The evolving distribution of relative humidity conditional upon daily maximum temperature in a warming climate

Jiacan Yuan[1], Michael L Stein[2], and Robert E Kopp[2]

[1]Fudan University
[2]Rutgers University

November 21, 2022

Abstract

The impacts of heat waves in a warming climate depend not just on changing temperatures but also on changing humidity. Using 35 simulations from the Community Earth System Model Large Ensemble (CESM LENS), we investigate the long-term evolution of the joint distribution of summer relative humidity (RH) and daily maximum temperature () in four U.S. cities (New York City, Chicago, Phoenix, New Orleans) under the high-emissions Representative Concentration Pathway (RCP) 8.5. We estimate the conditional quantiles of RH given by quantile regression models, using functions of temperature for each month and city for three time periods (1990-2005, 2026-2035, and 2071-2080). Quality of fit diagnostics indicate that these models accurately estimate conditional quantiles for each city. As expected, each quantile of increases from 1990-2005 to 2071-2080, while mean RH decreases modestly. For a fixed , the high quantiles of RH (and thus of heat index and dew point) increase from 1990-2005 to 2071-2080 in all four cities. This result suggests that the health impacts of a day of a given will increase in a warming climate due to the increase of RH. Conditional upon a fixed quantile of , the median and high quantiles of RH decrease, while those of heat index and dew point both increase. This result suggests that, despite a modest decrease in median relative humidity, heat stress impacts in a warming climate will increase faster than temperatures alone would indicate.

The evolving distribution of relative humidity conditional upon daily maximum temperature in a warming climate

Jiacan Yuan[1,2,3] *, Michael L. Stein[3,4,5] , Robert E. Kopp[2,3]

1. Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai, China

2. Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, New Jersey, United States

3. Rutgers Institute of Earth, Ocean, and Atmospheric Sciences, Rutgers University, New Brunswick, New Jersey, United States

4. Department of Statistics, University of Chicago, Chicago, Illinois, United States

5. Department of Statistics, Rutgers University, New Brunswick, New Jersey, United States

Corresponding author: Jiacan Yuan (jcyuan@fudan.edu.cn)

*Current addresses: Fudan University. 2005 Songhu Road, Environmental Building, Shanghai, 200438, Shanghai, China

Key Points:

1. The quantile-regression models accurately estimate the conditional distribution of relative humidity given temperature
2. The health impacts of a day at a fixed daily maximum temperature will increase due to the increase of relative humidity as climate warms
3. If considering relative humidity, heat stress impacts in a warming climate will increase faster than temperatures alone would indicate

Abstract

The impacts of heat waves in a warming climate depend not just on changing temperatures but also on changing humidity. Using 35 simulations from the Community Earth System Model Large Ensemble (CESM LENS), we investigate the long-term evolution of the joint distribution of summer relative humidity (RH) and daily maximum temperature (*Tmax*) in four U.S. cities (New York City, Chicago, Phoenix, New Orleans) under the high-emissions Representative Concentration Pathway (RCP) 8.5. We estimate the conditional quantiles of RH given *Tmax* by quantile regression models, using functions of temperature for each month and city for three time periods (1990-2005, 2026-2035, and 2071-2080). Quality of fit diagnostics indicate that these models accurately estimate conditional quantiles for each city. As expected, each quantile of *Tmax* increases from 1990-2005 to 2071-2080, while mean RH decreases modestly. For a fixed *Tmax*, the high quantiles of RH (and thus of heat index and dew point) increase from 1990-2005 to 2071-2080 in all four cities. This result suggests that the health impacts of a day of a given *Tmax* will increase in a warming climate due to the increase of RH. Conditional upon a fixed quantile of *Tmax*, the median and high quantiles of RH decrease, while those of heat index and dew point both increase. This result suggests that, despite a modest decrease in median relative humidity, heat stress impacts in a warming climate will increase faster than temperatures alone would indicate.

