Title
Observation-based Simulations of Humidity and Temperature Using Quantile Regression

Journal
Journal of Climate, 33(24)

ISSN
0894-8755

Authors
Poppick, Andrew
McKinnon, Karen A

Publication Date
2020

DOI
10.1175/jcli-d-20-0403.1

Peer reviewed

Observation-Based Simulations of Humidity and Temperature Using Quantile Regression

ANDREW POPPICK[a] AND KAREN A. MCKINNON[b]

[a] *Carleton College, Northfield, Minnesota*
[b] *University of California, Los Angeles, Los Angeles, California*

(Manuscript received 2 June 2020, in final form 11 August 2020)

ABSTRACT: The human impacts of changes in heat events depend on changes in the joint behavior of temperature and humidity. Little is currently known about these complex joint changes, either in observations or projections from general circulation models (GCMs). Further, GCMs do not fully reproduce the observed joint distribution, implying a need for simulation methods that combine information from GCMs with observations for use in impact studies. We present an observation-based, conditional quantile mapping approach for the simulation of future temperature and humidity. A temperature simulation is first produced by transforming historical temperature observations to include projected changes in the mean and temporal covariance structure from a GCM. Next, a humidity simulation is produced by transforming humidity observations to account for projected changes in the conditional humidity distribution given temperature, using a quantile regression model. We use the Community Earth System Model Large Ensemble (CESM1-LE) to estimate future changes in summertime (June–August) temperature and humidity over the continental United States (CONUS), and then use the proposed method to create future simulations of temperature and humidity at stations in the Global Summary of the Day dataset. We find that CESM1-LE projects decreases in summertime humidity across CONUS for a given deviation in temperature from the forced trend, but increases in the risk of high dewpoint on historically hot days. In comparison with raw CESM1-LE output, our observation-based simulation largely projects smaller changes in the future risk of either high or low humidity on days with historically warm temperatures.

KEYWORDS: Climate change; Climate variability; Humidity; Surface temperature; Statistical techniques

1. Introduction

Assessing the potential societal impacts of changes in future heat events requires an understanding of projected changes in both temperature and humidity. For example, high humidity is a contributor to human heat stress during heat events (Barreca 2012) and some projections show substantial increases by the end of century in the risk of humid heat events that exceed theoretical limits on the human body's ability to self-regulate temperature through evaporative cooling (Pal and Eltahir 2016; Coffel et al. 2017). See also Buzan and Huber (2020) for a recent review of the impacts of moist heat stress and its projected changes. In contrast, hot and dry conditions increase wildfire risk (Seager et al. 2015), among other impacts. Changes in hot and humid or hot and dry events can be affected by distributional changes beyond the means of each variable, such as changes in the underlying local relationship between temperature and humidity as well as changes in variability at multiple scales, implying a need for methods that are sensitive to these potentially complex changes.

There is limited existing work that addresses local joint changes in temperature and humidity either in observations or in general circulation models (GCMs). Many recent studies focus on univariate summaries that may be useful for a particular impact of interest, such as wet bulb or wet bulb globe temperature (as an indicator for human comfort) (Knutson and Ploshay 2016; Pal and Eltahir 2016; Coffel et al. 2017; Li et al. 2017; Lee and Min 2018; Li et al. 2020) or vapor pressure deficit (as an indicator for crop health or wildfire risk) (Seager et al. 2015; Hsiao et al. 2019). While univariate summaries can be useful for studying particular impacts, a more general approach is desirable and requires multivariate methods. Studies that do consider changes in both temperature and humidity have tended to focus on limited quantities, such as changes in temperature and humidity on the 1% warmest days (Fischer and Knutti 2013), for monthly averages (Simmons et al. 2010), in univariate quantities during specific definitions of hot and dry or hot and humid events (Schoof et al. 2017), or at a small number of spatial locations (Pryor and Schoof 2016; Yuan et al. 2020). Fischer and Knutti (2013) recognized a need for more work in understanding joint changes in temperature and humidity, but a detailed understanding remains lacking in the literature.

Moreover, future impact studies may require not only an understanding of projected changes in temperature and humidity in GCMs, but also realistic bivariate simulations of these variables. It is well understood that raw GCM output is insufficient for these purposes, because GCM output forced with historical forcings does not fully reproduce observed climate variable distributions; see John and Soden (2007), Brands et al. (2013), Tian et al. (2013), and Zhao et al. (2015) for examples specifically evaluating GCM simulations of temperature and humidity or heat stress [see also IPCC (2013, ch. 9)]. This fact is not specific to temperature or humidity simulations, and a number of methods have been proposed to combine

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0403.s1.

Corresponding author: Andrew Poppick, apoppick@carleton.edu

DOI: 10.1175/JCLI-D-20-0403.1

observations with model output to produce better calibrated future simulations, typically labeled ''bias correction'' methods [see, e.g., Ho et al. (2012), Hawkins et al. (2013), and Cannon et al. (2020) for reviews of the main types of methods]. Ho et al. (2012) separate popular methods into two classes: those that modify GCM output in an attempt to correct biases with observations, and those that modify observations in an attempt to account for GCM projected changes. Because the term *bias correction* connotes the former but not the latter type of methods, we instead refer to these methods as ''model-based'' or ''observation-based'' simulation methods. While no simulation method resolves all potential defects of future simulations (Dixon et al. 2016; Lanzante et al. 2018), observation-based simulations have the attractive property that they easily preserve most of the higher-order behavior of observational distributions and generally require more statistical modeling of distributions in GCM output (where available data are typically abundant and signal-to-noise ratios are relatively high) than in observations (where data are typically limited and signal-to-noise ratios are lower). Recent work has extended observation-based simulation methods to incorporate more complex changes from GCM output, such as distinct variability changes at different time scales (Leeds et al. 2015; Poppick et al. 2016) or seasonally varying changes to the full marginal distribution of a climate variable (Haugen et al. 2019).

Most of the aforementioned simulation methods are univariate, but there has been a recent increase in proposed methods for multivariate simulations. Indeed, Zscheischler et al. (2019) emphasizes the importance of multivariate simulation methods, using temperature and humidity as one example where there is such a need, because separately producing simulations of each variable fails to address intervariable dependencies that can have relevant effects on impacts studies. Recently, Schoof et al. (2019) proposed a model-based method for temperature and humidity and, to our knowledge, other existing bivariate or multivariate methods are model-based [e.g., Piani et al. 2010; Vrac and Friederichs 2015; Mehrotra and Sharma 2015, 2016; Cannon 2018; Vrac 2018; Guo et al. 2019; see also François et al. (2020) for a review of some methods] and so may struggle to realistically simulate higher-order distributional features that are not explicitly corrected in the simulation procedure. We therefore see a benefit in developing observation-based multivariate simulation methods.

This work makes two main contributions to the literature on joint changes in temperature and humidity. One is an analysis of projected changes in the relationship between summertime daily temperature and humidity over the continental United States (CONUS) in the Community Earth System Large Ensemble (CESM1-LE) (Kay et al. 2015). Our methodology is based on a quantile regression (Koenker and Bassett 1978) model for humidity given local and global temperature, allowing for a flexible and unifying approach to studying the relationship between humidity and temperature in different parts of the humidity distribution (i.e., dry vs humid days) over the course of the summer, and changes thereof with increasing global mean temperature. Quantile regression has been used previously in the climate literature, for example, to study distributional changes in observed temperatures (Reich 2012; Matiu et al. 2016; McKinnon et al. 2016; Rhines et al. 2017; Gao and Franzke 2017) and in temperatures from GCM output (Haugen et al. 2018), and recently, changes in the relationship between temperature and humidity in observations (McKinnon and Poppick 2020) and for a limited number of CONUS grid cells within CESM1-LE but without an explicit model for changes (Yuan et al. 2020).

The second and primary contribution of this work is a proposed observation-based bivariate simulation method for daily summertime humidity and temperature. A temperature simulation is first produced by transforming historical temperature observations to account for GCM projected changes in mean and temporal covariance, using a method based on those proposed in Leeds et al. (2015) and Poppick et al. (2016). A simulation of humidity is then produced by transforming humidity observations to account for changes both in temperature itself as well as in the underlying relationship between humidity and temperature. Our methodology relies on the aforementioned quantile regression analysis and can be thought of as an observation-based ''quantile-mapping'' approach to simulation [as in Haugen et al. (2019)]. To our knowledge, the proposed method is the first such observation-based method that can account for bivariate changes and is built specifically for simulating temperature and humidity.

The remainder of this paper is organized as follows. In section 2, we define the variables in our analysis and describe both the GCM data used to study changes in the relationship between humidity and temperature and the observations we use for producing our observation-based simulation. In section 3, we describe our method for producing an observation-based temperature simulation, which is an input into our humidity simulation, and illustrate how the simulation procedure works at one location. In section 4, the primary contribution of this work, we then describe the proposed humidity simulation and illustrate the bivariate simulation at the same location. We provide an analysis of changes over CONUS, as well as summaries of our resulting observation-based simulations, in section 5. In section 6, we provide a discussion and concluding remarks. Additional technical details may be found in appendixes A and B, and statistical model validation may be found in the online supplemental material.

2. Data

The simulation procedure that we propose in sections 3 and 4 requires temperature and humidity data from both GCM output and observations. Unlike for temperature, many different and related variables are used as measures of humidity. In this paper, we use the *dewpoint*, which is the temperature to which an air parcel would need to be cooled in order for the water vapor in it to condense. We use dewpoint for several reasons, including that it is directly measured by the weather stations in the observational data we use (see below), is a reasonable indicator of human comfort (Davis et al. 2016), and can be understood straightforwardly as a measure of humidity because it does not change with temperature if the moisture content of the air is fixed (unlike relative humidity). At some points in this paper, we also refer to the *dewpoint depression*, which is defined as the difference between (dry bulb) temperature and dewpoint. While our simulations and results are presented in terms of dewpoint, the dry bulb temperature and dewpoint determine the relative humidity value, so a relative humidity simulation is implicit in the proposed procedure.

a. *Climate model data*

Climate model output for temperature and humidity are from CESM1-LE, a 40-member initial condition ensemble run with the fully coupled 1° latitude–longitude version of CESM1. Historical forcings (Lamarque et al. 2010) are used for the years 1920–2005 and the RCP8.5 scenario (Meinshausen et al. 2011) is used for the years 2005–2100. In this study, we focus on boreal summer (JJA) daily average temperature and dewpoint over a subregion of the global grid that contains CONUS (latitudes between 25° and 50°N, and longitudes between 66° and 125°W, and grid cells at least 50% land).

Daily or subdaily near-surface humidity data are only available (as reference height specific humidity, QREFHT) in CESM1-LE as 6-hourly data during the time periods 1990–2005, 2026–35, and 2071–80. To calculate daily average dewpoint from the available data, we first use the QREFHT and surface pressure (PS) values to calculate 6-hourly values of dewpoint using the equations found in Table 1 of Willett et al. (2014); we then average the four 6-hourly values to compute a daily value. To be consistent with this calculation, daily temperatures are computed by averaging the four concurrent 6-hourly reference height temperature (TREFHT) values. In cases where the calculated dewpoint exceeds the daily temperature value (0.03% of cases), it is set to the value equal to the temperature minus the minimum positive calculated dewpoint depression at the grid cell.

