The Effects of Gridding Algorithms on the Statistical Moments and Their Trends of Daily Surface Air Temperature*

NICHOLAS R. CAVANAUGH

*Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, and Scripps Institution of Oceanography,
University of California, San Diego, La Jolla, California*

SAMUEL S. P. SHEN

*Department of Mathematics and Statistics, San Diego State University, San Diego, and Scripps Institution of Oceanography,
University of California, San Diego, La Jolla, California*

(Manuscript received 1 October 2014, in final form 29 July 2015)

ABSTRACT

This paper explores the effects from averaging weather station data onto a grid on the first four statistical moments of daily minimum and maximum surface air temperature (SAT) anomalies over the entire globe. The Global Historical Climatology Network–Daily (GHCND) and the Met Office Hadley Centre GHCND (HadGHCND) datasets from 1950 to 2010 are examined. The GHCND station data exhibit large spatial patterns for each moment and statistically significant moment trends from 1950 to 2010, indicating that SAT probability density functions are non-Gaussian and have undergone characteristic changes in shape due to decadal variability and/or climate change. Comparisons with station data show that gridded averages always underestimate observed variability, particularly in the extremes, and have altered moment trends that are in some cases opposite in sign over large geographic areas. A statistical closure approach based on the quasi-normal approximation is taken to explore SAT's higher-order moments and point correlation structure. This study focuses specifically on relating variability calculated from station data to that from gridded data through the moment equations for weighted sums of random variables. The higher-order and nonlinear spatial correlations up to the fourth order demonstrate that higher-order moments at grid scale can be determined approximately by functions of station pair correlations that tend to follow the usual Kolmogorov scaling relation. These results can aid in the development of constraints to reduce uncertainties in climate models and have implications for studies of atmospheric variability, extremes, and climate change using gridded observations.

1. Introduction

Understanding climate variability from a probabilistic distribution point of view requires detailed information about various orders of statistical moments. Owing to spatial heterogeneity, various orders of spatial correlations or the closely related structure functions are also critical information. A variety of spatiotemporal

resolutions are needed in the climate sciences, and numerous gridded observation data products focused on the first and second moments—that is, the mean and variance—have been produced to facilitate the objective validation of climate models and studies of global climate change [e.g., NOAA OISST (Reynolds et al. 2002), GPCP (Huffman et al. 2009), and NASA GISS (Hanson et al. 2010)]. These gridded datasets are convenient for many applications, for example, model forecast validation and studies of large-scale climate variability. It is, however, still unclear exactly how the variability of gridded data is quantitatively related to weather variability observed at the weather stations that contribute to the grid-scale averages. To quantify station-level variability, Cavanaugh and Shen (2014) documented the Northern Hemisphere climatology and trends of higher-order statistical moments from the

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-14-00668.s1>.

Corresponding author address: Nicholas R. Cavanaugh, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mailstop 74R316C, Berkeley, CA 94720.
E-mail: nrcavanaugh@lbl.gov

Global Historical Climatology Network–Daily (GHCND) surface air temperature (SAT) station dataset. A comparable study of grid-scale variability is also needed to facilitate climate model validation and climate risk analysis when higher moments are important, such as when studying high-frequency variability and climate extremes.

The importance of research on climate extremes and higher-order moments cannot be overstated. Changes to climate extremes, as a manifestation of low-frequency climate variability or climate change, are an important aspect of climate research and impact many facets of people's lives. The reliable quantification of high-frequency variability and a precise description of the relevant physical mechanisms are also critical in order to improve numerical weather and climate model predictions. However, these tasks are known to be both challenging and elusive in the climate community. While monthly and longer averages are approximately Gaussian (Stephenson et al. 2004), the non-Gaussianity of daily data adds another facet of complexity. Non-Gaussian high-frequency climate processes require that more care is taken toward understanding the relevant probability distributions, higher-order moments, tail characteristics, and the probabilities of certain extremes, as well as their relationship to spatial resolution. The main purpose of this paper is to examine these properties for daily SAT anomalies over the entire globe.

Numerous gridding algorithms have been developed to translate observed point data through weighted averaging to grid-scale approximations; canonical examples include kriging-based optimal interpolation (Gandin 1963) and angular distance weighting interpolation (Shepard 1968). Considering that each observation can be thought of as a realization of a random variable (RV) at a given time and location, it is possible that the distribution of the resultant interpolated RV might be altered compared to those of the contributing RVs observed at point scale. However, when using interpolated geophysical data, it is commonly assumed that the gridding effects on variability distributions are of second order compared to the variability of the underlying mean signal and thus that the resulting gridded data can be routinely used for studies of climate variability and extremes. This paper will demonstrate that this assumption is violated for daily data and that cautions need to be taken, particularly when dealing with non-Gaussian RV distributions.

Daily SAT is a smooth spatial anomaly field with large decorrelation distances. Cavanaugh and Shen (2014) examined the climatological moments of daily SAT anomalies (T_{avg}) up to fourth order and demonstrated

that in most locations and seasons T_{avg} distributions are non-Gaussian and that their spatial patterns in moments are highly coherent on regional scales, implying that the SAT variability distributions of nearby weather stations are (nearly) identically distributed. Caesar et al. (2006a) investigated the decorrelation distances for daily maximum and minimum SAT (T_{min} and T_{max} , respectively) and showed that e -folding correlation distances are approximately 1000 km globally, with variations of up to a couple hundred kilometers according to the variable, location, and season. Caesar et al. (2006a,b) also discussed the compilation of the Met Office Hadley Centre GHCND (HadGHCND) dataset, the most widely used gridded observation SAT dataset derived from GHCND (Menne et al. 2012a,b). Our current study further examines the moments of HadGHCND SAT up to fourth order through point correlations to demonstrate the effects on variability distributions resulting from data aggregation and interpolation onto a grid. We aim to approximate the relationships between variability observed at point scale by weather stations and that estimated by averages at grid scale. To accomplish these tasks, we first develop the basic mathematical theory necessary to understand non-Gaussian variability distributions when translated from points onto a grid. We then compare the moment statistics and trends from the gridded dataset, HadGHCND, with those taken directly from the GHCND stations themselves to illustrate their differences. As a by-product of these goals, we present both station and gridded SAT datasets of mean, variance, skewness, and kurtosis for the entire global land (where observations are available) for T_{avg} , T_{min} , and T_{max} .

In section 2, we derive the mathematical theory by moments for weighted sums of dependent RVs and interpret its meaning through the lens of a turbulence theory-based moment closure approximation. Section 3 discusses the GHCND weather station dataset and the general form of gridding algorithms commonly used in practice. We also briefly describe our methodology for calculating climatological moments as well as trends over 1950–2010. Section 4 evaluates the (higher order) point correlation terms originally identified in section 2. Section 5 describes climatological moments for the gridded dataset and their comparisons to station data. We present trends in moments over 1950–2010 and show how gridded variability trends can deviate from or oppose those estimated directly from station data. Section 6 explains our findings in the broader context of multiscale climate variability, and section 7 contains conclusions and discussion.