1. Introduction

Heat waves, events in which sweltering weather lasts days to weeks, negatively impact human health, ecosystems, crop yields, and physical infrastructure (Allen et al., 2010; Fontana, Toreti, Ceglar, & De Sanctis, 2015; Ramamurthy, Li, & Bou-Zeid, 2017; Zuo et al., 2015). As the climate continues to warm, heat waves are projected to increase in frequency and duration (Collins et al., 2013; Meehl & Tebaldi, 2004; Papalexiou, AghaKouchak, Trenberth, & Foufoula-Georgiou, 2018). In addition, the portion of the world's population and land area that are exposed to deadly heat (referring to extremely hot conditions that may cause death) is projected to increase under even a scenario with aggressive mitigation of greenhouse gas emissions (Mora et al., 2017). Thus, climate change raises serious concerns on the growing impact of heat waves (Stocker et al., 2013).

Although many studies have targeted heat waves and their impact on human health (Basu & Samet, 2002; Mazdiyasni et al., 2017; Meehl & Tebaldi, 2004; Mora et al., 2017; Orlowsky & Seneviratne, 2012; Tebaldi, Hayhoe, Arblaster, & Meehl, 2006), there is still debate on the most appropriate ways to identify extreme values of heat for the assessment of heat-related mortality and illness. Much of the literature only uses temperature to describe heat extremes (Carleton et al., 2019; Dosio, Mentaschi, Fischer, & Wyser, 2018; Kodra & Ganguly, 2014; Mazdiyasni et al., 2017; Meehl & Tebaldi, 2004; Tebaldi et al., 2006). However, other climate variables (e.g. humidity, solar radiation, wind speed) may also contribute to human discomfort and mortality. Mora et al. (2017) assessed multiple pairs of climate variables: surface temperature, relative humidity, solar radiation, and wind speed. They found the pair combining surface temperature and relative humidity most accurately identifies lethal conditions. High humidity is an important contributor to heat stress as it can reduce the human body's capability to remove metabolic heat by sweating. Some studies consider both temperature and humidity to identify the extreme value of heat. Fischer and Knutti (2013) suggested that the quantities jointly defined by temperature and relative humidity can reduce uncertainty of future projections of heat extremes. Heat index, which is broadly used in weather warning systems for heat stress, can be estimated by a multiple-regression model of temperature and relative humidity (Rothfusz, 1990). Russo et al. (2017) introduced a new Apparent Heat Wave Index (AHWI) that utilizes both daily maximum temperature and daily minimum relative humidity to define heat waves and physiologic stress.

In the context of global warming, the long-term trend of land surface temperature is positive in both present observations and future projections (Byrne & O'Gorman, 2018; Dai, 2006; Sutton, Dong, & Gregory, 2007). Although the specific humidity increases as climate warms(Stocker et al., 2013), the long-term trend of surface relative humidity (RH) over land decreases in both observations and future projections. Willett et al. (2014) found that RH averaged over land in-situ observations decreased from 2000 and 2013. Coupled Model Intercomparison Project Phase 5 (CMIP5) global climate models project that surface RH will decrease over most land areas (except for parts of tropical Africa and South Asia) by the end of this century, possibly due to the faster increase in surface air temperature over land than over the ocean (Byrne & O'Gorman, 2018; Flato et al., 2013; O'Gorman & Muller, 2010). Given projected increases in temperature and decreases in RH, the combined effects of temperature and RH on future heat extremes is unclear. In this study, we use both temperature and RH as contributing variables to identify heat extremes and investigate the long-term evolution of their joint distribution.

Conventional approaches assessing extreme events may apply an assumption about the tail of distribution of the contributing variable (e.g. Kharin, Zwiers, Zhang, & Wehner, 2013; Kodra & Ganguly, 2014), and assume only the parameters of the presumed distribution changes when climate gets warmer. However, the distributions of these contributing variables generally do not follow a standard distribution. Furthermore, shapes of distributions may change quite substantially in a warming climate (Haugen et al. 2018). In this study, we develop a statistical model to characterize the distribution of relative humidity (RH) conditional on daily maximum temperature (*Tmax*), as well as its evolution, without assuming any particular parametric form for this distribution. To reach this end, we applied quantile regression to the Community Earth System Model Large Ensemble (LENS, Kay et al., 2015), which provides sufficient volume of samples to allow accurate estimation for the tails of the distribution of a climate variable (Haugen, Stein, Moyer, & Sriver, 2018).