The statistical models we describe below for changes in the distribution of local temperature and dewpoint require an estimate of the forced trend in global mean temperature (GMT). An estimate from CESM1-LE is obtained by first averaging the annually averaged TREFHT values across grid cells (weighting by geographic area) and across the first 35 ensemble members (for the full 1920–2100 ensemble output). We exclude the last five runs, completed at the University of Toronto, in this calculation due to concerns about the consistency of the results when considering trends in the GMT (CESM Project 2020); however, they are included in the local temperature and dewpoint data described above, as we have found no evidence of a discrepancy in the distribution of these local variables over CONUS. The raw global annual-mean ensemble-mean temperature value is then further smoothed using a lowess smoother with a 5% span (approximately 9 years) and tricubic weighting to obtain our final estimate of the GMT forced trend. See Fig. S1 in the online supplemental material for the ensemble mean GMT anomalies and lowess smoothed estimated trend. While alternative methods for GMT trend estimation are available (e.g., Poppick et al. 2017), we use lowess smoothing here for simplicity since studying the global trend is not of primary interest in this work.

FIG. 1. Locations of GSOD stations used in our analysis.

b. *Observational data*

Observational temperature and dewpoint data are from the Global Summary of the Day (GSOD) database provided by the National Center for Environmental Information (NCEI). Data are typically collected on an hourly or 3-hourly basis and are averaged to daily values. We restrict our analysis to stations within CONUS and we use data from the years 1973–2018, where the record is most complete. We retain only stations that meet the relatively strict criterion that there are no missing temperature values during JJA over the time period studied; this corresponds to about 25% of CONUS stations in the database and appears to be geographically representative (see Fig. 1). The requirement of no missing temperature values makes the proposed temperature simulation (relying on fast Fourier transforms) more convenient; however, this requirement could be relaxed by using interpolation methods to fill in missing values or by using an alternative temperature simulation method that does not require complete data, since the proposed dewpoint simulation can be implemented with any temperature simulation method deemed suitable by the user.

3. Univariate simulation of temperatures

The novel observation-based dewpoint simulation that we propose in section 4 is conditional on a temperature simulation. Many methods exist for producing univariate simulations of temperature that combine information from GCM output with observations, as discussed in section 1. The procedure used here is a modified version of those proposed in Leeds et al. (2015) and Poppick et al. (2016), which are extensions of the so-called delta method. The delta method simulates future temperatures by adding a GCM projected future mean trend to historical observations; however, temperature variability changes are also potentially important and unaccounted for by the delta method. Variability changes can be dependent on time scale (e.g., day-to-day versus interannual variability changes can differ), implying full changes to the temporal covariance structure. The proposed procedure accounts for these temporal covariance structure changes by modifying the observed temperatures' spectral density to account for GCM projected changes.

Before describing the simulation method, we introduce some notation. Throughout, we write $\tilde{x}$ for a quantity in GCM output that is analogous to the quantity x in reality. We write $\mathbf{T}_y^{(h)} = (T_{y,1}, \ldots, T_{y,n_{\mathrm{JJA}}})$ for the vector of the $n_{\mathrm{JJA}} = 92$ observed JJA daily temperatures in the yth year of the observational record, and $\tilde{\mathbf{T}}_y^{(h)}$ and $\tilde{\mathbf{T}}_y^{(f)}$ for temperatures in historical and future portions of the GCM output (suppressing notation for the GCM ensemble member number for the sake of clarity).

We assume that temperatures can be separated into a mean forced response component and a residual component representing internal variability (also possibly changing in distribution in response to forcing):

$$\mathbf{T}_y^{(h)} = \boldsymbol{\mu}_y^{(h)} + \boldsymbol{\epsilon}_y^{(h)},$$

and similarly for $\tilde{\mathbf{T}}_y^{(h)}$ and $\tilde{\mathbf{T}}_y^{(f)}$, where $\boldsymbol{\mu}_y^{(h)}$ is the mean forced component in year y (not necessarily constant across days within year y) and $\boldsymbol{\epsilon}_y^{(h)}$ is the temperature deviation from that forced component. (In our terminology, the "mean forced component" can be thought of as the true average temperature over infinite realizations of internal variability, and would include both changes due to radiative forcing and the baseline average value.) Our temperature simulation procedure requires information about projected *changes* in both the mean forced component and internal temperature variability. The GCM change in mean forced component is denoted

$$\tilde{\Delta}_y = \tilde{\boldsymbol{\mu}}_y^{(f)} - \tilde{\boldsymbol{\mu}}_y^{(h)}. \tag{1}$$

We represent changes in the temporal covariance structure of the temperature deviations, ϵ_y, in terms of changes in spectral densities. Under the reasonable approximation that JJA internal temperature variability, captured by ϵ_y, is statistically stationary within a year, it has a spectral density that is denoted $a_y^{(h)}(\omega)$ for frequency ω, interpreted as the variance attributable to fluctuations at frequency ω. We write $\mathbf{a}_y^{(h)} = \{a_y^{(h)}(0/n_{\mathrm{JJA}}), a_y^{(h)}(1/n_{\mathrm{JJA}}), \ldots, a_y^{(h)}[(n_{\mathrm{JJA}}-1)/n_{\mathrm{JJA}}]\}$ for the vector of spectral densities on the Fourier frequencies [noting that the spectral density must be periodic with $a(\omega) = a(1-\omega)$ for $\omega \in (0, 1/2)$]. Likewise, we write $\tilde{\mathbf{a}}_y^{(h)}$ and $\tilde{\mathbf{a}}_y^{(f)}$, and we write

$$\tilde{\boldsymbol{\rho}}_y = \left\{\frac{\tilde{a}_y^{(f)}(0/n_{\mathrm{JJA}})}{\tilde{a}_y^{(h)}(0/n_{\mathrm{JJA}})}, \ldots, \frac{\tilde{a}_y^{(f)}[(n_{\mathrm{JJA}}-1)/n_{\mathrm{JJA}}]}{\tilde{a}_y^{(h)}[(n_{\mathrm{JJA}}-1)/n_{\mathrm{JJA}}]}\right\} \tag{2}$$

for the projected change in spectral density, expressed as a ratio of variances.

In section 3a, we first summarize the proposed procedure that uses the above quantities to produce an observation-based future simulation that captures projected changes in mean and temporal covariance. Because the true forced mean component and distribution of internal variability are not fully known in either the observations or GCM, these quantities must be estimated using statistical methods. In section 3b, we therefore describe the statistical models that are used to estimate the quantities required for the proposed procedure. We then illustrate the method at an example location in section 3c.

a. Simulation method

With the notation established above, our proposed simulation of daily JJA temperatures $\hat{\mathbf{T}}_y^{(f)}$ for the yth year of the future period is

$$\hat{\mathbf{T}}_y^{(f)} = \boldsymbol{\mu}_y^{(h)} + \tilde{\Delta}_y + \mathcal{F}^{-1}\mathrm{diag}(\sqrt{\tilde{\boldsymbol{\rho}}_y})\mathcal{F}[\mathbf{T}_y^{(h)} - \boldsymbol{\mu}_y^{(h)}], \tag{3}$$

where $\mathcal{F}$ is the discrete Fourier transform matrix [i.e., has entries $\mathcal{F}_{j,t} = e^{-2\pi i(j-1)(t-1)/n_{\mathrm{JJA}}}$]. (While written as matrix multiplication, in practice this is calculated using a fast Fourier transform algorithm.) In words, starting with observed JJA temperatures from the yth year of the historical record, we

1) subtract the estimated mean forced component from the observed temperatures to obtain the observed temperature deviation from the mean (herein simply the "temperature deviation"),
2) calculate the discrete Fourier transform of the observed temperature deviations,
3) multiply by frequency-dependent projected square root spectral density changes,
4) invert the Fourier transformation to obtain future temperature deviations reflecting GCM projected variability changes, and
5) add back the estimated forced response from the observations and the GCM projected change in mean to obtain the future temperature simulation.

This procedure produces a simulation that has the mean changes and (approximately) the covariance changes projected by the GCM [Eqs. (1) and (2)] but otherwise retains many features of the observed temperatures [including, e.g., spatial coherences; see Poppick et al. (2016)]. If $\tilde{\boldsymbol{\rho}}_y = 1$ (i.e., there are no projected variability changes), then the procedure reduces to the basic delta method. Other versions of the delta method extend the basic method to allow for *marginal* variability changes (i.e., same variability change at each frequency); step 3 is an extension of this idea, where instead we are separately multiplying by the change in standard deviation associated with fluctuations at each of the Fourier frequencies, producing changes to the full temporal covariance structure. As written, the simulation procedure is applied separately to temperatures in each year y, but in practice this can be done simultaneously in a vectorized fashion and the resulting simulation procedure can be implemented very quickly.

b. Statistical models for mean and variability changes

The quantities $\boldsymbol{\mu}_y^{(h)}$, $\tilde{\Delta}_y$, and $\tilde{\boldsymbol{\rho}}_y$ in Eq. (3) are unknown and must be estimated from the available data using statistical methods.

One approach to estimation, although not ours, would be to rely on methods that attempt to avoid strong assumptions about the functional form of the above quantities; as a simple example, one might estimate $\tilde{\Delta}_y$ as the difference in ensemble mean local temperatures between the two time points. However, this approach has a number of disadvantages. First, the ensemble mean local temperature (for example) is a very noisy estimate of the true model forced component, due to the presence of considerable internal variability in local temperatures and an ensemble of only 40 runs. This problem is even more pronounced when estimating more complex quantities (such as the changes in spectral densities $\tilde{\boldsymbol{\rho}}_y$) or when estimating $\boldsymbol{\mu}_y^{(h)}$ from observations (where we have only a single observational record). Moreover, these kinds of methods do not readily provide an interpretable summary of how the relevant functions are changing across years within the forcing scenario of interest.

Instead, our approach is based on the fact that some aspects of the forced response of local atmospheric variables scale approximately with the GMT forced response [e.g., Santer et al. (1990), Dai et al. (2015), and Li et al. (2020), among many others; see also IPCC (2013, section 12.4.2)]. We use this idea to constrain the functional forms of the changes. The resulting parametric statistical models make stronger assumptions but give estimates that are both less noisy and easier to interpret. Validation of the statistical models presented below, and comparison with empirical estimates obtained only by averaging across ensemble members, may be found in the online supplemental material in sections S2.1 and S2.2.

We model the mean forced component of local temperature on the dth day of the yth year in the historical observational record as

$$\mu_{y,d}^{(h)} = \beta_0 + \beta_1 \tilde{G}_y^{(h)} + \gamma(d) + \eta(d)\tilde{G}_y^{(h)}, \qquad (4)$$

where $\tilde{G}_y^{(h)}$ is the smoothed GMT anomaly value in year y of the historical period from the GCM, the β terms are unknown parameters, and $\gamma(d)$ and $\eta(d)$ are unknown functions. We use $\tilde{G}_y^{(h)}$ as our estimate of the GMT forced response in the historical period *for the observations* under the assumption that the GCM adequately captures the forced response of GMT. The function $\gamma(d)$ allows for seasonally varying mean temperatures and $\eta(d)$ allows for that seasonal cycle to change as the GMT changes; these functions are both parameterized using the first two seasonal harmonics. Since $\tilde{G}_y$ is smooth in time, the resulting estimates of forced mean changes in local temperature also vary smoothly in time, with additional flexibility to allow for seasonality in the mean changes. The GCM mean forced component is modeled similarly, so that the projected change in mean temperature on the dth day of the yth year (comparing future versus historical time periods) is

$$\tilde{\Delta}_{y,d} \equiv \tilde{\mu}_{y,d}^{(f)} - \tilde{\mu}_{y,d}^{(h)} = [\tilde{\beta}_1 + \tilde{\eta}(d)](\tilde{G}_y^{(f)} - \tilde{G}_y^{(h)}). \qquad (5)$$

Modeling local mean changes as proportional to global mean changes results in well-fitting mean functions for the time periods, region, and climate change scenario analyzed here. More complex local mean emulators (e.g., Castruccio et al. 2014) could be used if more a general simulation (e.g., for multiple climate change scenarios) were required.