2. Higher-order moments and their relations to higher-order point correlations

a. A “central limit theorem” by moments

SAT at a given location or grid cell can be treated as an RV with realizations in time indexed by t . For any discrete RV X , its first moment μ_1 is the expected value

$$\mu_1 = E[X] = \sum_i p_i x_i = \bar{X}, \quad (1)$$

where x_i is a random outcome with uniform probability p_i in the sample space of X (see any introductory statistics textbook; e.g., [Ross 2010](#)). The overbar represents an arithmetic ensemble mean or expected value and is equivalent to $E[\cdot]$. The n th-order moments of X , μ_n for $n \geq 2$, are defined as

$$\mu_n = E[(X - \bar{X})^n]. \quad (2)$$

The second-order moment is commonly referred to as the variance and has the same units as X^2 (and its square root is the standard deviation with the same units as X). The third- and fourth-order moments are directly related to the skewness and kurtosis, which are suitably nondimensionalized by the standard deviation [see Eqs. (10) and (11) below; also [Cavanaugh and Shen 2014](#); [Wilks 2006](#); [Ross 2010](#)]. The skewness measures the asymmetry of a distribution and is usually sufficient to demonstrate non-Gaussianity since Gaussian RVs necessarily have zero skewness. The kurtosis measures the tail thickness and peakedness of the distribution, with higher values indicating increased probabilities of large deviations from the mean in symmetric distributions. Gaussian distributions have kurtosis equal to 3.

The approximate transformation of station data into data on grids using weighted averages leads to a gridcell-based RV $\tilde{X}$, which is the weighted sum of any number of station RVs from the set $i = [1, 2, \dots, s]$ (which are now defined as the zero-centered anomalies without loss of generality) with weights w_i and a total number of stations s . We wish to explore how the moments of $\tilde{X}$, here denoted as $\dot{\mu}_n$, are quantitatively related to the data from the contributing stations.

For the mean, the expectation of the gridbox RV remains zero, unchanged from those of the station RVs, that is,

$$\dot{\mu}_1 = \sum_i w_i \bar{X}_i = \sum_i w_i \bar{X}_i = 0. \quad (3)$$

The n th-order moments $\dot{\mu}_n$ when $n \geq 2$, are more complex and are functions of the weights as well as various higher-order polynomial covariance terms, which we will refer to from here on out as n -point correlations. The

general expression for the moments [Eq. (4a)] and the expanded expressions for the second through fourth moments [Eqs. (4b)–(4d), respectively] are displayed below:

$$\dot{\mu}_n = \left(\sum_i w_i X_i \right)^n \quad \text{and} \quad (4a)$$

$$\dot{\mu}_2 = \sum_i w_i^2 \bar{X}_i^2 + \sum_{i \neq j} w_i w_j \bar{X}_i \bar{X}_j, \quad (4b)$$

$$\dot{\mu}_3 = \sum_i w_i \bar{X}_i^3 + 3 \sum_{i \neq j} w_i^2 w_j \bar{X}_i^2 \bar{X}_j + \sum_{i \neq j \neq k} w_i w_j w_k \bar{X}_i \bar{X}_j \bar{X}_k, \quad (4c)$$

and

$$\begin{aligned} \dot{\mu}_4 = & \sum_{i \neq j} w_i^4 \bar{X}_i^4 + 4 \sum_{i \neq j} w_i^3 w_j \bar{X}_i^3 \bar{X}_j + 3 \sum_{i \neq j} w_i^2 w_j^2 \bar{X}_i^2 \bar{X}_j^2 \\ & + 6 \sum_{i \neq j \neq k} w_i^2 w_j w_k \bar{X}_i^2 \bar{X}_j \bar{X}_k \\ & + \sum_{i \neq j \neq k \neq l} w_i w_j w_k w_l \bar{X}_i \bar{X}_j \bar{X}_k \bar{X}_l. \end{aligned} \quad (4d)$$

Equation (4b) defines the variance for the weighted sum of RVs and is equal to the weighted sum of all of the terms in the 2D covariance matrix of dimension $s \times s$. Similarly, Eq. (4c) defines the third-order moment for the weighted sum and is equal to the weighed sum of all the terms in the 3D coskewness matrix of dimension $s \times s \times s$. Last, Eq. (4d) defines the fourth-order moment for the weighted sum and is equal to the weighted sum of all the terms in the four-dimensional (4D) cokurtosis matrix of dimension $s \times s \times s \times s$. If $\dot{\mu}_3$ and $\dot{\mu}_4$ are normalized by $\dot{\mu}_2$ to be dimensionless, they define the skewness and kurtosis of the weighted sum, respectively. The last terms in Eqs. (4b)–(4d) are worth special attention as they are related to the colinearity among the observations. The nonzero values of these terms make it apparent that weighted aggregation is needed to form gridded data from station data. A traditional simple average using equal weights is the optimal average only for the case of a spatially white noise field where the point correlations are all zero. Climate researchers have noted colinearity in climate data for a long time. For example, empirical orthogonal functions (EOF) are a simple way to isolate data colinearity in the EOF spectral space ([North 1984](#); [Wilks 2006](#)) and to formulate optimal averaging and optimal gridding algorithms for globally inhomogeneous and correlated random fields, such as SAT and precipitation ([Shen et al. 1994, 1998, 2004, 2014](#)). While second-order covariances are familiar to the climate research community in the usual definition, the higher-order expectation matrices, here termed coskewness and cokurtosis, are less commonly seen in climate research. Additionally, their contained higher-order point correlation terms are routinely used

to examine the properties of random fields in the studies of statistical fluid mechanics and turbulence theory. For convenience, we will employ the simple notation $B_{o(i),o(j),o(k),o(l)}$ to denote point correlations terms of their respective order ($B_{2,1} = \overline{X_i^2 X_j}$, $B_{2,1,1} = \overline{X_i^2 X_j X_k}$, $B_{1,1,1,1} = \overline{X_i X_j X_k X_l}$, etc.).

Since we will see that the moments of $T_{\min}$ and $T_{\max}$ are nearly equal over large spatial areas, we can consider the RVs X_i in Eqs. (3) and (4a) to be identically distributed. In the case of independent RVs, B_2 is the only contributing term to $\dot{\mu}_2$; B_3 is the only contributing term to $\dot{\mu}_3$; and B_4 and $B_{2,2}$ are the only contributing terms to $\dot{\mu}_4$, with the dominance of $B_{2,2}$ increasing as s increases. In the special case when all X_i are perfectly dependent, then $\dot{\mu}_n = \mu_n$ and the probabilistic distribution of the weighted sum is preserved. With regard to SAT, this would imply homogeneous weather stations with identical time series. In practice, station SAT correlations usually decrease with distance (Caesar et al. 2006a) when the teleconnection range is not yet reached. As such, the Cauchy–Schwarz inequality implies that $\dot{\mu}_2 \leq \mu_2$, or more simply that the variance of the interpolated data is always less than that of the contributing stations (Shen et al. 2012). Although not rigorously proven, it is reasonable to also claim that $|\dot{\mu}_3/\dot{\mu}_2^{3/2}| \leq |\mu_3/\mu_2^{3/2}|$ and $|\dot{\mu}_4/\dot{\mu}_2^2 - 3| \leq |\mu_4/\mu_2^2 - 3|$ will hold for finite s . In short, gridded data will always be less variable and less extreme than station data.

b. The quasi-normal approximation

Here, time series of SAT anomaly measurements are approximated as stationary for each season and decade, but they are expected to be nonstationary for overlapping decades because of low-frequency variability and climate change; this treatment forms the basis for the quasi-stationarity assumption often used in studies of geophysical time series. While there are some notable counterexamples (e.g., the Cauchy distribution), for many probability distributions, the probability density function (pdf) of an RV can be determined by its moments or moment-generating function (Wackerly et al. 2008). Some pdfs can be determined by finitely many moments. For example, a Gaussian RV's distribution can be determined by the moments and point correlations up to the second order. However, for some RVs a nonzero third-order moment (as is the case for SAT) is important, indicating that higher-order expectation terms are nonnegligible and can require more simultaneous means, such as triplets and quadruplets [see Eq. (4a)]. For SAT, the first four moments are assumed estimable and are generally considered succinct measures of distribution; thus, we will truncate the terms at quadruplets in this paper.

