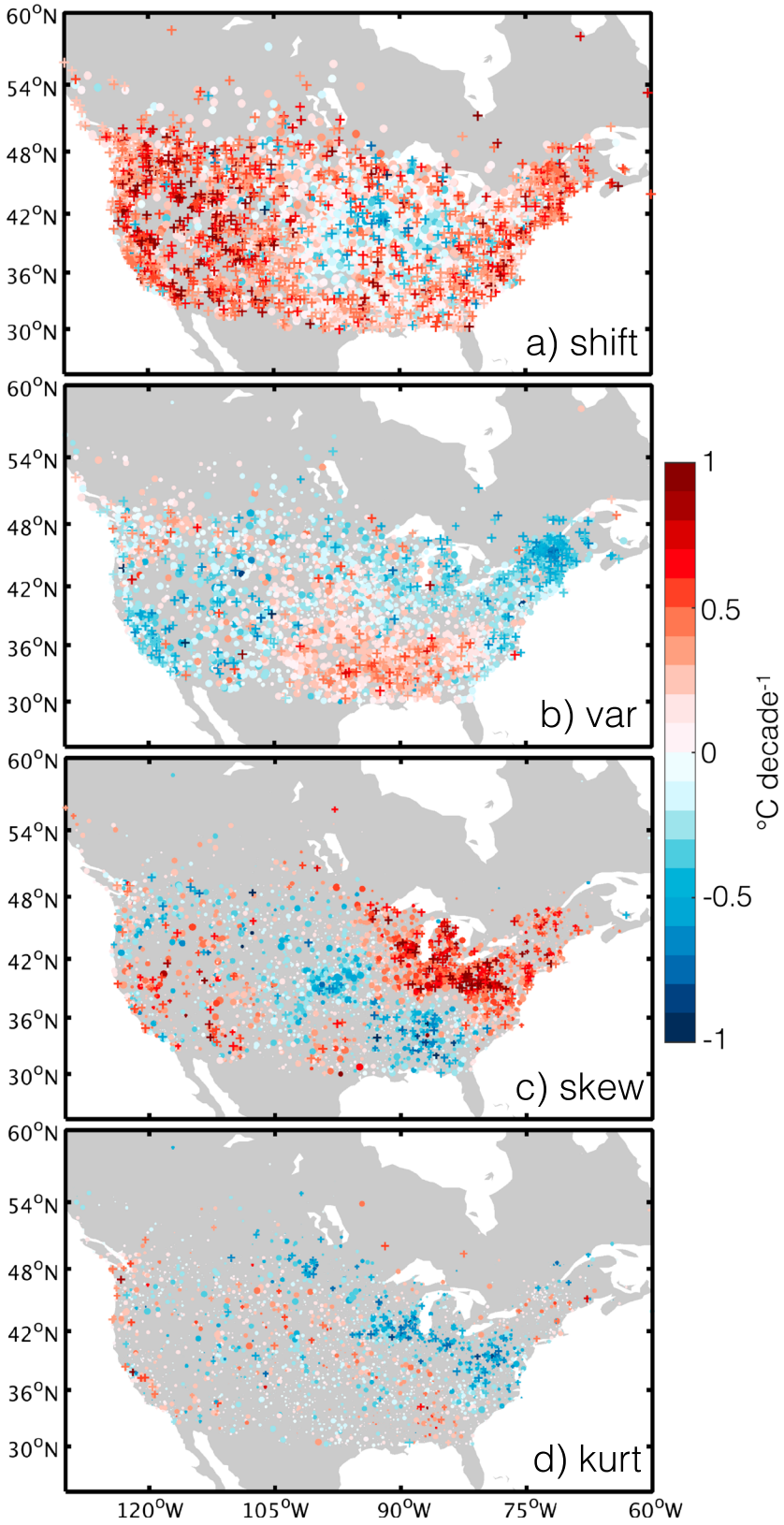

Figure 12. As in Figure 10 but for Tn.