Similarly, we model the changes in spectral densities at frequency ω as

$$\log[\tilde{\rho}_y(\omega)] = \tilde{\delta}(\omega)[\tilde{G}_y^{(f)} - \tilde{G}_y^{(h)}], \qquad (6)$$

where $\tilde{\delta}(\omega)$ is a smooth function in frequency (smoothing enforced via a kernel smoother). If $\tilde{\delta}(\omega)$ is constant, then model (6) says that only the marginal standard deviation of temperatures changes over time (and scales with GMT); otherwise, the model allows different changes at different frequencies and therefore changes to the full temporal covariance structure. Positive values of $\delta(\omega)$ correspond to frequencies with variability increasing with increasing GMT, and similarly for negative values of $\delta(\omega)$. See the sections a and b in appendix A for additional modeling and estimation details for models (4) and (6), respectively.

c. *Illustration of method*

Here we show an illustration of the procedure described above for a location near Minneapolis, Minnesota (MN); the observational data we use are from a station located at Minneapolis–Saint Paul International Airport, and the changes from CESM1-LE are estimated using the nearest grid cell. For an analysis over all of CONUS, see section 5. For illustration to directly compare the observation-based simulation with raw output from CESM1-LE, the years 1996–2005 of the observations are used to simulate the years 2071–80 of the RCP8.5 scenario, although the full data are used to estimate the required parameters for the simulation. For reference, the average change in the GMT forced mean between these two time periods in CESM1-LE is approximately 3.3°C.

First, we show the CESM1-LE projected changes in mean and variability at this grid cell. Figure 2a shows changes in local mean temperature per degree warming of GMT. Specifically, we show the function $\tilde{\beta}_1 + \tilde{\eta}(d)$ versus d from Eq. (5) along with approximate 95% pointwise confidence intervals; see appendix A, section a for a description of how the confidence intervals are obtained. Averaged over the summer, this grid cell shows warming of about 1.6°C per degree warming in GMT, but the later part of summer warms most quickly. Figure 2b shows the relative change in the square root spectral density at this grid cell per degree warming of GMT—that is, we show the function $\exp[\tilde{\delta}(\omega)/2] - 1$, again along with approximate 95% pointwise confidence intervals; see appendix A, section b—which can be interpreted as the relative change in standard deviation for fluctuations with frequency ω. In Fig. 2b, the value at $\omega = 0$ is labeled "JJA annual average"; this is because the discrete Fourier transform at $\omega = 0$ is proportional to the time average, so that $\sqrt{\tilde{\rho}_y(0)}$ is the change to the standard deviation of the JJA annual average in year y in Eq. (3). Summertime lower-frequency variability at this grid cell increases with GMT (e.g., the standard deviation of the JJA average temperature is estimated to increase by about 6% per degree increase in GMT), but the variability changes at higher frequencies are estimated to be negligible.

These changes in means and variability are then incorporated into the simulation procedure [Eq. (3)] to produce an observation-based simulation of future temperatures. Figures 2c–f show an illustration of observed temperatures and the resulting simulation. We show observed and simulated daily values over the first year of the record (Figs. 2c,d), as well as JJA average values across years (Figs. 2e,f), to illustrate the differences in variability changes at short and longer time scales, respectively. The observed temperatures and simulated series are shown on top (Figs. 2c,e), while deviations from the forced mean (i.e., $T_{y,d}^{(h)} - \mu_{y,d}^{(h)}$ and similar) are shown at bottom (Figs. 2d,f) to illustrate how the simulation procedure affects temperature variability. Because high-frequency variability

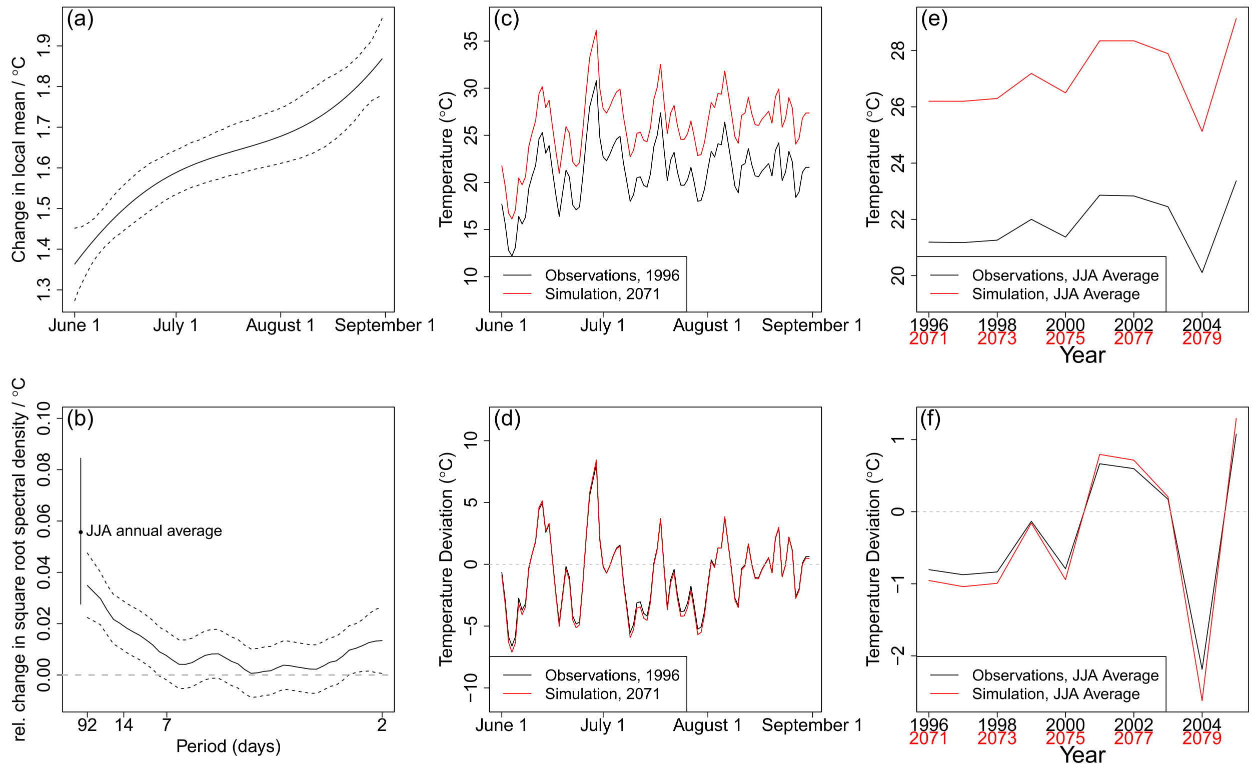

FIG. 2. (a) CESM1-LE projected change in mean temperature for a 1° increase in GMT (with approximate 95% confidence interval; dashed), by day of the summer, for a grid cell near Minneapolis, MN. The summertime mean temperatures at this location increase more quickly than GMT, and the increase is greater toward the end rather than the beginning of the summer. (b) Estimated relative change in the square root spectral density for a 1° increase in GMT at the same grid cell (with approximate 95% confidence interval; dashed). Variability increases at low frequencies, but changes are negligible for higher frequencies. The remaining panels show an illustration of the proposed temperature simulation method using observations from Minneapolis–Saint Paul International Airport. Observations are from the years 1996–2005, and the simulation is for the years 2071–80 of the RCP8.5 scenario. (c),(e) Temperatures; (d),(f) temperature deviations (i.e., temperature minus its estimated forced mean); first year of daily values in observations (black) and resulting simulation (red) are shown in (c) and (d); JJA average values over the observed record and simulation period are shown in (e) and (f).

changes at this location are projected to be negligible, simulated daily temperature deviations look very similar to the observed values in the first year and the resulting simulation is driven by mean changes; however, lower-frequency variability is projected to increase at this location, resulting in JJA average values that are more variable in the simulation than in the original observations.

4. Observation-based conditional quantile mapping of dewpoint given temperature

In this section, we propose an observation-based simulation method for dewpoint that depends on a temperature simulation such as that proposed in section 3. The proposed dewpoint simulation falls into the general category of ''quantile mapping'' simulation methods, and works by transforming observed dewpoint quantile levels given observed temperatures to new dewpoints given simulated temperatures, accounting for GCM projected changes in the underlying relationship between dewpoint and temperature.

As in section 3, we first summarize the proposed simulation procedure in section 4a, then describe the statistical models that are used to estimate the quantities required for the proposed procedure in section 4b, and finally provide an illustration of the simulation at one location in section 4c.

a. Simulation method

Conceptually, we want to simulate future dewpoint values, given a future temperature simulation, via a quantile mapping that transforms the observed distribution into a future distribution. Writing $\mathbf{D}^{(h)}$ for the observed dewpoint values and $\hat{\mathbf{D}}^{(f)}$ for the observation-based future simulation [and $\mathbf{T}^{(h)}$, $\hat{\mathbf{T}}^{(f)}$, $\tilde{\mathbf{G}}^{(h)}$, and $\tilde{\mathbf{G}}^{(f)}$ for the observed local temperature and future simulation, and historical and future smoothed GMT forced trend estimate, as above], our *ideal* simulation procedure would entail the following transformation:

$$\hat{\mathbf{D}}^{(f)} = F^{-1}_{\mathbf{D}^{(f)}}[\boldsymbol{\tau}^*|\hat{\mathbf{T}}^{(f)}, \tilde{\mathbf{G}}^{(f)}], \quad \text{where}$$

$$\boldsymbol{\tau}^* = F_{\mathbf{D}^{(h)}}[\mathbf{D}^{(h)}|\mathbf{T}^{(h)}, \tilde{\mathbf{G}}^{(h)}] \tag{7}$$

and where $F_X(\cdot|\cdot)$ denotes the (conditional) cumulative distribution function of the variable X and therefore $F_X^{-1}(\cdot|\cdot)$ is the conditional quantile function. In words, the ideal procedure would be the following:

1) determine the quantile levels associated with observed dewpoint values, given observed local temperatures and the GMT forced trend, and then
2) simulate future dewpoints using those quantile levels, but conditionally on future simulated local and global temperatures and under the projected future relationship between dewpoint and temperature.