As an early model for the relationships between point correlations for idealized turbulent fields, Millionschikov (1941) proposed that quasi-normal or quasi-Gaussian fields be characterized by vanishing fourth-order cumulants, leading to closure of the fourth-order moments of the following form:

$$\overline{X_i X_j X_k X_l} = \overline{X_i X_j} \overline{X_k X_l} + \overline{X_i X_k} \overline{X_j X_l} + \overline{X_i X_l} \overline{X_j X_k}, \quad (5)$$

implying that fourth-order correlations can be decomposed into functions of pair correlations. A thorough treatment of this theory is given by Monin and Yaglom (1971).

By combining the Cauchy–Schwarz inequality with the quasi-normal approximation [Eq. (5)], André et al. (1976a,b) showed analytically that bounds on the third-order point correlations that prohibit the nonphysical development of negative energy spectra take the form

$$|\overline{X_i X_j X_k}| \leq \min \left\{ \begin{array}{l} [\overline{X_i^2} (\overline{X_j^2} \overline{X_k^2} + \overline{X_j X_k}^2)]^{1/2}, \\ [\overline{X_j^2} (\overline{X_i^2} \overline{X_k^2} + \overline{X_i X_k}^2)]^{1/2}, \\ [\overline{X_k^2} (\overline{X_i^2} \overline{X_j^2} + \overline{X_i X_j}^2)]^{1/2} \end{array} \right\}. \quad (6)$$

Equation (6) implies that the triple correlations are bounded by functions of the smallest of the squared pair correlations for standardized variables. While the concept of quasi normality is widely considered out of date in modern turbulence literature, the relationships in Eqs. (5) and (6) remain instructive, as well as practically useful, and will be utilized in section 4 to shed light on the parallels between turbulence theory and climatological observations. In section 6, we will further discuss the quasi-normal approximation's relevance and shortcomings to the study of multiscale weather variability.

3. Data and methodology

a. Station GHCND data

GHCND contains in situ observational data from over 90 000 weather stations in 180 countries and territories worldwide (Menne et al. 2012a,b). The GHCND source data have been compiled with the goal of maximizing spatial coverage. Records from numerous agencies have been collected and quality controlled but not homogenized. Before new data are added to the dataset, they are first cross-checked with included data to confirm record uniqueness, spatial consistency, and temporal consistency (Menne et al. 2012a), which reduces the chances of redundancy. Each record is then subject to 19 quality assurance (QA) measures and flagged with appropriate tags so that data can be utilized or discarded as necessary by a user (Durre et al. 2010).

Our study leaves out any data that have been QA flagged. To define our anomaly time series, a running mean of 5 days is first calculated on the $T_{\max}$ and $T_{\min}$ records in order to fill in the holes of a few days in the climatology period 1961–90 and also to help smooth temporal inhomogeneities. Daily seasonal cycles are computed by averaging the values for each day, occurring once per year, from 5-day-centered running means over the 1961–90 period, producing daily $T_{\max}$ and $T_{\min}$ climatologies for each station. Each station's daily $T_{\max}$ and $T_{\min}$ climatology is then subtracted from the raw $T_{\max}$ and $T_{\min}$ data to produce $T_{\max}$ and $T_{\min}$ anomalies. Only stations that include at least 20 yr of data (7300 days) over the 1961–90 period are included in this study. This procedure is also used for the HadGHCND dataset, described in the next section.

b. Gridded GHCND dataset

HadGHCND is a recent attempt at producing a gridded surface observation-based climate database at daily resolution focused on providing information relevant to temperature and precipitation extremes and climate model evaluation (Caesar et al. 2006a,b). It is composed of GHCND stations that have been interpolated to a grid based on an angular distance-weighting algorithm (Alexander et al. 2006; Caesar et al. 2006a). The goal of performing weighted average interpolation is to leverage shared information summarized by the dependencies between nearby stations so that weighted combinations of data sources contribute to a more spatiotemporally complete record. Often in station records incompleteness and inconsistencies arise from equipment malfunctions, human error, and systematic changes, which manifest themselves in the record as unbelievable or missing data, instrumental drift, and discontinuous jumps, among other imperfections. Weighted interpolation seeks to aggregate stations contained within some distance scale D —in the case of HadGHCND, an e -folding correlation distance—which are thought to contain at least some component of the information that should represent a grid cell centered at a certain longitude–latitude location ϕ' and θ' and weight them as a function usually of their distance from the center of the grid cell d_i (favoring closer stations) and their geographic location ϕ_i and θ_i ,

$$X'(\phi', \theta') = \sum_i w_i(d_i, \phi_i, \theta_i) X_i; \quad d_i \leq D \quad \text{and} \quad \sum_i w_i = 1. \quad (7)$$

Since varying station density near a grid cell can bias this interpolation procedure, it is common to require a minimum number of stations (3 for HadGHCND) and

FIG. 1. Diagrammatic depiction of the HadGHCND gridding algorithm. The solid rectangle indicates the grid cell station data are interpolated to and the small solid circle indicating the grid cell center, whose longitude and latitude are ϕ' and θ' . The large dashed circle indicates the farthest extent of the search radius D , inside of which stations may be used for interpolation. Variable X_i indicates the locations of weather stations being aggregated for interpolation, whose distance to the grid center is $d_i < D$ and whose longitude and latitude are ϕ_i and θ_i .

also to limit the maximum number of contributing stations (10 for HadGHCND). Figure 1 shows a general diagrammatic representation of the HadGHCND gridding procedure. A diagrammatic depiction of this procedure is illustrated in Fig. 1. Other optimal- or variogram-based interpolation methods do not explicitly impose a D scale, but rather they utilize the covariance structure between stations to implicitly define this limit by giving lower weights to less correlated stations (Gandin 1963). The relationship of Eq. (7) to the theory outlined in section 2 is immediately evident.