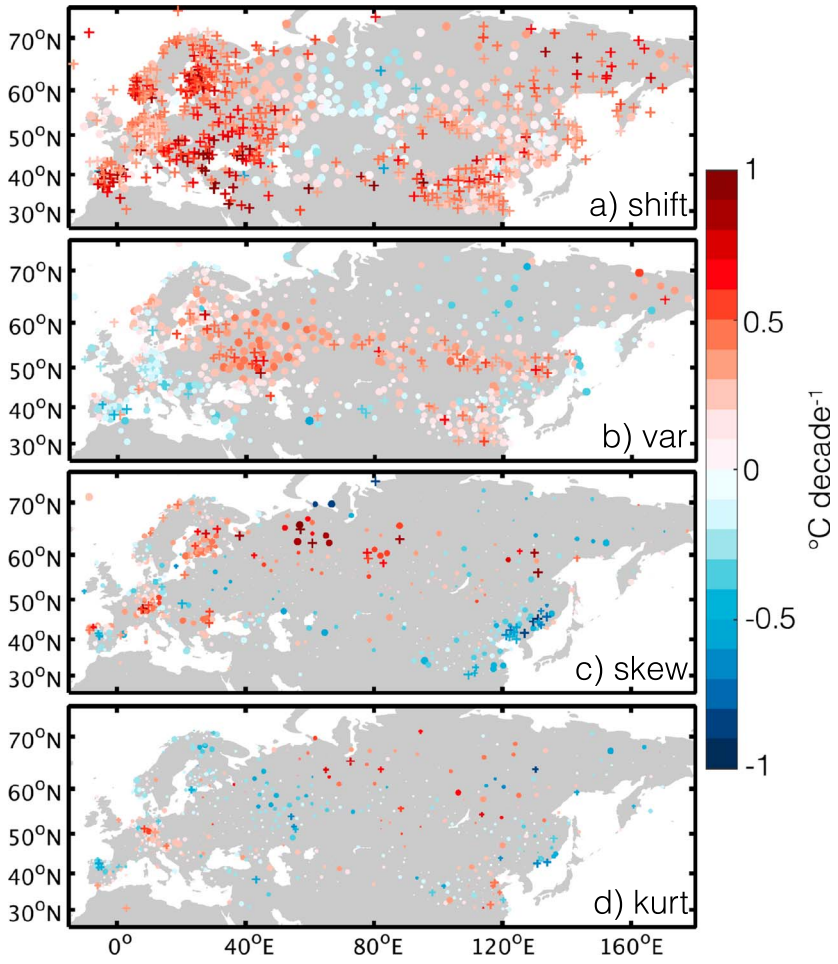

Figure 13. As in Figure 11 but for Tn.

In contrast to North America, the majority of Eurasian stations show increases in variability for Tx and Tn. The structure is regionally coherent: most central Eurasian stations show positive projections onto the variance basis function, while western Europe generally exhibits negative projections. For Tx only, trends project negatively onto the variance basis function in Sweden, eastern China, and Japan. Few of the projections onto the variance basis function are significant at the 95% level, including the large positive values in eastern Europe because of the high variability of the underlying climate. The inflation of variance in eastern Europe is colocated with large positive shifts in the full distribution, leading to large increases in the probability of hot extremes (Figure 14b).

Changes in the basis functions associated with the higher-order moments of skewness and kurtosis tend to have a smaller spatial scale and do not, on average, explain as much variance in the trends. For Tx, the fraction of all stations with significant projections onto the skewness and kurtosis bases is 0.05, which is the rate of expected false positives given our use of 95% confidence intervals. For Tn, the fraction is higher—0.10 for skewness and 0.07 for kurtosis—slightly exceeding that which would be expected by chance.

We can now compare our results to prior regional studies, with the important caveat that the prior studies used longer time periods than our analysis, as well as a variable set of methods. Existing results suggested that changes in the distribution of daily summer temperatures in the Iberian Peninsula, Italy, and Sweden were consistent with a simple shift of the full distribution [*Della-Marta et al.*, 2007; *Simolo et al.*, 2010; *Acero et al.*, 2014]. We also find that changes in the shape of the distribution of Tx and Tn are not significant in either Italy or Sweden but identify significant decreases in variance in the Iberian Peninsula. While *Della-Marta et al.* [2007]