However, in practice we do not know $F^{-1}_{\mathbf{D}^{(f)}}(\cdot|\cdot)$, the conditional quantile function of true future dewpoints, so our proposed method will replace this function with one that incorporates changes from the GCM, analogous to other observation-based simulation procedures.

b. Statistical model for conditional quantile functions

We model the relationship between dewpoint and temperature (and changes thereof) through a quantile regression model of the log dewpoint depression, with effects for GMT, seasonality, and the local temperature deviation from its estimated forced mean. Writing $\tilde{Y}_{y,d} = \log(\tilde{T}_{d,y} - \tilde{D}_{d,y})$ for the log dewpoint depression in the GCM (again suppressing notation for the run, and here not distinguishing between historical versus future time periods), our quantile regression model for the GCM is

$$F^{-1}_{\tilde{Y}_{d,y}}(\tau|\tilde{T}_{d,y}, \tilde{G}_y) = \tilde{\alpha}_{0,\tau} + \tilde{g}_\tau(d) + \tilde{\alpha}_{1,\tau}\tilde{G}_y + \tilde{h}_\tau(\tilde{T}_{d,y} - \tilde{\mu}_{d,y}), \quad (8)$$

where the function $\tilde{g}_\tau(\cdot)$ characterizes the seasonal cycle in the τth quantile for a fixed GMT and local temperature deviation, $\tilde{\alpha}_{1,\tau}$ characterizes the change in dewpoint depression with GMT for a fixed local temperature deviation and day of the year, and the function $\tilde{h}_\tau(\cdot)$ characterizes the relationship between the dewpoint depression and local temperature deviation for a fixed day of year and GMT. We use the two leading seasonal harmonics to parameterize $\tilde{g}_\tau(\cdot)$ and use a natural spline to parameterize $\tilde{h}_\tau(\cdot)$; see appendix A, section c for details. Working on the scale of log dewpoint depression allows us to easily enforce the constraint that the dewpoint depression must be positive, and therefore that dewpoint must be less than the temperature value. However, model (8) implies an equivalent model on the scale of the dewpoint itself, $\tilde{D}_{y,d} = \tilde{T}_{d,y} - e^{\tilde{Y}_{d,y}}$, that is,

$$F^{-1}_{\tilde{D}_{d,y}}(\tau|\tilde{T}_{d,y}, \tilde{G}_y) = \tilde{T}_{d,y} - \exp[\tilde{\alpha}_{0,1-\tau} + \tilde{g}_{1-\tau}(d) + \tilde{\alpha}_{1,1-\tau}\tilde{G}_y + \tilde{h}_{1-\tau}(\tilde{T}_{d,y} - \tilde{\mu}_{d,y})]. \quad (9)$$

So while in practice the model is estimated on the scale of Eq. (8) (where the model is linear), changes in the log dewpoint depression can be translated into changes in dewpoint (but the τth quantile of dewpoint depression corresponds to the $1-\tau$th quantile of dewpoint). See supplemental material section S2.3 for information about the quality of fit of this quantile regression model in CESM1-LE.

In model (8), changes in the conditional quantile function associated with changes in GMT are captured by the $\tilde{\alpha}_{1,\tau}$ parameters at each quantile level τ. In the spirit of univariate observation-based simulation procedures, we assume that the GCM captures changes with GMT through this parameter, but that the terms describing the seasonal cycle and dependence on local temperature in observations may be different from those in the GCM. As such, we take the future quantile function for our simulation to be

$$F^{-1}_{\hat{D}^{(f)}_{d,y}}[\tau|\hat{T}^{(f)}_{d,y}, \tilde{G}^{(f)}_y] = \hat{T}^{(f)}_{d,y} - \exp\{\alpha_{0,1-\tau} + g_{1-\tau}(d) + \tilde{\alpha}_{1,1-\tau}\tilde{G}^{(f)}_y + h_{1-\tau}[\hat{T}^{(f)}_{d,y} - \hat{\mu}^{(f)}_{d,y}]\}.$$

This assumption would imply that the ideal simulation procedure (7) may be rewritten as

$$\hat{D}^{(f)}_{y,d} = \hat{T}_{d,y} - \exp\{\alpha_{0,1-\tau^*_{y,d}} + g_{1-\tau^*_{y,d}}(d) + \tilde{\alpha}_{1,1-\tau^*_{y,d}}\tilde{G}^{(f)}_y + h_{1-\tau^*_{y,d}}[\hat{T}^{(f)}_{d,y} - \hat{\mu}^{(f)}_{d,y}]\},$$

where

$$\tau^*_{y,d} = F_{D^{(h)}_{y,d}}[D^{(h)}_{y,d}|T^{(h)}_{y,d}, \tilde{G}^{(h)}_y], \quad (10)$$

which is therefore our proposed procedure. Note that we estimate the quantile function for the observations over the historical period using the $\tilde{\alpha}_{1,\tau}$ values estimated from the GCM; this is consistent with the overall simulation approach and the idea that the GCM will have much more information about this term than is contained in the observational output, because of both the larger amount of data in the GCM as well as the larger changes in GMT.

To better understand the procedure (10), it is instructive to consider first its behavior in the special case that $\tilde{\alpha}_{1,\tau} = 0$ for all τ (i.e., the relationship between dewpoint depression and the temperature deviation does not change with GMT) and $\tilde{\rho}(\omega) = 1$ for all ω (i.e., temperature variability does not change). In this setting, it is straightforward to show that the simulation resulting from (10) is $\hat{D}^{(f)}_{y,d} = D^{(h)}_{y,d} + \tilde{\Delta}_{y,d}$, where $\tilde{\Delta}_{y,d}$ is the projected change in local mean temperature [Eq. (5)]. That is, in this simplest setting, the future dewpoint value ''follows'' the mean temperature change, and also no modeling of the observed relationship between temperature and dewpoint is explicitly needed for the simulation.

In the general case, the procedure (10) can therefore be understood as an adjustment to that simple setting to account for changes with GMT and in the simulated temperature deviation value. The reason that an explicit quantile regression model for *observed* dewpoints is now needed is because the future simulated temperature deviation differs from the observed deviation value if $\tilde{\rho}(\omega) \neq 1$, therefore requiring an estimate of the observed relationship $h_\tau(\cdot)$ to carry out the simulation. This implies that, if projected temperature variability changes are small, the dewpoint simulation will not be very sensitive to errors in estimating $h_\tau(\cdot)$ from observations, and primarily depends on the estimated changes from the GCM. This is important because there is more information available from the GCM about changes than there is

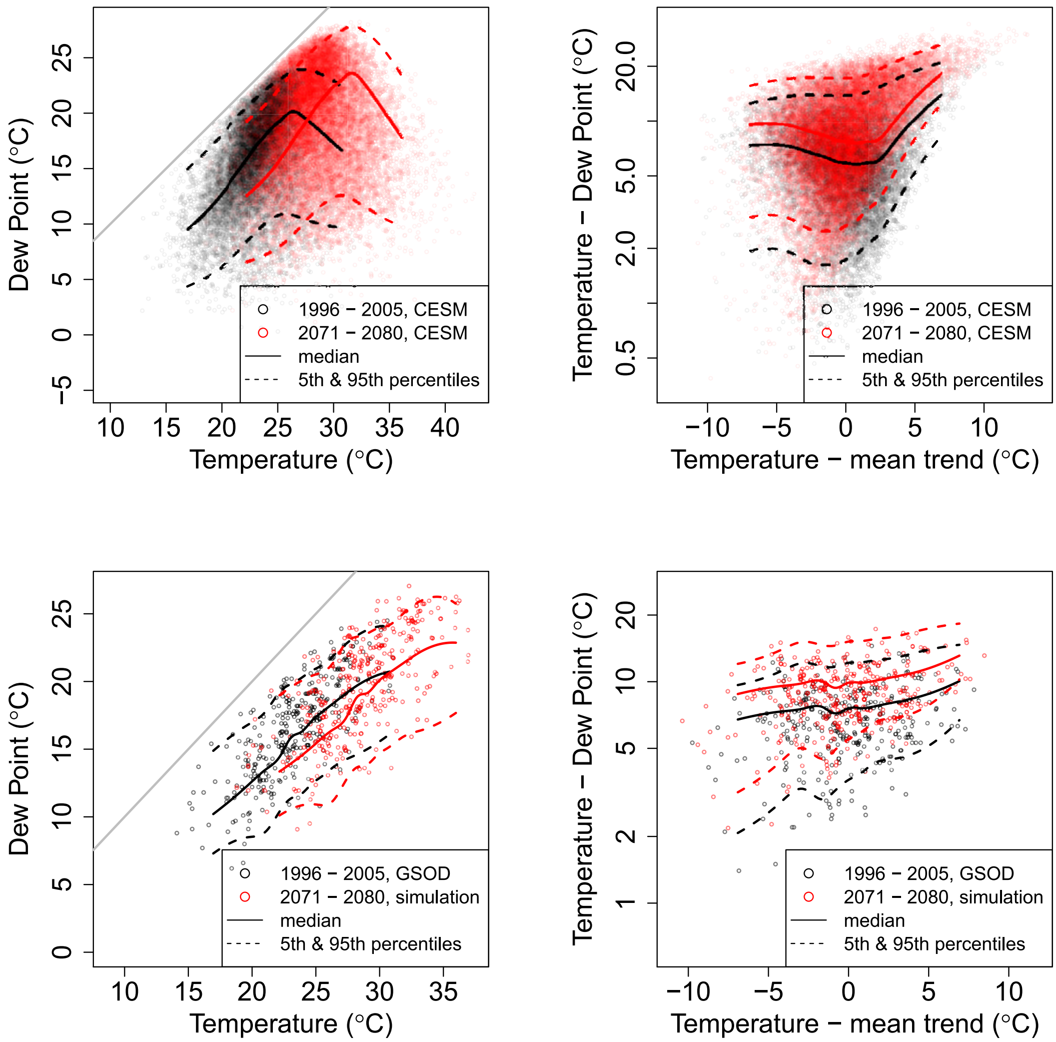

FIG. 3. Comparison of historical (black) vs future (red) bivariate distributions in CESM1-LE and observations vs simulated values during the time periods 1996–2005 and 2071–80 for the month of July. (top) CESM1-LE; (bottom) observations and simulations; (left) dewpoint vs temperature; (right) dewpoint depression (i.e., temperature minus dewpoint) vs temperature deviation (i.e., temperature minus its estimated forced mean). Lines show the fitted median and 5th/95th percentile curves on 16 July when GMT is equal to its average value in the relevant time period.

information in the observed record about the baseline distribution. Thus the simulation procedure (10) shares the important feature with other observation-based simulation methods that it is not as dependent on explicitly modeling the behavior of observations as it is on modeling changes in the GCM output.

c. Illustration of method

Figure 3 shows the bivariate distributions of temperature and dewpoint in CESM1-LE, as well as in the observations and observation-based simulations, for the same location near Minneapolis, MN, used in Fig. 2. We show observed and simulated values for the years 1996–2006 and 2071–80 during the month of July, along with the fitted median and 5th/95th percentile curves on 16 July when the GMT is equal to the average value in the respective time interval. (The displayed quantile curves are those for a fixed day of year, so are not strictly comparable to the data values shown, which vary in distribution over the month and across years; however, since this variation is relatively small compared to the differences between the future and historical periods, we still find it helpful to display for reference.)