c. Methods for calculating higher moments and their trends

Following Cavanaugh and Shen's (2014) approach, we calculate higher statistical moments over the climatological period 1961–90 and over the moving time window of each decade during the 1950–2010 period for both station and gridded data. We also calculate the moments' trends over the 1950–2010 period for stations that pass additional completeness criteria (see Cavanaugh and Shen 2014). We include only the $T_{\max}$ and $T_{\min}$ data streams with least 20 yr of data over the climatological period 1961–90. The data streams are separated into four seasons: December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). These seasonal records

are then decomposed into their first four moments. The arithmetic mean measures the first moment

$$\bar{X} = \frac{1}{\tau} \sum_{t=1}^{\tau} x_t, \quad (8)$$

with time index t for nonmissing data x_t contained in the 30-yr distribution for each season. We use the standard deviation to represent the second moment

$$\hat{\sigma} = \left[\frac{1}{\tau} \sum_{t=1}^{\tau} (x_t - \bar{X})^2 \right]^{1/2}, \quad (9)$$

so that the units of the first two moments are equal. The third and fourth standardized moments are characterized by the sample skewness and sample kurtosis, respectively, defined as

$$\text{Skew}(X) = \frac{1}{\tau} \sum_{t=1}^{\tau} \frac{(x_t - \bar{X})^3}{\hat{\sigma}^3} \quad \text{and} \quad (10)$$

$$\text{Kurt}(X) = \frac{1}{\tau} \sum_{t=1}^{\tau} \frac{(x_t - \bar{X})^4}{\hat{\sigma}^4}, \quad (11)$$

which are scaled by the standard deviation to be dimensionless. Note that $\bar{X} = 0$ in our case, since anomaly time series have zero mean for the entire climatology period.

To examine trends in the moments over the 1950–2010 period, the same moment decomposition methodology described above is applied to decade-length moving bins, the first being 1950–59, then 1951–60, and so on, for a total of 52 bins. Each decade is required to be at least 75% complete for its moments to be computed. The resultant 52-yr time series for each moment are then fit with a generalized linear regression model with a first-order autoregressive [AR(1)] process,

$$\mu_{t+1} = \alpha + \beta t + u_t, \quad (12a)$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad (12b)$$

$$\varepsilon_t \sim N(0, \sigma^2), \quad (12c)$$

for the seasonal moment μ for year t , with y-intercept α , slope β , error term u , AR(1) coefficient ρ , and random Gaussian noise ε of zero mean and standard deviation σ . Thus, for each moment and for each season, β summarizes the change rate per year for that moment. The AR(1) structure was chosen so that more realistic confidence intervals on the regression parameters could be calculated. We opted for a parametric approach, rather than a nonparametric approach like a Mann–Kendall test, so that the autocorrelation structure and autoregressive parameters could also be examined, although those results are not reported here.

4. SAT point correlations

Each station i that has at least 20 yr of climatological data is used in this study, both at the station level and in the gridding procedure. A maximum search radius of 4000 km is used to identify which other stations are within the proximity of the station i in question. From the subset of stations identified for station i , all possible combinations of stations are used to estimate the point correlation terms that arise in Eqs. (4b)–(4d) using the definition of the expectation in Eq. (1). This process is repeated for each contributing station (about 8200 in total) for both $T_{\min}$ and $T_{\max}$. The third- and fourth-order point correlations are subsampled in order to make the problem more computationally feasible, resulting in about 36–38 million estimates for each correlation term. For higher-order point correlations, station records are first normalized by their standard deviations to ease interpretation.

a. Pair correlation

Figure 2 displays elements contributing to the second term in Eq. (4b)—that is, the station pair correlations $B_{1,1}$ as a function of the distance between stations normalized by an interpolated e -folding correlation distance for station i , λ_i (see Fig. 3). Following Hansen and Lebedeff (1987), we separate correlation plots into 30° latitude bands. Because of the very large number of data points, station correlations are plotted here as a probability density estimated from station pairs that fall within each 0.1-wide bin spanning $d/\lambda = [0, 4]$. Solid and dashed lines represent the mean and uncertainty bounds estimated for each bin and thus can be interpreted as a continuous box-and-whisker plot.

In agreement with Hansen and Lebedeff (1987), Caesar et al. (2006a), and others, our station pair correlations for $T_{\min}$ and $T_{\max}$ decay smoothly as a function of the distance separating the stations. Also, when normalizing by a locally determined e -folding correlation distance, the decay rates globally collapse to the same functional form for both $T_{\min}$ during DJF and $T_{\max}$ during JJA, as well as for other seasons (not shown). This functional form is best approximated as a power law corresponding to the usual Kolmogorov relation. We discuss this point further in section 6. Uncertainty bounds increase as station density decreases and the magnitude and shape of the uncertainty envelope is similar across each latitude bin and for both $T_{\min}$ and $T_{\max}$.

Figure 3 shows the e -folding correlation distances for each station estimated by spline interpolation from the subset of stations within 4000 km plotted geographically. The e -folding correlation distances vary globally largely

FIG. 2. Pair correlations $\overline{X_i X_j}$ between weather station data time series X_i and X_j for (a) $T_{\min}$ during DJF and (b) $T_{\max}$ during JJA as a function of station separation distance d_{ij} normalized by a local e -folding correlation distance λ_i (illustrated in Fig. 3). Shading corresponds to the spread of observations and can be interpreted as the uncertainty. The solid black line indicates the mean correlation. Dashed lines represent the 5th and 95th percentiles estimated from observations and can be interpreted as uncertainty bounds. Means for latitude bands are further broken out by colors: 60°–90°N (red), 30°–60°N (magenta), 0°–30°N (yellow), 30°S–0° (green), and 60°–30°S (blue).

in response to regional geography, particularly for $T_{\min}$. SAT decorrelation distances tend to be larger during each stations' respective synoptically active seasons (fall, winter, and spring), which is indicative of more pronounced large-scale meteorological conditions affecting or determining SAT at those times. Additional e -folding correlation maps for each season are provided in the supplementary material (Figs. S1–S6).

b. Coskewness

Figure 4 displays elements of Eq. (4c) that contribute to the coskewness matrix, scaled by the third moment of station i to facilitate understanding. Station pairs or triplets whose scaled higher-order correlation is between -1 and 1 indicate that the third moment of the weighted sum of those stations is closer to zero than that of station i ; this also implies (in most cases) that the skewness of the weighted sum is closer to zero since the

B_3 term in Eq. (4c) decays faster than the B_2 term in Eq. (4b). Only stations whose standardized third moment is significantly different from zero, based on the standard error of skewness (see Cavanaugh and Shen 2014) are plotted. Figures 4a,b show $B_{2,1}$ correlations plotted against their respective $B_{1,1}$ correlations. The $B_{2,1}$ scales proportionately with $B_{1,1}$, where higher $B_{1,1}$ indicates that $B_{2,1}$ is nearer to the observed third moment for station i .

Figures 4c,d display the $B_{1,1,1}$ correlations plotted against the minimum pair correlation between each of the three stations. Plotting against the minimum pair correlation is inspired by the inequality in Eq. (6), which is satisfied by our observations. The $B_{1,1,1}$ scales approximately linearly with the minimum pair correlation, where higher average pair correlations indicate that triple correlations are nearer to the observed third moment for station i .

FIG. 3. Spline-interpolated local e -folding correlation distance λ for each GHCND station according to daily data for (a) $T_{\min}$ during DJF and (b) $T_{\max}$ during JJA. The color scale is limited to between 200 and 2500 km.