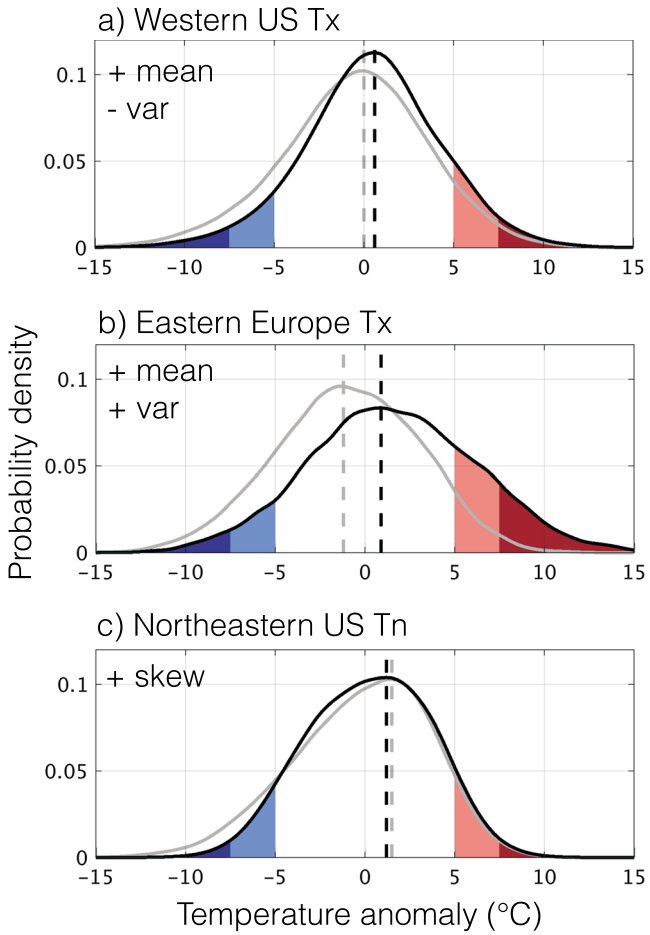

Figure 14. Changes in daily temperature distributions for (a) the western United States (35–50°N, 115–130°W, Tx, 415 stations), (b) eastern Europe (45–60°N, 30–45°E, Tx, 66 stations), and (c) the northeast United States (38–48°N, 70–90°W, Tn, 652 stations). The gray distribution is calculated using daily temperature from July and August during 1980–1989, and the black distribution is calculated using 2006–2015. Anomalies are calculated with respect to each station's average summer (JA) temperature during 1980–1989. Dark red (blue) shading indicates the probability of anomalies greater than (less than) 7.5°C above (below) normal; light red (blue) shading is for anomalies greater than (less than) 5°C. The vertical dashed lines indicate the median of each distribution. The western United States exhibits a large positive shift but a decrease in variance, leading to compensating behavior with respect to the probability of hot extremes. In contrast, Eastern Europe exhibits a large positive shift as well as an increase in variance. The northeastern United States shows a minimal shift or variance change, but a large increase in skewness decreases the probability of cold extremes. The figure is intended primarily for visualization purposes.

identified an increase in variance around Germany, we instead find little change in variance, although we do identify large and positive projections onto the skewness and kurtosis basis functions in Tx. Positive projections onto the skewness and kurtosis basis functions fatten the upper tail of the temperature distribution in a manner similar to increases in variance.

Trends project significantly and positively onto the skewness basis function around southern Germany and in the northeastern United States for Tx, as well as in the midwestern and northeastern United States for Tn (e.g., Figure 14c). Trends project negatively onto the skewness basis function for Tn in the southeastern United States and southeastern Russia. Northern German stations exhibit positive projections onto the kurtosis basis function for Tx that explain much of the variance of the overall changes in shape. Small bands of stations in the interior western United States also show large positive projections onto the kurtosis basis function, whereas a band of stations across southern Kansas and Missouri exhibit negative values. There is a tendency toward negative projections onto the kurtosis basis function across the northeastern and midwestern United States.

6. Discussion and Conclusions

Across the Northern Hemisphere midlatitudes, there are spatially coherent and significant changes in the distributions of daily maximum and minimum summer temperatures. While a large majority of the changes can be explained by shifts in the distributions, there are also significant changes in the shape of the distributions related to the probability of extreme events. We have demonstrated that the changes can be summarized through projections onto the first four Legendre polynomials. Furthermore, the Legendre polynomials are good approximations of the dominant modes of variability across percentiles and stations in the data, which themselves are similar to those that would result from enforcing an increase in the mean, variance, skewness, and kurtosis of a distribution.

The projections of trends in Tx and Tn onto the basis functions were compared to those that would be expected from interannual variability of a stationary distribution. Trends associated with each basis function are increasingly consistent with a stationary distribution as higher-order moments are considered. Changes associated with a mean shift in the distribution are the most significant overall, and nearly all of the significant shifts in the distributions are positive. Consistent with previous work [*Donat et al.*, 2013], we find a greater fraction of stations in each region (North America, eastern Eurasia, and western Eurasia) that show a significant positive shift in daily summer temperature for Tn than Tx. There does not appear to be a clear difference between Tx and Tn regarding the fraction of stations that show significant changes in other moments.