Focusing first on CESM1-LE, the bivariate distributions are complex, as are the projected changes. The estimated $\tilde{\alpha}_{1,\tau}$ values are positive (Fig. 3, top right), and larger in magnitude for the smaller dewpoint depression quantiles (or equivalently, the larger dewpoint quantiles); for the three displayed quantile curves, the estimated values are $\tilde{\alpha}_{1,0.05} = 0.13$°C °C^{-1}, $\tilde{\alpha}_{1,0.05} = 0.081$°C °C^{-1}, and $\tilde{\alpha}_{1,0.95} = 0.067$°C °C^{-1}. This means that for a fixed temperature *deviation* [i.e., $\tilde{T}_{d,y} - \tilde{\mu}_{d,y}$, the covariate in model (8)], the dewpoint depression tends to be larger in the future time period than in the historical period (Fig. 3, top

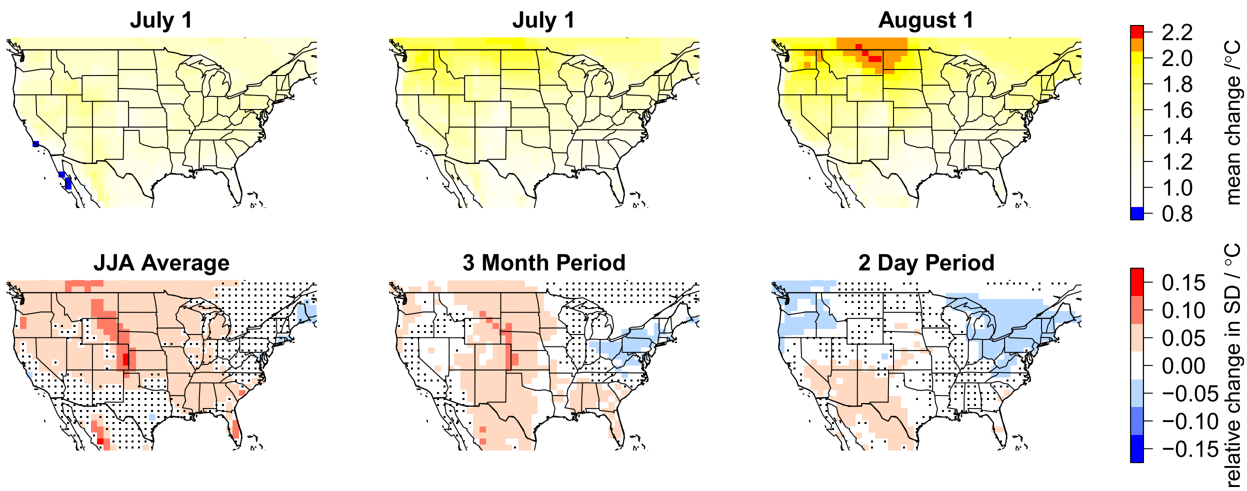

FIG. 4. (top) Projected mean temperature changes per degree increase in GMT on 1 Jun, 1 Jul, and 1 Aug [i.e., $\tilde{\beta}_1 + \tilde{\eta}(d)$ from Eq. (5), for d corresponding to the dates in question]. Projected mean changes are stronger toward the end rather than the beginning of summer, as also illustrated for the example grid cell in Fig. 2a. (bottom) Projected changes in temperature spectral densities per degree increase in GMT, for three different frequencies. We are displaying the values $\exp[\tilde{\delta}(\omega)/2] - 1$, where $\tilde{\delta}(\omega)$ is as in Eq. (6); this can be interpreted as the relative change, per degree increase in GMT, in the standard deviation attributable to fluctuations at the frequency ω. The left plot is for $\omega = 0$, which corresponds to the change in the standard deviation of the JJA average temperature; the center and right plots correspond to frequencies $\omega = 1/92$ and $\omega = 1/2$, respectively. Grid cells are stippled if the estimated change is not significantly different from zero, controlling the false discovery rate at 5% (see appendix A for details).

right), especially on days that were expected to be more humid; that is, for a fixed realization of temperature variability around the forced mean, we see relative drying especially on the most humid days. However, since the mean temperature increases from the historical to the future time periods, a fixed temperature deviation corresponds to a higher temperature value; for example, we still see combined heat and humidity events that are more severe in the future than in the historical period, albeit less so than if the $\tilde{\alpha}_{1,\tau}$ values were negative (Fig. 3, top left).

When we compare the bivariate distributions in CESM1-LE to those in the observations and our observation-based simulation (Fig. 3, bottom), we see clear discrepancies between the two. For example, dewpoints show less variability in the observations for a fixed temperature than is apparent from CESM1-LE, and the estimated median and 5th/95th percentile curves are nearly monotonic in the observations whereas not in CESM1-LE. The simulation procedure (10) transfers the projected changes with GMT to the observed bivariate relationship, generating a simulation whose changes look similar to those in CESM but that is consistent with the original observations' distribution (and better preserves higher-order features from the observations, like spatiotemporal relationships). Because of this difference in the underlying relationship between dewpoint and temperature, the aforementioned increase in future risk of high humidity at a historically high temperature value is smaller in the observation-based simulation compared to in CESM1-LE.

5. Results

The preceding sections describe our simulation procedure and illustrate both the projected changes from CESM1-LE and resulting simulations at one location near Minneapolis, MN. Here we provide a summary of projected changes from CESM1-LE over a region containing CONUS, and discuss the resulting simulations using GSOD station data. Simulations are produced using projected changes from the grid cell that is nearest to the GSOD station in question and at least 50% land, and the years 1996–2005 are used to produce a simulation of the years 2071–80 in the RCP8.5 scenario in order to directly compare the observation-based simulation with raw output from CESM1-LE, although the full data is used to estimate the required parameters for the simulation.

Recall that the temperature simulations require estimates of both local mean changes and changes in spectral densities. Figure 4 (top) shows projected local mean temperature changes per degree warming in GMT, at three time points throughout the summer. That is, we show $\tilde{\beta}_1 + \tilde{\eta}(d)$ from Eq. (5), for d corresponding to the dates in question. Across CONUS, summertime local temperatures in CESM increase more quickly than in the GMT, and the increases are stronger toward the end rather than the beginning of the summer (especially in the Northwest). That is, the changes over CONUS are broadly similar to those shown in Fig. 2a at the grid cell near Minneapolis.

Figure 4 (bottom) shows projected changes in local temperature spectral densities per degree increase in GMT, at three different frequencies. As in Fig. 2b, we show the values $\exp[\tilde{\delta}(\omega)/2] - 1$, where $\tilde{\delta}(\omega)$ is as in Eq. (6); this corresponds to the relative change, per degree increase in GMT, in the standard deviation attributable to fluctuations at frequency ω. Likewise as in Fig. 2b, the changes at $\omega = 0$ are labeled "JJA average" because these can be interpreted as changes in the standard deviation of the JJA average temperature in the resulting simulation. Summertime low-frequency temperature

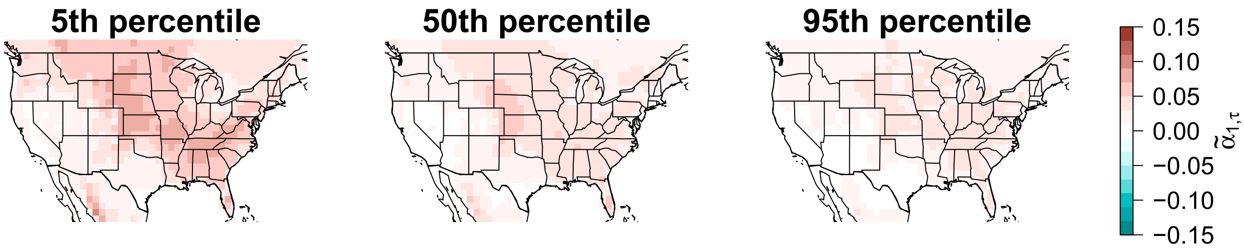

FIG. 5. Projected changes in log dewpoint depression quantiles per degree increase in GMT, fixing the temperature deviation from the forced trend [i.e., $\tilde{\alpha}_{1,\tau}$ from Eq. (8)]. Recall that the 5th percentile of dewpoint depression corresponds to the 95th percentile of dewpoint, and vice versa. Over CONUS, summertime dewpoint depression increases with GMT, fixing the local temperature deviation, and more so on more humid days.

variability is largely projected to increase with increasing GMT, although changes in the Northeast and Southwest are estimated as negligible. Higher-frequency variability changes are more spatially variable, with decreases in the Northeast and Northwest and increases in the Southwest and Plains states.

After estimating changes in temperature distributions for our temperature simulation, the dewpoint simulation requires information about changes in the conditional quantile function of dewpoint. Figure 5 shows the estimated values of $\tilde{\alpha}_{1,\tau}$ from Eq. (8) for three quantile levels; recall that these are the changes in log dewpoint depression quantiles per increase in GMT, fixing the day of the year and the temperature deviation from its forced mean, and that the τth quantile for dewpoint depression corresponds to the $1-\tau$th quantile for dewpoint. As in Fig. 3, top right, we see increases in dewpoint depression with increasing GMT, particularly in the lower quantiles. That is, for a fixed temperature deviation, we see relative drying particularly on the most humid days. The increases are smaller in the West and Southwest.

The observation-based simulations inherit changes in dewpoint depression for a fixed temperature *deviation* projected from the GCM (i.e., the parameter $\tilde{\alpha}_{1,\tau}$), but the resulting changes in dewpoint for a fixed temperature *value* can differ between the GCM and observation-based simulation because the same temperature value corresponds to different temperature deviations in Eqs. (9) versus (10) and involve different relationships in the GCM and observations, $\tilde{h}(\cdot)$ versus $h(\cdot)$. Here we consider changes in the risk of high or low dewpoint values for a fixed temperature value. Figure 6 shows changes in the risk of historically high humidity events at a fixed historically high temperature, for CESM1-LE and the observation-based simulation. We express changes in risk in terms of odds ratios, comparing the future (2071–80) odds of exceeding the historical (1990–2005) 95th percentile of dewpoint, conditional on local temperature being at the historical 95th percentile, to the historical odds. Overall, the probability of historically high humidity on a historically warm day increases in both CESM1-LE and the observation-based simulation, particularly in the Southeast. However, the increase in risk is typically larger in CESM1-LE than in the observation-based simulation. This result may stand in apparent contrast with Fig. 5, which shows dewpoint depression increases (and so dewpoint decreases) for a fixed temperature deviation in the future; however, a historically high temperature is associated with a smaller (or possibly negative) temperature deviation in the future, explaining the difference (e.g., compare the left and right panels of Fig. 3 for a fixed x-axis value). See appendix B for more details on how the quantities we are showing here are calculated.

Figure 7 is analogous to Fig. 6 but instead shows changes in the risk of historically *low*-humidity events at a fixed historically high temperature. Here we compare the 2071–80 odds of dewpoint less than the 1990–2005 5th percentile of dewpoint, fixing local temperature at its 1990–2005 95th percentile, to the 1990–2005 odds. Unlike for humid heat events, the risk of dry heat events appears to decrease over CONUS except in the Northeast (particularly in later summer). However, the decrease is stronger in CESM1-LE than in the observation-based simulation, indicating that future low-humidity heat events may be a greater risk than indicated by CESM1-LE.

6. Discussion

In this work, we propose an observation-based joint simulation of future dewpoint and temperature that accounts for estimated changes from CESM1-LE in 1) mean temperature, 2) temperature variability at multiple time scales, and 3) changes in the relationship between dewpoint and temperature. We believe that the proposed simulation method is preferable compared to those based on attempting to statistically correct GCM runs, both because observation-based simulation procedures retain higher-order distributional features of the observations and because they require less explicit statistical modeling of observations (where data are limited) than of GCMs (where more data are available).