Quantile regression is a form of regression analysis that estimates conditional quantile functions – models in which, for any given quantile (τ) between 0 and 1, the τ^{th} conditional quantile of the response variable is expressed as a linear function of predictor variables (Koenker & Hallock, 2001), where the coefficients in this linear function can vary with τ. Haugen et al. (2018) applied quantile regression to temperature in an ensemble of simulations from CESM (Sriver, Forest, & Keller, 2015) to study the evolving distribution of temperature in a warming climate. They constructed a quantile regression model which continuously represents the smooth evolution of temperature distributions both day-to-day over an annual cycle and year-to-year over longer temporal trends over North America. Quantile regression provides a natural way to study the nuanced changes in distributions and, especially tails, which is essential for studies of weather extremes. Haugen et al. (2018) used temperature as the response and function of days in a year and years as predictors. Here, we focus on using functions of temperature as predictors for RH. Estimating the distribution of RH conditional on *Tmax* (RH | *Tmax*) provides flexibility for quantifying heat extremes by RH and *Tmax* simultaneously. This way allows us to look at RH at given temperature or RH at given temperature quantile. In addition, it allows us to assess any variables that take both temperature and RH into account, such as dew point, heat index, and wet bulb temperature.

We developed models for four major U.S. cities (Fig. S1 in Supplement Information) in different climate settings: New York City (NYC) has a humid subtropical climate; Chicago (CHI) has a hot-summer humid continental climate; Phoenix (PHX) has a hot-desert climate; New Orleans (NOLA) also has a humid subtropical climate, but has warmer and more humid winters than New York City. These four cities were selected to test the sensitivity of the approach in a variety of climates. Our approach can be applied to different locations with different climate background.

Section 2 describes the data sources used in this study, as well as the approaches used to calculate key metrics. Section 3 numerically diagnoses the joint distribution of RH and *Tmax,* identifying features used to help select the form of basis functions used in the statistical models. Section 4 describes the statistical methodology. Section 5 presents the approaches used to validate the quality of fit of the statistically estimated joint distribution. Section 6 presents key results, including the estimated quantiles of RH | *Tmax*, and discusses the implications of the estimated quantiles of RH | *Tmax* for extreme heat in future projections. Details on the selection of basis functions for the statistical models are described in the Appendix.

2. Data Sources and meteorological metrics

2.1 Data

The CESM LENS provides sufficiently large samples to yield accurate estimates of conditional quantiles via quantile regression. The LENS dataset is a 40-member initial-conditions ensemble forced by Representative Concentration Pathway (RCP) 8.5, which represents a high-emissions scenario in which emissions continue to rise through the 21st century (van Vuuren et al., 2011). Of the 40 ensemble members, 35 were run on the Yellowstone supercomputer, while the other 5 members were run on University of Toronto supercomputer. To avoid systematic biases, we used the 35 members obtained from the same machine. The principal variables used in this study were six-hourly surface temperature (TREFHT) and six-hourly surface specific humidity (QREFHT, kg/kg). These six-hourly data are only available for three periods: 1990-2005, 2026-2035, and 2071-2080. We extract the *Tmax* based on the 6-hourly data. RH was determined from the specific humidity q and surface pressure P at the time at which the *Tmax* was observed (Lawrence, 2005).

$$RH = \frac{q}{q_{sat}} = \frac{q\ (P - e_{sat})}{0.622 \times e_{sat}} \tag{1}$$

$$e_{sat} = C \exp\left(\frac{A\ Tmax}{Tmax + B}\right) \tag{2}$$

Where A = 17.625, B = 243.04℃, C = 610.94℃. These coefficients are evaluated by Alduchov and Eskridge (1996), who recommend that the equation (2) provides an estimation of e_{sat} with a relative error of <0.4% over the range $-40℃ \leq Tmax \leq 50℃$.