The spatial scale of changes associated with a shift of the distributions of Tx and Tn is similar to that of summer atmospheric circulation, suggesting that some of the structure may be related to trends in circulation. The trends, however, need not be forced: regional trends in temperature over time periods as long as five decades can be heavily influenced by sampling natural atmospheric variability [*Deser et al.*, 2012], particularly in the cold season. As such, the current analysis does not illuminate whether the present trends are due to anthropogenic influence and thus whether or not they are likely to continue in a similar manner.

The present results show regions of both increasing and decreasing daily temperature variance. The spatially mixed signal of changes in variance is in line with the conclusions drawn by *Huntingford et al.* [2013] for the interannual variance of summer average temperatures. However, the spatial patterns of change in interannual variance bear little resemblance to the trends we find when considering trends in the distribution of daily data, which take into account both within-season variance and interannual variance. For example, while we find increasing variability in east Texas and decreasing variability in the western and midwestern United States (Figures 10), *Huntingford et al.* [2013] identify trends in interannual variance of opposite sign. With the caveat that *Huntingford et al.* [2013] consider changes before and after 1980, whereas we consider trends from 1980 to 2015, together these studies may indicate that trends in interannual variance and within-season variance have been of opposite sign. The different behaviors on the two timescales likely result from the different physical mechanisms important for synoptic versus interannual processes and highlight the importance of apples-to-apples comparisons between studies. It would be useful to parse the contribution from the different timescales in influencing trends in variability in future work.

Although there is not a hemispherically consistent trend in temperature variance, there are coherent trends across large regions. Stations in eastern Europe and southern Russia show positive and significant projections of Tx onto the variance basis function. Some of the same stations also exhibit large, positive shifts in the distribution of Tx, and the two effects combined suggest a large increase in the probability of hot extremes (Figure 14b). Changes in variance, however, do not always act as an amplifier of shifts in the distribution. The significant decreases in the variance of Tx in the western United States act to damp the increase in hot extremes that would result from the positive shift in the distribution as a whole (Figure 14a).

Further work is required to identify the causal mechanisms leading to the quantile trends identified here. Nevertheless, we have introduced a method that summarizes trends across the full distribution of temperature and have demonstrated how it can be used to cleanly separate shifts in distributions from other changes in shape. While the temperature trends observed since 1980 can primarily be characterized by a simple shift, we identify large regions that also show significant shape changes that will influence probabilities of extreme events.

Acknowledgments
We thank Clara Deser for her helpful comments. The authors acknowledge funding from the NCAR Advanced Study Program and NSF grant 1304309. The computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. GHCND and ISD data are available through the NCDC at http://www1.ncdc.noaa.gov/pub/data/ghcn/daily/ and https://www.ncdc.noaa.gov/isd/data-access, respectively.

References

Acero, F., J. A. García, M. C. Gallego, S. Parey, and D. Dacunha-Castelle (2014), Trends in summer extreme temperatures over the Iberian Peninsula using nonurban station data, *J. Geophys. Res. Atmos.*, *119*, 39–53, doi:10.1002/2013JD020590.

Alexander, L., and S. Perkins (2013), Debate heating up over changes in climate variability, *Environ. Res. Lett.*, *8*(4), 41001.

Barriopedro, D., E. M. Fischer, J. Luterbacher, R. M. Trigo, and R. García-Herrera (2011), The hot summer of 2010: Redrawing the temperature record map of Europe, *Science*, *332*(6026), 220–224.

Brown, S., J. Caesar, and C. A. Ferro (2008), Global changes in extreme daily temperature since 1950, *J. Geophys. Res.*, *113*, D05115, doi:10.1029/2006JD008091.

Cade, B. S., and B. R. Noon (2003), A gentle introduction to quantile regression for ecologists, *Front. Ecol. Environ.*, *1*(8), 412–420.

Cavanaugh, N. R., and S. S. Shen (2014), Northern Hemisphere climatology and trends of statistical moments documented from GHCN-daily surface air temperature station data from 1950 to 2010, *J. Clim.*, *27*(14), 5396–5410.

Della-Marta, P. M., M. R. Haylock, J. Luterbacher, and H. Wanner (2007), Doubled length of western European summer heat waves since 1880, *J. Geophys. Res.*, *112*, D15103, doi:10.1029/2007JD008510.

Deser, C., R. Knutti, S. Solomon, and A. S. Phillips (2012), Communication of the role of natural variability in future North American climate, *Nat. Clim. Change*, *2*(11), 775–779.

Director, H., and L. Bornn (2015), Connecting point-level and gridded moments in the analysis of climate data*, *J. Clim.*, *28*(9), 3496–3510.