That said, observation-based simulations are of course limited by the observational record. The simulation procedure as proposed is nonstochastic and is determined by the observational data and the estimated changes from the GCM; however, it may be possible to produce longer or multiple simulations by using a resampling procedure on the observations or through developing a statistical model of the variability [as in, e.g., McKinnon et al. (2017) and McKinnon and Deser (2018)]. In situ station measurements are also spatially limited, can contain systematic errors [e.g., issues discussed in Brown and DeGaetano (2009), Durre et al. (2010), Dunn et al. (2014),

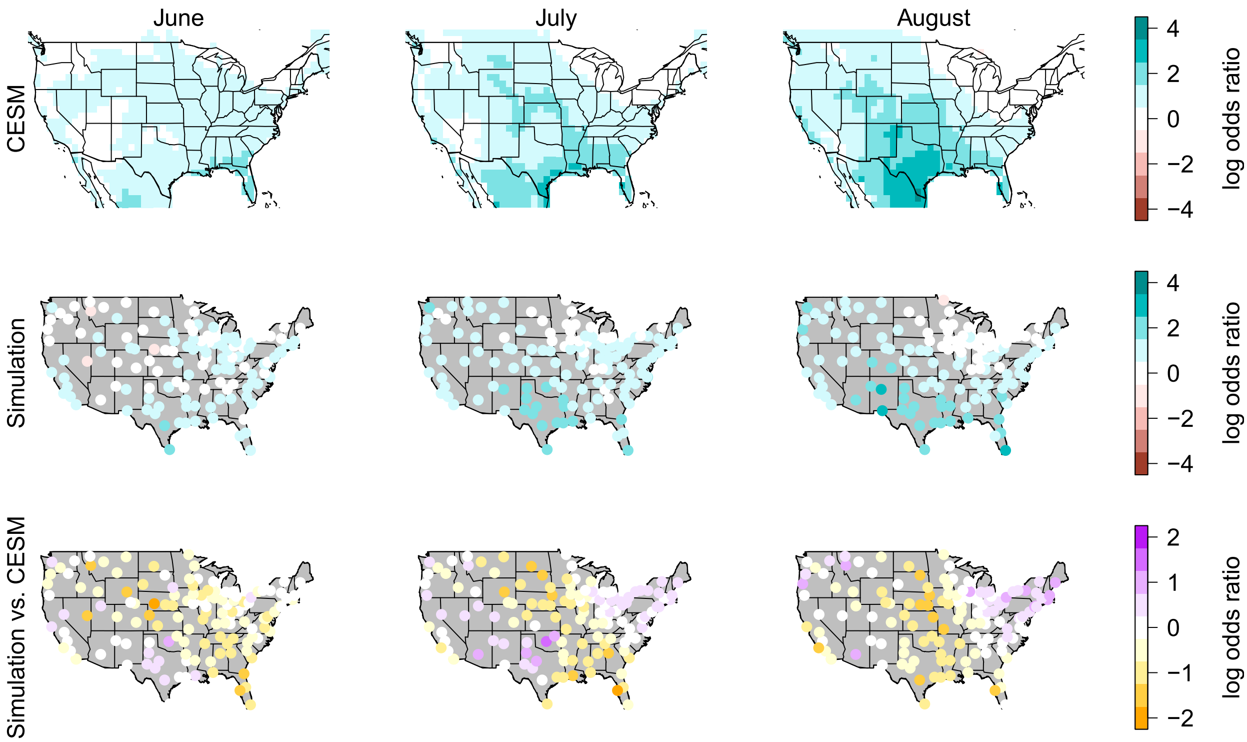

FIG. 6. Change in risk of historically high (95th percentile) dewpoint on historically high temperature (95th percentile) days, by month, comparing the years 2071–80 vs 1990–2005 in (top) CESM1-LE and (middle) the observation-based simulation, and (bottom) the difference between the two. The top two rows show the log (base 10) odds ratio comparing the future vs historical risk; positive values correspond to an increase in risk. The bottom row shows the difference between the middle and top rows; positive values indicate a larger future risk in the observation-based simulation than in CESM1-LE. See appendix B for details.

Rhines et al. (2015), and many others] and are point measurements that cannot be straightforwardly compared to a GCM grid cell, presenting potential challenges for any simulation method that requires information from observations (including model-based simulation procedures). We do not address these issues here, but we believe that it is advisable to repeat the simulation procedure with multiple data sources (if available) to address the effects of observational uncertainties.

The analysis presented here uses one GCM ensemble, CESM1-LE. While outside the scope of the current work, it would be important when applying these methods using changes from other climate model output to verify that the underlying statistical models remain appropriate for the climate model output in question (as e.g., we do in section S2 of the supplemental material). Multimodel comparisons are advisable if producing simulations for impacts assessments in order to account for intermodel variability.

Simulation procedures that combine information from GCM output with observations inherently require users to choose either what changes from the GCM should be reflected in the simulation (for observation-based procedures) or what features of the GCM output should be corrected (for model-based procedures). These notions become arguably more ambiguous in more complex settings (e.g., for multivariate simulations and where the simulation involves nonlinear transformations). The underlying assumption that the GCM captures relevant changes also becomes more challenging to evaluate in the presence of limited observational data if the changes are complex but small over the historical period, as in our setting. While this ambiguity is difficult to fully overcome, we believe that one additional advantage of the approach taken here, which involves parametric statistical models to characterize GCM projected changes, is that such models make transparent what changes are inherited in the resulting simulation and these estimated changes are on their own a relevant summary of the GCM's behavior.

Our understanding of changes in compound extreme events, such as humid or dry heat events, remains limited. We hope that the statistical methods developed in this paper aid in the development of coherent and interpretable comparisons across other regions, forcing scenarios, and climate models, to ultimately enrich our understanding of these complex but important projected changes.

Acknowledgments. AP received support from the Hewlett Mellon Fellowship from Carleton College. Code to reproduce the analyses may be found at https://github.com/apoppick/temp_dewp_simulations.

Data availability statement. The CESM1-LE data are available for download from the Climate Data Gateway at NCAR

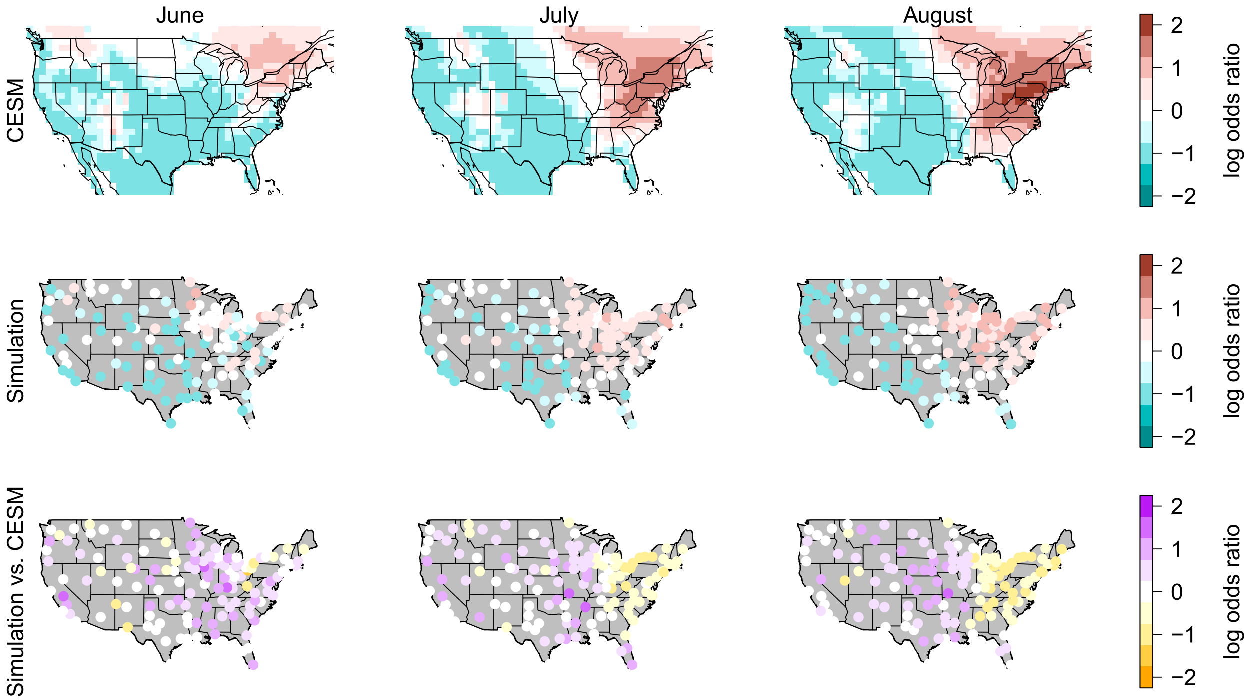

FIG. 7. Change in risk of historically low (5th percentile) dewpoint on historically high temperature (95th percentile) days, by month, comparing the years 2071–80 vs 1990–2005 in (top) CESM1-LE and (middle) the observation-based simulation, and (bottom) the difference between the two. The top two rows show the log (base 10) odds ratio comparing the future vs historical risk; positive values correspond to an increase in risk. The bottom row shows the difference between the middle and top rows; positive values indicate a larger future risk in the observation-based simulation than in CESM1-LE. See appendix B for details. Compare to Fig. 6 but note the reversal of the color bars to follow the intuition that red corresponds to drying.

(formally known as the Earth System Grid) (https://doi.org/10.5065/d6j101d1). The GSOD weather station data are available at https://www1.ncdc.noaa.gov/pub/data/gsod/.

APPENDIX A

Details on Statistical Models

In this appendix, we provide more detail on the statistical models used in this paper and how the relevant parameters are estimated.

a. Mean temperature model

The functions $\gamma(d)$ and $\eta(d)$ in the mean temperature model (4) control seasonality and changes thereof with GMT, respectively. These functions are parameterized using the first two seasonal harmonics, that is,

$$\gamma(d) = \sum_{k=1}^{2} [\gamma_{k,1}\cos(2\pi kd/365) + \gamma_{k,2}\sin(2\pi kd/365)], \quad \text{(A1)}$$

and similarly for $\eta(d)$. We have found that using only the first two harmonics is sufficient for modeling seasonality within JJA (see supplemental material section S2.1 for model validation) since most seasonal structure is dealt with by only examining JJA temperatures.

The parameters in the mean model (4) are estimated via least squares (separately for CESM1-LE and GSOD, but using all data within each respective dataset). Uncertainties in CESM1-LE projected mean changes (i.e., as shown in Fig. 2a) are assessed using a residual block bootstrap, blocking by run in the 40-member ensemble. That is, we

1) calculate the residuals from the mean model, $\tilde{\mathbf{e}} = \tilde{\mathbf{T}} - \tilde{\boldsymbol{\mu}}$,
2) resample residuals by resampling 40 runs with replacement from the 40 ensemble members, and
3) add the resampled residuals to the originally estimated mean function to obtain a bootstrap dataset, and refit the mean model (4).

This process is repeated 1000 times and we use a 95% bootstrap percentile interval to display uncertainties.

b. Temperature spectrum model

The function $\tilde{\delta}(\omega)$ in the temperature variability change model (4) controls the change in spectral density at frequency ω with GMT changes. This function is estimated similarly to the methods described in Poppick et al. (2016). We briefly describe the fundamental elements of the procedure, but refer readers to Poppick et al. (2016, sections 4.1.1 and S2 and S3 therein) for more detail.