We select four grid cells corresponding to the four focal cities (Figure S1): (41.00°N, 73.75°W) for NYC, (41.00°N, 271.25°W) for CHI, (29.69°N, 270.00°W) for NOLA, and (33.46°N, 247.50°W) for PHX. Note that we use a grid cell southwest of the urban area of Chicago to represent CHI. The urban Chicago cell, of which Lake Michigan comprises about half, produces some extremely hot and humid days, for which the corresponding heat indices in the period of 1990-2005 are substantially higher than the historical record at Chicago (Fig. S2). This mismatch may arise due to the poor representation of Lake Michigan in CESM (Subin, Riley, & Mironov, 2012), and motivates our choice of a more inland cell.

To validate the model data with historical observational data, we use version 3.0 of Hadley Centre Global Sub-Daily Station Observation's (HadISD) air temperature and dew point temperature (T_d). HadISD is a quality controlled global sub-daily dataset that contains weather data at the station level (Dunn et al., 2012). HadISD has hourly temporal resolution, but was converted to six hourly data by taking the instantaneous temperature and dew-point temperature (T_d) at 0:00, 6:00, 12:00, and 18:00 of the Coordinated Universal Time, to be consistent with the LENS data. Then, we select *Tmax* based on the 6-hourly data and the T_d at time when the *Tmax* is selected. The RH is obtained from *Tmax* and the corresponding T_d (Lawrence, 2005):

$$RH = \exp\left(A\left(\frac{T_d}{T_d + B} - \frac{Tmax}{Tmax + B}\right)\right) \tag{3}$$

2.2 Computation of meteorological metrics

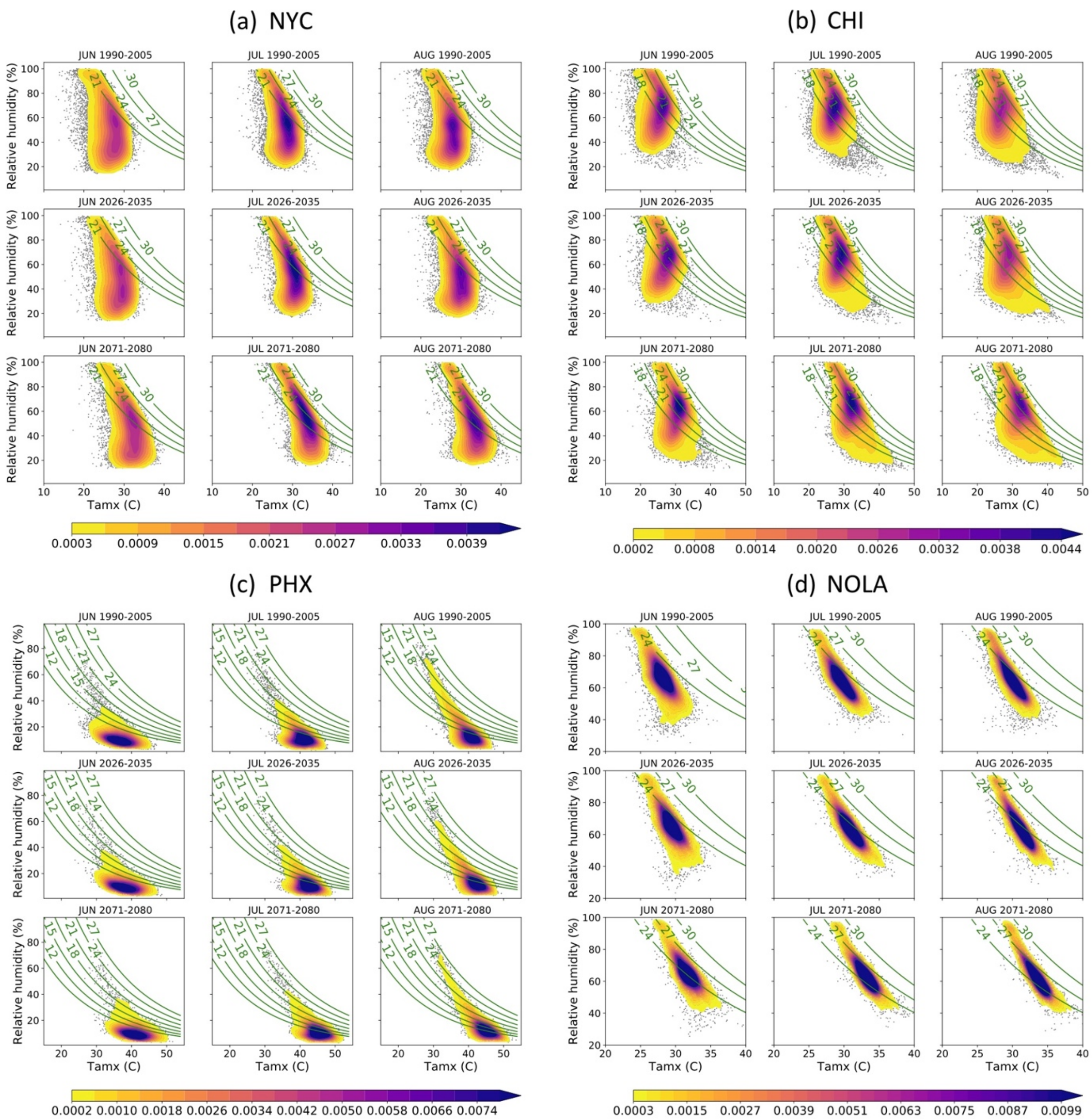
The dew-point temperature (T_d) is the temperature when an air particle reaches saturation by cooling the air isobarically. We use an empirical metric called Magnus formula (Alduchov & Eskridge, 1996; Gibbins, 1990) to calculate T_d from *RH* and *Tmax*.