In all cases, normalized third-order point correlation distributions are unimodal and appear to be approximately symmetric about the mean (solid line). Uncertainty bounds for third-order terms are large relative to those of the other moments. Some portion of this is due to the method of scaling the observed third-order correlations by the third moment of station i . Third moments close to zero may inflate the ratio between the third moment and the higher-order point correlation in question, thereby amplifying the uncertainty bounds.

c. Cokurtosis

Figure 5 shows the elements of the cokurtosis matrix defined in Eq. (4d) plotted against functions of the pair correlations predicted by the quasi-normal

approximation [see Eq. (5)]. Cokurtosis terms are scaled by the fourth moment of station i ; a scaled fourth-order correlation less than 1 indicates that the fourth moment of the weighted sum of those stations is closer to 0 than that of station i ; this also signifies (in most cases) that the kurtosis for the weighted sum is closer to that of a Gaussian distribution since B_4 term in Eq. (4d) decays faster than B_2 term in Eq. (4b). In all cases, scaled $B_{3,1}$ (Figs. 5a,b), $B_{2,2}$ (Figs. 5c,d), $B_{2,1,1}$ (Figs. 5e,f) and $B_{1,1,1,1}$ (Figs. 5g,h) have a one-to-one relationship with their respective quasi-normal pair correlation function and have very small error bounds relative to the equality signal. This result signifies that SAT at stations can be locally (i.e., up to grid scale) approximated as quasi normal with fourth-order moment closure [see Eq. (5)].

FIG. 4. Higher-order correlation terms in coskewness matrices for (left) $T_{\min}$ during DJF and (right) $T_{\max}$ during JJA normalized by the third moment $\bar{X}^3$ of station i plotted as functions of pair correlations (a),(b) $B_{2,1}$ and (c),(d) $B_{1,1,1}$. Figure shading and lines as in Fig. 2.

5. Observed multiscale SAT variability

In light of the observed dependence measures documented in section 4, we examine the climatological SAT moments estimated from GHCND seasonal data during 1961–90 and compare these distribution characteristics to those computed using HadGHCND. The trends of these moments over 1950–2010 computed from GHCND and HadGHCND datasets are also compared. Significant disagreement between trends and/or trend patterns exists between the two datasets. This paper only shows the results for $T_{\min}$ during DJF and $T_{\max}$ during JJA to maximize interest among the climate community; however, additional plots for all seasons are included in the supplementary material (Figs. S13–S18).

a. Climatological $T_{\min}$ and $T_{\max}$

Figures 6 and 7 illustrate the spatial distribution climatological moments for $T_{\min}$ during DJF and $T_{\max}$ during JJA, respectively. The left columns of Figs. 6 and 7 are the moments computed from GHCND station data, and the right columns are the moments computed from the gridded HadGHCND data minus those from the GHCND data. The station anomaly pdfs for both $T_{\min}$ and $T_{\max}$ are regionally approximately identical, as evidenced by the smooth and coherent spatial structure of moments at regional scales (Figs. 6 and 7a,c,e; discussed in detail in Cavanaugh and Shen 2014). Additionally, distributions of these SAT variables are non-Gaussian and have somewhat complex spatial structures that are

influenced by regional geography and land–atmosphere characteristics.

Immediately evident from Figs. 6 and 7b is that the standard deviations estimated from gridded observations are greatly reduced from those estimated directly from station data. Globally, reductions in standard deviations are largest in regions where station density is low and lowest where station density is high. The clearest example of this can be observed at the border between the United States and Mexico, where station density goes from high to low when crossing the border going south. This effect is due to the contribution of the $B_{1,1}$ terms in Eq. (4b); in areas with high station density, only stations that are geographically close to the center of the grid cell are used in interpolation, and thus their pair correlations are close to one (see Figs. 2a,b). Large reductions in standard deviation arise in large-scale regions where the standard deviations are highest in the climatological record—for example, $T_{\min}$ (Fig. 6a) in northern Europe and Russia and $T_{\max}$ (Fig. 7a) in the northwestern United States, northern Europe, and Russia. These reductions occur resultant from the $B_{1,1}$ terms in Eq. (4b), where the magnitude of the reduction should arise proportionally to the climatological standard deviations of the contributing stations. There are also more localized regions where there are large reductions in standard deviation—for example, $T_{\min}$ along the west coast of North America and $T_{\max}$ on the Iberian Peninsula and in Mexico. These reductions result from reduced $B_{1,1}$ terms in Eq. (4b) arising from the short decorrelation distances in those areas (Figs. 3a,b).

FIG. 5. Higher-order correlation terms in cokurtosis matrices for (left) $T_{\min}$ during DJF and (right) $T_{\max}$ during JJA normalized by the fourth moment $\overline{X^4}$ of station i plotted as functions of the pair correlations prediction by the quasi-normal approximation for (a),(b) $B_{3,1}$; (c),(d) $B_{2,2}$; (e),(f) $B_{2,1,1}$; and (g),(h) $B_{1,1,1,1}$. Figure shading and lines as in Fig. 2.

In some relatively rare cases, the climatological standard deviation from the gridded data is slightly increased relative to the station record. These increases might arise as a result of local-scale station inhomogeneity [e.g., those arising from variable station exposures or wind-chill effects, see Fall et al. (2011), and/or landscapes and land-cover variability, see Montandon et al. (2011)], weighting effects of the chosen interpolation algorithm (which might occur, for instance, in the case where poorly sited stations are

isolated in space), and/or from nonstatic observation records or recording techniques, all of which are likely the major contributors in the United States and western Europe. In other areas that are less station dense, station inhomogeneity within a large decorrelation length scale is the more likely contributor, as is probably the case, for example, in the Middle East (see Fig. 6b).

At large scale, skewness values in the gridded dataset for both $T_{\min}$ and $T_{\max}$ (Figs. 6 and 7d) tend to be closer to zero than in the station observations (i.e., the

FIG. 6. (left) Station-level climatological moments and (right) the bias of gridded data for T_{min} during DJF: (a),(b) standard deviation, (c),(d) skewness, and (e),(f) kurtosis. The biases at each station are represented as the moments of the HadGHCND grid cell minus the moments from station data inside the same grid cell. Grid cells in (c) and (e) are marked if at least 50% of the contained stations have moments that significantly differ from 0 based on the standard error of skewness and kurtosis at the 5% significance level.

deviations are negatively correlated with the station observations illustrated in Figs. 6 and 7c, indicating that gridded variability is more symmetric than in the stations themselves. Spatial correlations between Figs. 6c,d and 7c,d are -0.435 and -0.502 , respectively. Similarly, kurtosis deviations (Figs. 6 and 7f) also tend to be negatively correlated with observations (Figs. 6 and 7e), indicating that the peakedness of gridded variability is closer to that of a Gaussian distribution than the stations themselves, as one would expect from Eq. (4d). Spatial correlations between Figs. 6e,f and 7e,f are -0.532 and -0.559 , respectively. Both of these effects can be

anticipated from Figs. 4 and 5. As can be deduced by comparing the broad distributions of third-order moments (Fig. 4) to the narrow distributions of the fourth-order moments (Fig. 5), the magnitude of the anticorrelations between station skewness and gridded skewness are smaller than those of kurtosis. Gridded skewness and kurtosis biases tend to be largest both where the magnitudes of the station moments are the largest and where the decorrelation distances are the shortest. It is difficult to identify the more dominant contribution, since these regions tend to overlap when comparing the station climatologies (Figs. 6 and 7c,e) to the e -folding distances

FIG. 7. As in Fig. 6, but for T_{max} during JJA.