A preliminary estimate of $\tilde{\delta}(\omega)$ is first obtained by maximizing a block composite Whittle likelihood function (Dahlhaus 1997) based on the separate periodograms of each year of JJA temperatures. The estimator based on maximizing this approximate likelihood combines information from the yearly periodograms across runs and yearly blocks, but still produces a rather rough estimate across frequency ω. This is therefore further smoothed across frequencies using an Epanechnikov kernel to obtain the final estimate. The bandwidth of the kernel smoother is chosen to minimize the sum of squared errors from leave-one-out cross-validation. We use a fixed bandwidth for all nonzero frequencies, but do not smooth the estimate of $\tilde{\delta}(0)$ beyond the initial estimate to avoid contaminating information about interannual variability changes with intraseasonal changes.

Approximate standard errors of the resulting estimate of $\tilde{\delta}(\omega)$ can then be derived through a tedious but straightforward calculation involving the second derivative of the composite likelihood function and the kernel smoother weights. The 95% intervals shown in Fig. 2b are the ± 2SE intervals around the final estimate of $\tilde{\delta}(\omega)$. The determination of statistical significance shown in Fig. 4, bottom, is based on first calculating a p value from the Z statistic associated with the estimate of $\tilde{\delta}(\omega)$ and its standard error, and then controlling the false discovery rate (FDR) at 5% using the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995; Wilks 2016). The FDR control procedure accounts for the multiple-testing problem arising from the fact that we are assessing significance at each of the 1344 grid cells (and so would expect many small p values even if there were no grid cells with nonzero true changes).

c. Dewpoint quantile regression model

The function $g_\tau(\cdot)$ in the quantile regression model (8) controls seasonality in dewpoint (after controlling for the local temperature deviation and GMT trend) and the function $h_\tau(\cdot)$ controls the relationship between dewpoint and the local temperature deviation for a fixed day and GMT. As for the mean temperature model, we parameterize $g_\tau(\cdot)$ using the first two seasonal harmonics [see Eq. (A1)]. The function $h_\tau(\cdot)$ is parameterized using a natural spline with 10 degrees of freedom (i.e., 8 knots). The knots are chosen as the empirical quantiles of the local temperature deviation, with boundary knots equal to the minimum and maximum observed value.

Model (8) is estimated for the quantile levels $\tau = 0.005, 0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99$, and 0.995. If quantile functions cross, this is dealt with by reordering fitted values (Chernozhukov et al. 2010). Values for other quantile levels needed for the dewpoint simulation are obtained via linear interpolation. That is, the value $\tau^*_{y,d}$ in Eq. (10) is calculated by linearly interpolating between the two quantile levels yielding fitted values closest to the observed value $D^{(h)}_{y,d}$; the simulation $\hat{D}^{(f)}_{y,d}$ is then calculated by linearly interpolating between the two corresponding future simulation values in Eq. (10).

Finally, note that the reported GSOD data are first rounded to the nearest tenth of a degree Fahrenheit, then converted to Celsius and rerounded to the nearest tenth of a degree Celsius; to avoid numerical issues that arise when estimating quantile regression models with discrete response variables (Machado and Silva 2005), we therefore add a random jitter to each observed dewpoint depression value when fitting the model to observations: dewpoint depression is produced by jittering both temperature and dewpoint separately by $U_F \times 5/9 + U_C$, where U_F and $U_C \overset{\text{indep}}{\sim} \text{Unif}(-0.05, 0.05)$.

APPENDIX B

Details on Calculations of Changes in Risk Probabilities

Figures 6 and 7 show changes in the risk of historically high and low humidity heat events, respectively, at a fixed historically high temperature. Here we provide more details on the quantities we are showing.

Write $\overline{G}^{(h)}$ and $\overline{G}^{(f)}$ for the average GMT forced trend over the years 1990–2005 and 2071–80, respectively. Write $\tilde{\mu}^{(h)}_d$ and $\tilde{\mu}^{(f)}_d$ for the estimated local mean temperature on day d [from Eq. (4)] in CESM1-LE given those historical or future GMT forced trend values. Write $\tilde{T}^{(h)}_{95,m}$ for the empirical 95th percentile of local temperatures in month m over the years 1990–2005 in CESM1-LE (calculated across days in the month, years, and the 40 ensemble members). Finally, write $\tilde{q}^{(h)}_{95,d}$ for the estimated 95th percentile of dewpoints in CESM1-LE on day d from model (8), given local temperature $\tilde{T}^{(h)}_{95,m}$ and GMT anomaly $\overline{G}^{(h)}$; that is,

$$\begin{aligned}\tilde{q}^{(h)}_{95,d} &= F^{-1}_{\tilde{D}_{d,y}}[0.95|\tilde{T}_{d,y} = \tilde{T}^{(h)}_{95,m}, \tilde{G}_y = \overline{G}^{(h)}] \\ &= \tilde{T}^{(h)}_{95,m} - \exp\{\tilde{\alpha}_{0,1-0.95} + \tilde{g}_{1-0.95}(d) \\ &\quad + \tilde{\alpha}_{1,1-0.95}\overline{G}^{(h)} + \tilde{h}_{1-0.95}[\tilde{T}^{(h)}_{95,m} - \tilde{\mu}^{(h)}_d]\},\end{aligned}$$

and write $q^{(h)}_{95,d}$ for the analogous quantity from observations (replacing all of the above relevant quantities with their analogs from the observations).

The *future risk* of a historically high humidity heat event in CESM1-LE is defined as the probability on day d of exceeding $\tilde{q}^{(h)}_{95,d}$ given a historically high local temperature, $\tilde{T}^{(h)}_{95,m}$, but future GMT anomaly $\overline{G}^{(f)}$. That is,

$$\begin{aligned}\tilde{p}^{(f)}_{95,d} &= \Pr[\tilde{D}_{d,y} > \tilde{q}^{(h)}_{95,d}|\tilde{T}_{d,y} = \tilde{T}^{(h)}_{95,m}, \tilde{G}_y = \overline{G}^{(f)}] \\ &= 1 - F_{\tilde{D}_{d,y}}[\tilde{q}^{(h)}_{95,d}|\tilde{T}_{d,y} = \tilde{T}^{(h)}_{95,m}, \tilde{G}_y = \overline{G}^{(f)}].\end{aligned}$$

By definition, the *historical* risk is 0.05 (because $\tilde{q}^{(h)}_{95,d}$ is the estimated historical 95th percentile). We measure the change in risk in terms of the odds ratio,

$$\tilde{\omega}_{95,d} = \frac{\tilde{p}^{(f)}_{95,d}/[1 - \tilde{p}^{(f)}_{95,d}]}{0.05/(1-0.05)},$$

and similarly $\omega_{95,d}$ for the observation-based simulation. If the odds ratio is greater than 1, the future risk is greater than the historical risk. Figure 6 shows $\log_{10}(\tilde{\omega}_{95,d})$ (first row), $\log_{10}(\omega_{95,d})$ (second row), and $\log_{10}(\omega_{95,d}/\tilde{\omega}_{95,d})$ (bottom row) on the days 1 June, 1 July, and 1 August.

Figure 7 is similar except it shows changes in the risk of historically low-humidity heat events, which we define as dewpoints falling below the historical 5th percentile (rather than above the 95th percentile) on days with local temperature equal to the historical 95th percentile value.

REFERENCES

Barreca, A. I., 2012: Climate change, humidity, and mortality in the United States. *J. Environ. Econ. Manage.*, **63**, 19–34, https://doi.org/10.1016/j.jeem.2011.07.004.

Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **57B**, 289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Brands, S., S. Herrera, J. Fernández, and J. M. Gutiérrez, 2013: How well do CMIP5 Earth system models simulate present climate conditions in Europe and Africa? *Climate Dyn.*, **41**, 803–817, https://doi.org/10.1007/s00382-013-1742-8.

Brown, P. J., and A. T. DeGaetano, 2009: A method to detect inhomogeneities in historical dewpoint temperature series. *J. Appl. Meteor. Climatol.*, **48**, 2362–2376, https://doi.org/10.1175/2009JAMC2123.1.

Buzan, J. R., and M. Huber, 2020: Moist heat stress on a hotter Earth. *Annu. Rev. Earth Planet. Sci.*, **48**, 623–655, https://doi.org/10.1146/annurev-earth-053018-060100.

Cannon, A. J., 2018: Multivariate quantile mapping bias correction: An *n*-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dyn.*, **50**, 31–49, https://doi.org/10.1007/s00382-017-3580-6.

——, C. Piani, and S. Sippel, 2020: Bias correction of climate model output for impact models. *Climate Extremes and Their Implications for Impact and Risk Assessment*, Elsevier, 77–104.

Castruccio, S., D. J. McInerney, M. L. Stein, F. Liu Crouch, R. L. Jacob, and E. J. Moyer, 2014: Statistical emulation of climate model projections based on precomputed GCM runs. *J. Climate*, **27**, 1829–1844, https://doi.org/10.1175/JCLI-D-13-00099.1.

CESM Project, 2020: Known issues. UCAR, accessed 6 February 2020,http://www.cesm.ucar.edu/projects/community-projects/LENS/known-issues.html.

Chernozhukov, V., I. Fernández-Val, and A. Galichon, 2010: Quantile and probability curves without crossing. *Econometrica*, **78**, 1093–1125, https://doi.org/10.3982/ECTA7880.

Coffel, E. D., R. M. Horton, and A. de Sherbinin, 2017: Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. *Environ. Res. Lett.*, **13**, 014001, https://doi.org/10.1088/1748-9326/aaa00e.

Dahlhaus, R., 1997: Fitting time series models to nonstationary processes. *Ann. Stat.*, **25** (1), 1–37, https://doi.org/10.1214/aos/1034276620.

Dai, A., J. C. Fyfe, S.-P. Xie, and X. Dai, 2015: Decadal modulation of global surface temperature by internal climate variability. *Nat. Climate Change*, **5**, 555–559, https://doi.org/10.1038/nclimate2605.

Davis, R. E., G. R. McGregor, and K. B. Enfield, 2016: Humidity: A review and primer on atmospheric moisture and human health. *Environ. Res.*, **144**, 106–116, https://doi.org/10.1016/j.envres.2015.10.014.

Dixon, K. W., J. R. Lanzante, M. J. Nath, K. Hayhoe, A. Stoner, A. Radhakrishnan, V. Balaji, and C. F. Gaitán, 2016: Evaluating the stationarity assumption in statistically downscaled climate projections: Is past performance an indicator of future results? *Climatic Change*, **135**, 395–408, https://doi.org/10.1007/s10584-016-1598-0.

Dunn, R., K. Willett, C. Morice, and D. Parker, 2014: Pairwise homogeneity assessment of HadISD. *Climate Past*, **10**, 1501–1522, https://doi.org/10.5194/cp-10-1501-2014.

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, https://doi.org/10.1175/2010JAMC2375.1.

Fischer, E. M., and R. Knutti, 2013: Robust projections of combined humidity and temperature extremes. *Nat. Climate Change*, **3**, 126–130, https://doi.org/10.1038/nclimate1682.

François, B., M. Vrac, A. J. Cannon, Y. Robin, and D. Allard, 2020: Multivariate bias corrections of climate simulations: Which benefits for which losses? *Earth Syst. Dyn.*, **11**, 537–562, https://doi.org/10.5194/esd-11-537-2020.

Gao, M., and C. L. Franzke, 2017: Quantile regression–based spatiotemporal analysis of extreme temperature change in China. *J. Climate*, **30**, 9897–9914, https://doi.org/10.1175/JCLI-D-17-0356.1.