$$T_d = \frac{B\,\gamma}{A - \gamma} \tag{4}$$

$$\gamma = \ln RH + A\frac{Tmax}{Tmax + B} \tag{5}$$

The United States National Weather Service (NWS) uses a heat index in their heat stress early warning system. Similarly, we use the empirical equation developed by Rothfusz (1990) who performed a multiple regression analysis on the data of heat index from Steadman's comfort model (Steadman, 1979). The heat index is calculated using temperature and RH (Rothfusz, 1990). The NWS defines several categories for heat wave by heat index: 27℃ $-$ 32℃ is caution;

Figure 1. Joint distribution of RH and *Tmax* in summer months of three periods at (a) New York (NYC), (b) Chicago (CHI), (c) Phoenix (PHX), (d) New Orleans (NOLA). Gray dots are observations of *Tmax* and RH from CESM LENS simulations. Shading represents the density of the observations. Contours are the dew point temperature calculated from RH and *Tmax* (Units: ℃)



32°C-41°C is extreme caution; 41°C-54°C is danger; >54°C is extreme danger (NWS Weather Forecast Office 2011).

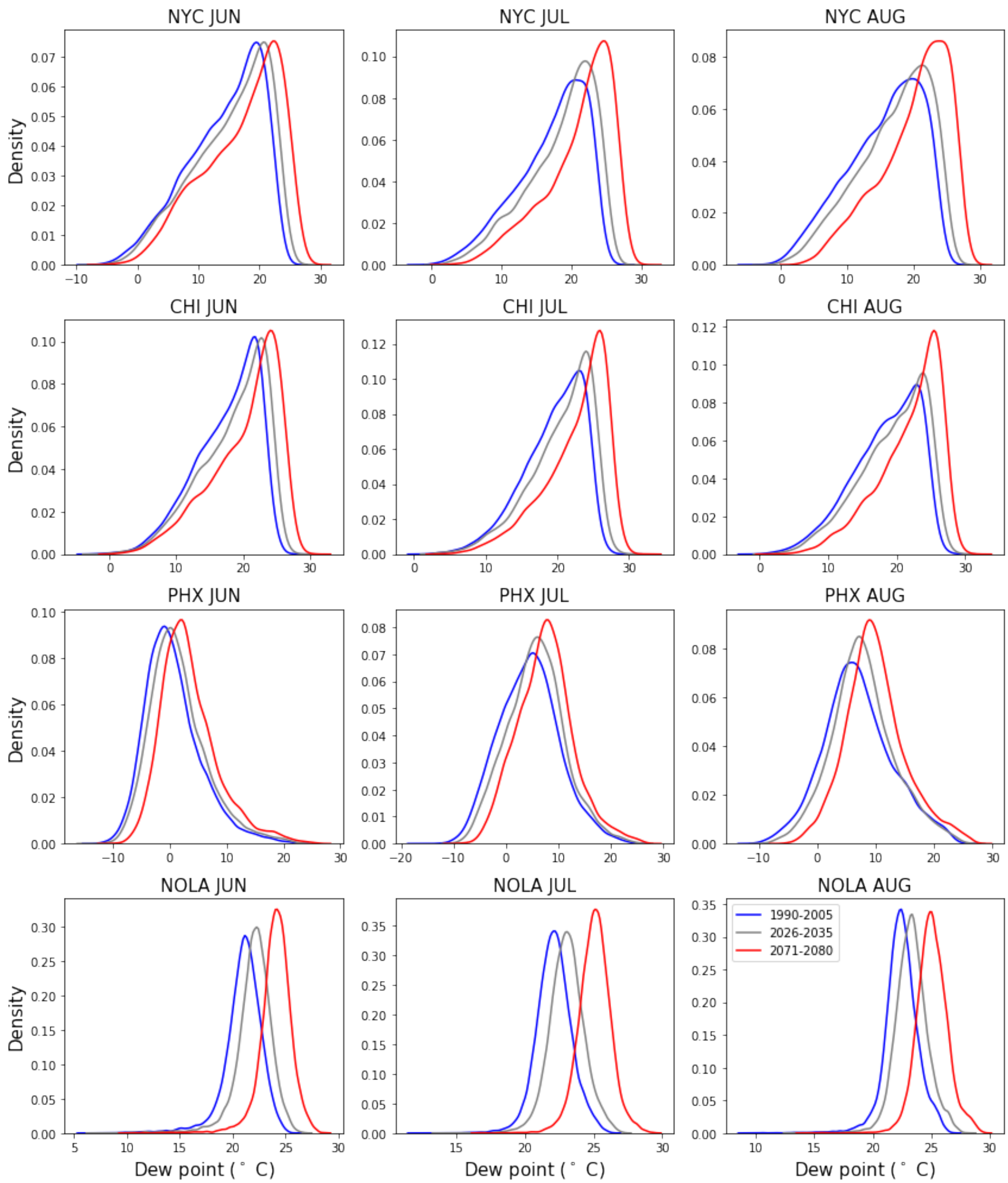
3. Joint distribution between RH and *Tmax*

To determine the form of basis functions used in the statistical models, we examine numerically the joint distribution of RH and *Tmax* for three different periods at four cities in the three summer months (June, July, August) (Fig. 1). From 1990-2005 to 2071-2080, the shape of joint distributions in each month and city changes, and the location of the joint distribution shifts toward hotter dew point. In NYC, CHI and some months in NOLA (e.g. June in three periods, August in 1990-2005), the joint distribution is constrained by two boundaries at the high end of each variable: a cap at 100% relative humidity and a cutoff (a sharp boundary of the scatters preventing the RH reaching 100%) at the hot end of the temperature. A kink is observed at the intersection of these two boundaries. Above the temperature at the kink (T_0), the maximum observed RH decreases in a fairly linear fashion.