(Figs. 3a,b). These results can be expected from Figs. 4 and 5. Third- and fourth-order point correlations are related to pair correlations of the contributing stations and are thus related to the distances separating the stations used for interpolation. Since we have chosen to normalize the higher-order correlations by the third and fourth moments of station i , in each instance where the normalized third- or fourth-order point correlation is less than one (which is where the bulk of the observations lie), we can expect from Eqs. (4a) and (5) that the higher-order moments of the gridded observations should be closer to those of a Gaussian. Similar to the explanation of results regarding the second moment, gridded data are most similar to Gaussian both where station density and/or local decorrelation distances (Figs. 3a,b) are low.

Weighted area averaging over homogeneous regions—that is, gridding data—minimizes bias with respect to the mean at the cost of changes to the distribution around it. While it is commonplace to use gridded data in studies of extremes, reductions and modulations to the observed variability resultant from area averaging have the largest impacts on the tails of pdfs that dictate extremal probabilities. In the least biased scenario of approximately Gaussian station distributions with very high station density (e.g., over parts of the United States), the probability of exceeding two standard deviations at the station level is generally about twice that estimated from gridded data located over the same region (not shown). By contrast, in greatly skewed regions with low station density (e.g., over parts of Russia), these probabilities can exceed

FIG. 8. (left) Trends in moments over the 1950–2010 period for $T_{\min}$ during DJF computed from GHCND data. and (right) trends in moments for stations differenced from the trends in moments estimated from the HadGHCND dataset: (a),(b) standard deviation, (c),(d) skewness, and (e),(f) kurtosis. The trend biases at each station are represented as the moment trends of the HadGHCND grid cell minus the moment trends from station data inside the same grid cell. Grid cells for (a),(c),(e) are marked if at least 50% of the contained stations have statistically significant trends estimated using an AR(1) procedure at the 5% significance level. Grid cells for (b),(d),(f) are marked if the direction of the trend for the gridded data opposes the direction of the average trend of each station within the same grid cell.

5 times those estimated from gridded data located over the same region (not shown). It is thus quite important to stress that caution must be taken when utilizing gridded data to study climate extremes.

b. Moment trends 1950–2010

The observed climatological distributions of SAT are non-Gaussian; additionally, the differences in moments between station and gridded SAT data are of the same order of magnitude as the moments themselves (Figs. 6

and 7). The first result necessitates the inclusion of higher-order moment effects in studies of nonstationary high-frequency climate variability. The second result suggests that trends in the moments of gridded data might also differ from those of the station data, potentially polluting inferences concerning the effects of climate change on weather variability. This section describes linear trends in the moments of $T_{\min}$ during DJF and $T_{\max}$ during JJA from 1950 to 2010 for both station data and gridded data using the methodology

FIG. 9. As in Fig. 8, but for $T_{\max}$ during JJA.

described in section 3c (for details see Cavanaugh and Shen 2014). We would like to explore what effect a gridding algorithm may have on trends compared with station data.

Figures 8 and 9 illustrate trends in moment statistics in the left column and biases in these trends caused by interpolation to a grid in the right column. Both $T_{\min}$ and $T_{\max}$ show large-scale spatially coherent trends in moments. For $T_{\min}$, the standard deviation trends are largely negative over the Northern Hemisphere, indicating less variability in winter nighttime temperatures. Surprisingly, trends in standard deviations for gridded data are mostly more positive than the trends from the station data; in many cases (particularly over Russia, western Canada, and Alaska), the direction of

the trends oppose each other. Standard deviation trends from $T_{\max}$ are more mixed globally; although again, in many cases (particularly over the northern United States and western Russia), gridded trends oppose those taken from the stations directly.

SAT trends in the skewness and kurtosis from station data are mixed globally, but they are largely spatially coherent and are in many regions significantly nonzero. For $T_{\min}$, the largest magnitude trends in skewness tend to be positive (over the northwestern and southeastern United States and in western Europe) and also tend to coincide with the areas of dominant negative trends in kurtosis. For $T_{\max}$, higher-order moment trends are variable at the regional level; and in some regions, areas with large trends in skewness and kurtosis tend to

coincide. Under these circumstances, this may indicate a particularly large change in the structure of SAT extremes at these locations over decadal time scales. Similar to standard deviations, biases in the gridded data higher-order moment trends are mostly regionally coherent, and in some cases, they contribute to trends that oppose those observed in the station records themselves. These differences in moment trends are obviously of practical significance when considering the likelihood of events in a nonstationary climate.

While it is possible that intrinsic variability at grid scale could have different trends than variability at point scale, this possibility is unlikely. As such, the opposite signs of the trends observed in station data and gridded data in some regions make it necessary to identify the mechanisms responsible for those differences, although pinpointing situation-specific mechanisms is a very challenging task because many factors appear to contribute simultaneously. A few of these mechanisms are listed below.

- (i) The number of stations contributing to the area average increases (decreases) over the trend period. Trends in gridded standard deviations are more negative (positive) than in the stations themselves. Trends in higher moments are biased toward becoming closer to (farther from) Gaussian in moments than in the stations themselves. Both changes are resultant from an increased (decreased) contribution of off-diagonal terms in Eqs. (4b)–(4d).
- (ii) The distances from the grid center to the stations contributing to the area average increase (decrease) over the trend period due to an increase in station density, resulting in greater station heterogeneity (homogeneity). Trends in gridded standard deviations are more negative (positive) than in the stations themselves. Trends in higher moments are biased toward becoming closer to (farther from) Gaussian in moments than in the stations themselves. Both changes are resultant from a decreased (increased) magnitude of off-diagonal terms in Eqs. (4b)–(4d), assuming stations are uniformly distributed.
- (iii) The distances between stations contributing to the area average increase (decrease) over the trend period due to a spatial redistribution of stations, resulting in greater station heterogeneity (homogeneity). Trends in gridded standard deviations are more negative (positive) than in the stations themselves. Trends in higher moments are biased toward becoming closer to (farther from) Gaussian in moments than in the stations themselves. Both

changes are resultant from a decreased (increased) magnitude of off-diagonal terms in Eqs. (4b)–(4d).

- (iv) The station weights contributing to the area average change over time as a result of changing station locations and data availability. Gridded moment trends can be biased in either direction.
- (v) Correlations (higher order) between stations contributing to the area average are nonstationary. Gridded moment trends can be biased in either direction.

Forcing the interpolation scheme to be constant in time might partially alleviate these artifacts; however, it might also come at the cost of reducing accuracy with respect to the probabilistic mean and sacrificing some potentially useful information.