Guo, Q., J. Chen, X. Zhang, M. Shen, H. Chen, and S. Guo, 2019: A new two-stage multivariate quantile mapping method for bias correcting climate model outputs. *Climate Dyn.*, **53**, 3603–3623, https://doi.org/10.1007/s00382-019-04729-w.

Haugen, M. A., M. L. Stein, E. J. Moyer, and R. L. Sriver, 2018: Estimating changes in temperature distributions in a large ensemble of climate simulations using quantile regression. *J. Climate*, **31**, 8573–8588, https://doi.org/10.1175/JCLI-D-17-0782.1.

——, ——, R. L. Sriver, and E. J. Moyer, 2019: Future climate emulations using quantile regressions on large ensembles. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **5**, 37–55, https://doi.org/10.5194/ascmo-5-37-2019.

Hawkins, E., T. M. Osborne, C. K. Ho, and A. J. Challinor, 2013: Calibration and bias correction of climate projections for crop modelling: An idealised case study over Europe. *Agric. For. Meteor.*, **170**, 19–31, https://doi.org/10.1016/j.agrformet.2012.04.007.

Ho, C. K., D. B. Stephenson, M. Collins, C. A. Ferro, and S. J. Brown, 2012: Calibration strategies: A source of additional uncertainty in climate change projections. *Bull. Amer. Meteor. Soc.*, **93**, 21–26, https://doi.org/10.1175/2011BAMS3110.1.

Hsiao, J., A. L. Swann, and S.-H. Kim, 2019: Maize yield under a changing climate: The hidden role of vapor pressure deficit. *Agric. For. Meteor.*, **279**, 107692, https://doi.org/10.1016/J.AGRFORMET.2019.107692.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. T. F. Stocker et al., Eds., Cambridge University Press, 1535 pp., https://doi.org/10.1017/CBO9781107415324.

John, V., and B. J. Soden, 2007: Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophys. Res. Lett.*, **34**, L18704, https://doi.org/10.1029/2007GL030429.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Knutson, T. R., and J. J. Ploshay, 2016: Detection of anthropogenic influence on a summertime heat stress index. *Climatic Change*, **138**, 25–39, https://doi.org/10.1007/s10584-016-1708-z.

Koenker, R., and G. Bassett Jr., 1978: Regression quantiles. *Econometrica*, **46**, 33–50, https://doi.org/10.2307/1913643.

Lamarque, J.-F., and Coauthors, 2010: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application. *Atmos. Chem. Phys.*, **10**, 7017–7039, https://doi.org/10.5194/acp-10-7017-2010.

Lanzante, J. R., K. W. Dixon, M. J. Nath, C. E. Whitlock, and D. Adams-Smith, 2018: Some pitfalls in statistical downscaling of future climate. *Bull. Amer. Meteor. Soc.*, **99**, 791–803, https://doi.org/10.1175/BAMS-D-17-0046.1.

Lee, S.-M., and S.-K. Min, 2018: Heat stress changes over East Asia under 1.5° and 2.0°C global warming targets. *J. Climate*, **31**, 2819–2831, https://doi.org/10.1175/JCLI-D-17-0449.1.

Leeds, W. B., E. J. Moyer, and M. L. Stein, 2015: Simulation of future climate under changing temporal covariance structures. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **1** (1), 1–14, https://doi.org/10.5194/ascmo-1-1-2015.

Li, C., X. Zhang, F. Zwiers, Y. Fang, and A. M. Michalak, 2017: Recent very hot summers in Northern Hemispheric land areas measured by wet bulb globe temperature will be the norm within 20 years. *Earth's Future*, **5**, 1203–1216, https://doi.org/10.1002/2017EF000639.

Li, D., J. Yuan, and R. B. Kopp, 2020: Escalating global exposure to compound heat-humidity extremes with warming. *Environ. Res. Lett.*, **15**, 064003, https://doi.org/10.1088/1748-9326/ab7d04.

Machado, J. A. F., and J. S. Silva, 2005: Quantiles for counts. *J. Amer. Stat. Assoc.*, **100**, 1226–1237, https://doi.org/10.1198/016214505000000330.

Matiu, M., D. P. Ankerst, and A. Menzel, 2016: Asymmetric trends in seasonal temperature variability in instrumental records from ten stations in Switzerland, Germany and the UK from 1864 to 2012. *Int. J. Climatol.*, **36**, 13–27, https://doi.org/10.1002/joc.4326.

McKinnon, K. A., and C. Deser, 2018: Internal variability and regional climate trends in an observational large ensemble. *J. Climate*, **31**, 6783–6802, https://doi.org/10.1175/JCLI-D-17-0901.1.

——, and A. Poppick, 2020: Estimating changes in the observed relationship between humidity and temperature using noncrossing quantile smoothing splines. *J. Agric. Biol. Environ. Stat.*, **25**, 292–314, https://doi.org/10.1007/s13253-020-00393-4.

——, A. Rhines, M. P. Tingley, and P. Huybers, 2016: The changing shape of Northern Hemisphere summer temperature distributions. *J. Geophys. Res. Atmos.*, **121**, 8849–8868, https://doi.org/10.1002/2016JD025292.

——, A. Poppick, E. Dunn-Sigouin, and C. Deser, 2017: An "observational large ensemble" to compare observed and modeled temperature trend uncertainty due to internal variability. *J. Climate*, **30**, 7585–7598, https://doi.org/10.1175/JCLI-D-16-0905.1.

Mehrotra, R., and A. Sharma, 2015: Correcting for systematic biases in multiple raw GCM variables across a range of timescales. *J. Hydrol.*, **520**, 214–223, https://doi.org/10.1016/j.jhydrol.2014.11.037.

——, and ——, 2016: A multivariate quantile-matching bias correction approach with auto- and cross-dependence across multiple time scales: Implications for downscaling. *J. Climate*, **29**, 3519–3539, https://doi.org/10.1175/JCLI-D-15-0356.1.

Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, https://doi.org/10.1007/s10584-011-0156-z.

Pal, J. S., and E. A. Eltahir, 2016: Future temperature in southwest Asia projected to exceed a threshold for human adaptability. *Nat. Climate Change*, **6**, 197–200, https://doi.org/10.1038/nclimate2833.

Piani, C., G. Weedon, M. Best, S. Gomes, P. Viterbo, S. Hagemann, and J. Haerter, 2010: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol.*, **395**, 199–215, https://doi.org/10.1016/j.jhydrol.2010.10.024.

Poppick, A., D. J. McInerney, E. J. Moyer, and M. L. Stein, 2016: Temperatures in transient climates: Improved methods for simulations with evolving temporal covariances. *Ann. Appl. Stat.*, **10**, 477–505, https://doi.org/10.1214/16-AOAS903.

——, E. J. Moyer, and M. L. Stein, 2017: Estimating trends in the global mean temperature record. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **3**, 33–53, https://doi.org/10.5194/ascmo-3-33-2017.

Pryor, S., and J. Schoof, 2016: Evaluation of near-surface temperature, humidity, and equivalent temperature from regional climate models applied in type II downscaling. *J. Geophys. Res. Atmos.*, **121**, 3326–3338, https://doi.org/10.1002/2015JD024539.

Reich, B. J., 2012: Spatiotemporal quantile regression for detecting distributional changes in environmental processes. *J. Roy. Stat. Soc.*, **61C**, 535–553, https://doi.org/10.1111/j.1467-9876.2011.01025.x.

Rhines, A., M. P. Tingley, K. A. McKinnon, and P. Huybers, 2015: Decoding the precision of historical temperature observations. *Quart. J. Roy. Meteor. Soc.*, **141**, 2923–2933, https://doi.org/10.1002/qj.2612.

——, K. A. McKinnon, M. P. Tingley, and P. Huybers, 2017: Seasonally resolved distributional trends of North American temperatures show contraction of winter variability. *J. Climate*, **30**, 1139–1157, https://doi.org/10.1175/JCLI-D-16-0363.1.

Santer, B. D., T. M. Wigley, M. E. Schlesinger, and J. F. Mitchell, 1990: Developing climate scenarios from equilibrium GCM results. Max-Planck-Institut für Meteorologie, Tech. Rep. 47, 31 pp.

Schoof, J., T. Ford, and S. Pryor, 2017: Recent changes in U.S. regional heat wave characteristics in observations and reanalyses. *J. Appl. Meteor. Climatol.*, **56**, 2621–2636, https://doi.org/10.1175/JAMC-D-16-0393.1.

——, S. Pryor, and T. Ford, 2019: Projected changes in United States regional extreme heat days derived from bivariate quantile mapping of CMIP5 simulations. *J. Geophys. Res. Atmos.*, **124**, 5214–5232, https://doi.org/10.1029/2018JD029599.

Seager, R., A. Hooks, A. P. Williams, B. Cook, J. Nakamura, and N. Henderson, 2015: Climatology, variability, and trends in the us vapor pressure deficit, an important fire-related meteorological quantity. *J. Appl. Meteor. Climatol.*, **54**, 1121–1141, https://doi.org/10.1175/JAMC-D-14-0321.1.

Simmons, A., K. Willett, P. Jones, P. Thorne, and D. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets. *J. Geophys. Res.*, **115**, D01110, https://doi.org/10.1029/2009JD012442.

Tian, B., E. J. Fetzer, B. H. Kahn, J. Teixeira, E. Manning, and T. Hearty, 2013: Evaluating CMIP5 models using airs tropospheric air temperature and specific humidity climatology. *J. Geophys. Res. Atmos.*, **118**, 114–134, https://doi.org/10.1029/2012JD018607.

Vrac, M., 2018: Multivariate bias adjustment of high-dimensional climate simulations: The Rank Resampling for Distributions and Dependences (R^2D^2) bias correction. *Hydrol. Earth Syst. Sci.*, **22**, 3175–3196, https://doi.org/10.5194/hess-22-3175-2018.

——, and P. Friederichs, 2015: Multivariate—intervariable, spatial, and temporal—bias correction. *J. Climate*, **28**, 218–237, https://doi.org/10.1175/JCLI-D-14-00059.1.

Wilks, D., 2016: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. *Bull. Amer. Meteor. Soc.*, **97**, 2263–2273, https://doi.org/10.1175/BAMS-D-15-00267.1.

Willett, K., R. Dunn, P. Thorne, S. Bell, M. de Podesta, D. Parker, P. Jones, and C. Williams Jr., 2014: HadISDH land surface multi-variable humidity and temperature record for climate monitoring. *Climate Past*, **10**, 1983–2006, https://doi.org/10.5194/cp-10-1983-2014.

Yuan, J., M. L. Stein, and R. E. Kopp, 2020: The evolving distribution of relative humidity conditional upon daily maximum temperature in a warming climate. *J. Geophys. Res. Atmos.*, **125**, e2019JD032100, https://doi.org/10.1029/2019jd032100.

Zhao, Y., A. Ducharne, B. Sultan, P. Braconnot, and R. Vautard, 2015: Estimating heat stress from climate-based indicators: Present-day biases and future spreads in the CMIP5 global climate model ensemble. *Environ. Res. Lett.*, **10**, 084013, https://doi.org/10.1088/1748-9326/10/8/084013.

Zscheischler, J., E. M. Fischer, and S. Lange, 2019: The effect of univariate bias adjustment on multivariate hazard estimates. *Earth Syst. Dyn.*, **10**, 31–43, https://doi.org/10.5194/esd-10-31-2019.