Doesken, N. J. (2005), The National Weather Service MMTS (maximum-minimum temperature system)-20 years after, in *15th Applied Climatology Conference*, Am. Meteorol. Soc., Savanna, Ga. [Available online at http://ams.confex.com/ams/pdfpapers/91613.pdf.]

Donat, M., et al. (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, *J. Geophys. Res. Atmos.*, *118*, 2098–2118, doi:10.1002/jgrd.50150.

Donat, M. G., and L. V. Alexander (2012), The shifting probability distribution of global daytime and night-time temperatures, *Geophys. Res. Lett.*, *39*, L14707, doi:10.1029/2012GL052459.

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose (2010), Comprehensive automated quality assurance of daily surface observations, *J. Appl. Meteorol. Climatol.*, *49*(8), 1615–1633.

Fischer, E., and R. Knutti (2014), Detection of spatially aggregated changes in temperature and precipitation extremes, *Geophys. Res. Lett.*, *41*, 547–554, doi:10.1002/2013GL058499.

Gibson, J. F., J. D. Farmer, M. Casdagli, and S. Eubank (1992), An analytic approach to practical state space reconstruction, *Physica D*, *57*(1), 1–30.

Haylock, M., N. Hofstra, A. Klein Tank, E. Klok, P. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.*, *113*, D20119, doi:10.1029/2008JD010201.

Huntingford, C., P. D. Jones, V. N. Livina, T. M. Lenton, and P. M. Cox (2013), No increase in global temperature variability despite changing regional patterns, *Nature*, *500*(7462), 327–330.

Karl, T., et al. (2012), US temperature and drought: Recent anomalies and trends, *Eos Trans. AGU*, *93*(47), 473–474.

Koenker, R., and G. Bassett Jr. (1978), Regression quantiles, *Econometrica*, *46*, 33–50.

Machado, J. A. F., and J. S. Silva (2005), Quantiles for counts, *J. Am. Stat. Assoc.*, *100*(472), 1226–1237.

Mearns, L. O., R. W. Katz, and S. H. Schneider (1984), Extreme high-temperature events: Changes in their probabilities with changes in mean temperature, *J. Clim. Appl. Meteorol.*, *23*(12), 1601–1613.

Menne, M. J., C. N. Williams Jr, and R. S. Vose (2009), The U.S. Historical Climatology Network monthly temperature data, version 2, *Bull. Am. Meteorol. Soc.*, *90*(7), 993–1107.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An overview of the global historical climatology network-daily database, *J. Atmos. Oceanic Technol.*, *29*(7), 897–910.

Pearson, K. (1895), Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material, *Philos. Trans. R. Soc. A*, *186*, 343–414.

Reich, B. J. (2012), Spatiotemporal quantile regression for detecting distributional changes in environmental processes, *J. R. Stat. Soc.*, *61*(4), 535–553.

Rhines, A., and P. Huybers (2013), Frequent summer temperature extremes reflect changes in the mean, not the variance, *Proc. Natl. Acad. Sci.*, *110*(7), E546–E546.

Rhines, A., M. P. Tingley, K. A. McKinnon, and P. Huybers (2015), Decoding the precision of historical temperature observations, *Q. J. R. Meteorol. Soc.*, *141*(693), 2923–2933.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller (2004), The role of increasing temperature variability in European summer heatwaves, *Nature*, *427*(6972), 332–336.

Simolo, C., M. Brunetti, M. Maugeri, T. Nanni, and A. Speranza (2010), Understanding climate change-induced variations in daily temperature distributions over Italy, *J. Geophys. Res.*, *115*, D22110, doi:10.1029/2010JD014088.

Smith, A., N. Lott, and R. Vose (2011), The integrated surface database: Recent developments and partnerships, *Bull. Am. Meteorol. Soc.*, *92*(6), 704–708.

Trewin, B. (2013), A daily homogenized temperature data set for Australia, *Int. J. Climatol.*, *33*(6), 1510–1529.

Venema, V. K., et al. (2012), Benchmarking homogenization algorithms for monthly data, *Clim. Past*, *8*(1), 89–115.

Vincent, L. A., X. L. Wang, E. J. Milewska, H. Wan, F. Yang, and V. Swail (2012), A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis, *J. Geophys. Res.*, *117*, D18110, doi:10.1029/2012JD017859.

Wang, K. (2014), Sampling biases in datasets of historical mean air temperature over land, *Sci. Rep.*, *4*, 4637.

Xu, W., Q. Li, X. L. Wang, S. Yang, L. Cao, and Y. Feng (2013), Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices, *J. Geophys. Res. Atmos.*, *118*, 9708–9720, doi:10.1002/jgrd.50791.