6. General implications for grid-scale SAT variability

While weather station observations from the GHCND dataset are taken at approximately one point in space and time, the moment statistics presented in section 4 can shed additional light on larger-scale SAT variability. Specifically, the observed moment statistics, which can be approximated as quasi normal, suggest that the first four moments are sufficient in order to approximate SAT distributions up to grid scale. The descriptive moment climatologies presented in Figs. 6 and 7 (and in Figs. S7–S12 of the supplementary material) quantify the upper bounds on the degree of non-Gaussianity in moments for these SAT distributions since averaging station data results in distributions whose moments closer resemble those of a Gaussian. As discussed in section 4, spatial averages over homogeneous areas, which should approximate the means in probability over those areas, are merely weighted sums (or integrals with a weighting function) of dependent RVs. The resulting distributions will tend to be closer to Gaussian than the distributions at point locations in space, an idea encapsulated by the central limit theorems for weakly dependent RVs (e.g., the small block–big block concept; Bernstein 1927).

By treating stations as approximately quasi normal in an interpolation algorithm, the implication is that by keeping track of which stations contribute to averages in gridding algorithms, their respective assigned weights, and their higher-order correlations, a mapping between the averaged variability and variability at a point in space can be deduced. For spatial averages, this of course can practically be accomplished (approximately) by merely keeping track of the distances between contributing stations. This result is critical for understanding statistical downscaling and helps explain its success when applied to

SAT since it is now clear how spatial variability is approximately related to variability at a point.

A very important shortcoming of the quasi-normal approximation is that it is at odds with the power-law structure observed ubiquitously in geophysics. Indeed, the normalized pair correlations illustrated in Fig. 2 decay most closely as a power law, not as an exponential as would be implied by a truly quasi-Gaussian random field. The raw pair correlations decay on average as a $-2/3$ power law (not shown) but vary geographically largely as a function of elevation and/or topography, where more mountainous regions tend to decay more quickly; analogously, SAT second-order structure functions are characterized on average as having a $2/3$ power law (not shown) prior to a scale break at approximately the e -folding correlation distance. These observed power-law exponents can be associated with a Hurst exponent of $1/3$ and a spectral slope corresponding to the classical Kolmogorov $-5/3$ scaling for isotropic 3D turbulence. Our analysis does not indicate a scale-break or transition region sometimes observed at the edge of the mesoscale in aircraft and model data, which is consistent with the arguments of Lovejoy et al. (2009) and Pinel et al. (2012). Station data follow the same power-law weather regime until transitioning to the macro-weather regime (Lovejoy 2013) at approximately 2000 km, depending on location. Lovejoy and Schertzer (2013) provide a comprehensive review of geophysical turbulence theory, multifractal power-law cascades, and observed scaling laws in the oceans and atmosphere.

7. Conclusions and discussion

We have taken a bottom-up approach toward understanding variability of SAT in weather station and gridded observations through an examination of the first four statistical moments. We focus on observations of $T_{\min}$ during DJF and $T_{\max}$ during JJA as examples of random fields observed approximately as points in space and instantaneously in time. We first derive the moments for weighted sums of dependent RVs, which contain both linear and nonlinear correlations up to fourth order. We then examine these higher-order correlations through the lens of the quasi-normal approximation where third- and fourth-order point correlations are analytically related to the pair correlations of contributing stations. We also report on station pair correlation decay with distance, and we illustrate the spatial distribution of pair correlation e -folding distance, globally.

Changes to moment statistics when moving from point scale to grid scale are expected from the observed higher-order point correlations and from scaling theory.

We compare a gridded observation dataset with station data as points of reference to examine these effects. The second moments of gridded observation data are greatly reduced, globally, compared to the second moments of contributing stations. Higher-order moments become closer to those of a Gaussian distribution than in the contributing stations themselves. More plainly, gridded data are always less variable and less extreme than station data. It is shown that these changes can be expected from mathematical theory, and that they are physically related to the distances between the contributing stations in a gridding algorithm.

We also examine trends in moments over a 1950–2010 regression period and show that gridding algorithms that change over time can contribute to trends in moments that can deviate from (or even oppose) those estimated from station data directly. These results in particular are relevant to assessments of weather risk and the study of extremes within the context of climate change. As an example, wintertime trends in SAT standard deviation are significantly negative over much of the world and appear to be driven by reduced variability in nighttime temperatures (which also has significantly negative trends in standard deviation). This result is consistent with climate change theory. It is uncertain at this time whether these differences arise from biases due to gridding algorithms, discussed in section 5b, or whether distribution trends are or should be fundamentally different across spatiotemporal scales.

This article aims both to shed light on SAT variability and to demonstrate some potential pitfalls in the analysis of high-frequency climate variability in general, particularly when studying extremes. Advances in turbulence theory—in particular, a deeper understanding of power-law scaling relationships in spatiotemporally dense and uniform atmospheric data in the weather regime (e.g., Lovejoy and Schertzer 2011)—have largely addressed the problem of relating variability observed at different scales in gridded random fields, including extensions to the higher-order moments, point correlations, and structure functions. However, further extending these theories to help address practical problems routinely faced by applied atmospheric scientists and meteorologists still requires additional research. In the case of this paper, we develop and explore an approximation to relate the distributions of spatiotemporally sparse and nonuniform in situ weather observations to the distributions of gridded observations most commonly used in applied research; and as a result, we quantitatively show how network heterogeneity and nonstationarity can lead to bias and inconsistencies between the two forms of data.

An overview of current gridding methods for daily data is available in Shen et al. (2001); however, it may be important to modify current or innovate new methods in order to optimize the study of high-frequency variability. Further research is needed to identify robust and unbiased methods for studying high-frequency climate variability and extremes, particularly in the context of climate change. For example, index-based gridded datasets for extremes, such as the recently released gridded temperature and precipitation climate extremes indices (Donat et al. 2013), provide a promising avenue for distributionally consistent examinations of high-frequency climate variability. Since station variability distributions can reasonably be considered identical at subgrid scale, it is more sensible to calculate statistics (moments, quantiles, maximum/minimum values, etc.) from the stations themselves and then average as opposed to vice versa, when the goal is to study weather variability as it applies to society, since humans do not observe average weather conditions on a grid. A recently proposed distribution-preserving gridding algorithm, the hybrid interpolation technique (Shen et al. 2001), is also well supported by these analyses, particularly if it is extended to include the higher moments. Because of changes in the historical observational network (e.g., number of stations and locations of stations), the gridbox data's aggregation weight for each used station varies with respect to time and thus does not guarantee that the aggregated data's statistical moments are not systematically affected by network changes. It is therefore important to account for and eliminate the network's effect when using gridded observations to study climate variability and change.

Acknowledgments. This research was supported in part by the National Science Foundation (NSF) under Grants AGS-1015957, AGS-1015926, AGS-1419526, OCE0960770, and OCE1419306. The authors thank the three anonymous reviewers and the editor for their comments, which greatly helped improve the quality of the manuscript.

REFERENCES

- Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, **111**, D05109, doi:10.1029/2005JD006290.
- André, J. C., G. De Moor, P. Lacarrère, and R. du Vachat, 1976a: Turbulence approximation for inhomogeneous flows: Part I. The clipping approximation. *J. Atmos. Sci.*, **33**, 476–481, doi:10.1175/1520-0469(1976)033<0476:TAFIFP>2.0.CO;2.
- , —, —, and —, 1976b: Turbulence approximation for inhomogeneous flows: Part II. The numerical simulation of a penetrative convection experiment. *J. Atmos. Sci.*, **33**, 482–491, doi:10.1175/1520-0469(1976)033<0482:TAFIFP>2.0.CO;2.
- Bernstein, S., 1927: Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.*, **97**, 1–59, doi:10.1007/BF01447859.
- Caesar, J., L. Alexander, and R. Vose, 2006a: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *J. Geophys. Res.*, **111**, D05101, doi:10.1029/2005JD006280.
- , —, and —, 2006b: HadGHCND. Met Office Hadley Centre, accessed 28 May 2014. [Available online at <http://www.metoffice.gov.uk/hadobs/hadghcnd/download.html>.]
- Cavanaugh, N. R., and S. S. P. Shen, 2014: Northern Hemisphere climatology and trends of statistical moments documented from GHCN-Daily surface air temperature station data from 1950 to 2010. *J. Climate*, **27**, 5396–5410, doi:10.1175/JCLI-D-13-00470.1.
- Donat, M. G., L. V. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar, 2013: Global land-based datasets for monitoring climatic extremes. *Bull. Amer. Meteor. Soc.*, **94**, 997–1006, doi:10.1175/BAMS-D-12-00109.1.
- 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. Meteor. Climatol.*, **49**, 1615–1633, doi:10.1175/2010JAMC2375.1.
- Fall, S., A. Watts, J. Nielsen-Gammon, E. Jones, D. Niyogi, J. R. Christy, and R. A. Pielke, 2011: Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends. *J. Geophys. Res.*, **116**, 1–15, doi:10.1029/2010JD015146.
- Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields*. R. Hardin, Ed., Israeli Program for Scientific Translations, 242 pp.
- Hansen, J., and S. Lebedeff, 1987: Global trends of measured surface air temperature. *J. Geophys. Res.*, **92**, 13 345–13 372, doi:10.1029/JD092iD11p13345.
- , R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, 1–29, doi:10.1029/2010RG000345.
- Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global precipitation record: GPCP Version 2.1. *Geophys. Res. Lett.*, **36**, L17808, doi:10.1029/2009GL040000.
- Lovejoy, S., 2013: What is climate? *Eos, Trans. Amer. Geophys. Union*, **94**, 1–2, doi:10.1002/2013EO010001.
- , and D. Schertzer, 2011: Space-time cascades and the scaling of ECMWF reanalyses: Fluxes and fields. *J. Geophys. Res.*, **116**, D14117, doi:10.1029/2011JD015654.
- , and —, 2013: *The Weather and Climate: Emergent Laws and Multifractal Cascades*. Cambridge University Press, 475 pp.
- , A. F. Tuck, D. Schertzer, and S. J. Hovde, 2009: Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys. Discuss.*, **9**, 3871–3920, doi:10.5194/acpd-9-3871-2009.
- Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012a: An overview of the Global Historical Climatology Network-Daily database. *J. Atmos. Oceanic Technol.*, **29**, 897–910, doi:10.1175/JTECH-D-11-00103.1.
- , and Coauthors, 2012b: Global Historical Climatology Network - Daily (GHCN-Daily), version 3. NOAA National Climatic Data Center, accessed 23 April 2014, doi:10.7289/V5D21VHZ.
- Millionshchikov, M. D., 1941: On the theory of homogeneous isotropic turbulence. *Dokl. Akad. Nauk SSSR*, **32**, 611–614.
- Monin, A. S., and A. M. Yaglom, 1971: *Statistical Fluid Mechanics: Mechanics of Turbulence*. Vol. 1. The MIT Press, 770 pp.

- Montandon, L. M., S. Fall, R. A. Pielke, and D. Niyogi, 2011: Distribution of landscape types in the Global Historical Climatology Network. *Earth Interact.*, **15**, 1–24, doi:[10.1175/2010EI371.1](https://doi.org/10.1175/2010EI371.1).
- North, G., 1984: Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, **41**, 879–887, doi:[10.1175/1520-0469\(1984\)041<0879:EOFANM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1984)041<0879:EOFANM>2.0.CO;2).
- Pinel, J., S. Lovejoy, D. Schertzer, and A. F. Tuck, 2012: Joint horizontal-vertical anisotropic scaling, isobaric and isoheight wind statistics from aircraft data. *Geophys. Res. Lett.*, **39**, L11803, doi:[10.1029/2012GL051689](https://doi.org/10.1029/2012GL051689).
- Reynolds, R., N. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, doi:[10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2).
- Ross, S., 2010: *A First Course in Probability*. 8th ed. Prentice Hall, 530 pp.
- Shen, S. S. P., G. R. North, and K.-Y. Kim, 1994: Spectral approach to optimal estimation of the global average temperature. *J. Climate*, **7**, 1999–2007, doi:[10.1175/1520-0442\(1994\)007<1999:SATOE0>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1999:SATOE0>2.0.CO;2).
- , T. Smith, C. Ropelewski, and R. E. Livezey, 1998: An optimal regional averaging method with error estimates and a test using tropical Pacific SST data. *J. Climate*, **11**, 2340–2350, doi:[10.1175/1520-0442\(1998\)011<2340:AORAMW>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2340:AORAMW>2.0.CO;2).
- , P. Dzikowski, G. Li, and D. Griffith, 2001: Interpolation of 1961–97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. *J. Appl. Meteor.*, **40**, 2162–2177, doi:[10.1175/1520-0450\(2001\)040<2162:IODTAP>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<2162:IODTAP>2.0.CO;2).
- , A. N. Basist, G. Li, C. Williams, and T. R. Karl, 2004: Prediction of sea surface temperature from the global historical climatology network data. *Environmetrics*, **15**, 233–249, doi:[10.1002/env.638](https://doi.org/10.1002/env.638).
- , C. K. Lee, and J. Lawrimore, 2012: Uncertainties, trends, and hottest and coldest years of U.S. surface air temperature since 1895: An update based on the USHCN V2 TOB data. *J. Climate*, **25**, 4185–4203, doi:[10.1175/JCLI-D-11-00102.1](https://doi.org/10.1175/JCLI-D-11-00102.1).
- , N. Tafolla, T. M. Smith, and P. A. Arkin, 2014: Multivariate regression reconstruction and its sampling error for the quasi-global annual precipitation from 1900 to 2011. *J. Atmos. Sci.*, **71**, 3250–3268, doi:[10.1175/JAS-D-13-0301.1](https://doi.org/10.1175/JAS-D-13-0301.1).
- Shepard, D., 1968: A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference*, ACM Publ. P-68, Brandon/Systems Press, 517–524.
- Stephenson, D. B., A. Hannachi, and A. O'Neill, 2004: On the existence of multiple climate regimes. *Quart. J. Roy. Meteor. Soc.*, **130**, 583–605, doi:[10.1256/qj.02.146](https://doi.org/10.1256/qj.02.146).
- Wackerly, D., W. Mendenhall, and R. Scheaffer, 2008: *Mathematical Statistics with Applications*. 7th ed. Cengage Learning, 912 pp.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 91, Academic Press, 627 pp.