The cutoff may be due to a maximum in surface air moisture availability, constrained by local meteorological conditions. Since saturated vapor pressure increases with temperature, if the water vapor supplied to the air does not change, relative humidity will decrease with temperature. The relation of the cutoff to contours of constant dew point supports this hypothesis (Fig. 1). In NYC and CHI, the cutoff generally tracks a dew point contour, with the level of the dew point contour increasing with time. For instance, the cutoff in NYC is near a dew point of 25°C during 1990-2005, 26°C during 2026-2035, and 29°C during 2071-2080. As the diurnal variability of dew point is small in most regions of North America (Schwartzman, Michaels, & Knappenberger, 1998), the dew point paralleled by the cutoff is approximately equivalent to the dew point in the morning when the air parcel is saturated. In other words, the morning dew point largely determines the moisture for the day when the water supply is not the limiting factor. As the climate warms, the daily attainable dew point elevates because of the increase in capacity of containing water vapor in the air. Therefore, the cutoffs shift toward higher dew point in a warmer climate. Other factors may also influence the local moisture, e.g. regional enhancement of convection, which could pump moisture from the boundary layer to the free atmosphere (Schwartzman et al., 1998; Sherwood, Roca, Weckwerth, & Andronova, 2010). By contrast, the cutoffs in NOLA are not parallel with the dew point contours. This may be because factors other than dew point (e.g. convection) play an important role in constraining the air moisture. A specific investigation is out of the scope of this study. In PHX, the local moisture supply is insufficient to produce saturated air due to the desert climate, so the cutoffs and kinks are not observed in the joint distributions of PHX.

To test if these features are also shown in the observations, we examined the joint distribution of RH and *Tmax* at stations near the four cities, respectively (Fig. S3). As we just have one climate realization for the observations, the station data (gray dots) are sparse. We used a period from 1980 to 2005 for the station data. This period is longer than the historical period 1990-2005 in the LENS simulations, because this choice was a compromise between having more data and wanting to match the LENS historical period reasonably well. The general pattern for the joint distributions based on the HadISD station data is similar whether one uses observations during 1980-2005 or 1990-2005. The patterns from the HadISD data in the period of 1980-2005 qualitatively resemble the patterns from the LENS simulations in the period of 1990-2005 in all months and at all four cities. In particular, a fairly clear cutoff is observed in the joint distribution from station data for all summer months in NYC and in June and July in CHI. These results suggest that the patterns of joint distribution between RH and *Tmax* simulated by LENS are real physical patterns.

Figure 2. Frequency distribution of dew point temperatures calculated from RH and *Tmax* in June, July, and August at New York (NYC), Chicago (CHI), Phoenix (PHX), New Orleans (NOLA) using the CESM LENS data (Units: ℃). Blue lines denote the period 1990-2005, grey 2026-2035, red 2071-2080.



We note that LENS generally shows lower maximum dew points than HadISD. Here, the focus is on changes in heat extremes from the historical to future periods, so we do not consider bias correction. Haugen et al. (2019) show how to use quantile regressions to combine observational records with simulations of present and future climate to produce bias-corrected future climate simulations.

As the kinks generally correspond to the maximum dew point of the simulations in a time period, we investigate the density of dew points in each month of each time period to find an objective criterion for the kink temperature (Fig. 2). The densities of dew point in NYC and CHI display a cliff shape at the hot end. The cliff feature is also seen in some months of NOLA, but is not as sharp as that in NYC and CHI. There is no cliff feature observed in PHX. This is consistent with the cutoff of the joint distribution between RH and *Tmax*. Therefore, we select the kink temperature (T_0) based on the cliff feature in the density distribution of the dew point. The criteria for T_0 selection are described in *Appendix A*.

4. Statistical methodology

Standard parametric distributions can not well capture the features of joint distributions between *Tmax* and RH, especially when a kink occurs. Here we construct conditional quantile regression models of RH | *Tmax* for each quantile. The quantile regression approach is applied to the joint distribution of RH and *Tmax* in summer months (June, July, August) over the three periods, respectively. Based on the empirical characteristics of the joint distributions, we use two kinds of basis functions in the quantile regression models: one is a kink function to capture the kink when it is apparent in the dew point density; the other is a set of cubic-spline basis functions of *Tmax*, which is used to capture smooth variation in RH as *Tmax* varies. Cubic splines in *Tmax* were used to flexibly model RH as a function of *Tmax* within each month. Our models for the τ^{th} quantile of RH | *Tmax* are of the form:

$$\widehat{RH}_\tau(Tmax) = \theta + \gamma\,(Tmax - T_0)_+ + \sum_{j=1}^{m} \eta_j\, K_j(Tmax) \tag{6}$$

where T_0 is the temperature at the kink. $(Tmax - T_0)_+$ denotes the basis functions that capture the kink, and $(\,)_+$ is defined as

$$(x)_+ = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{7}$$

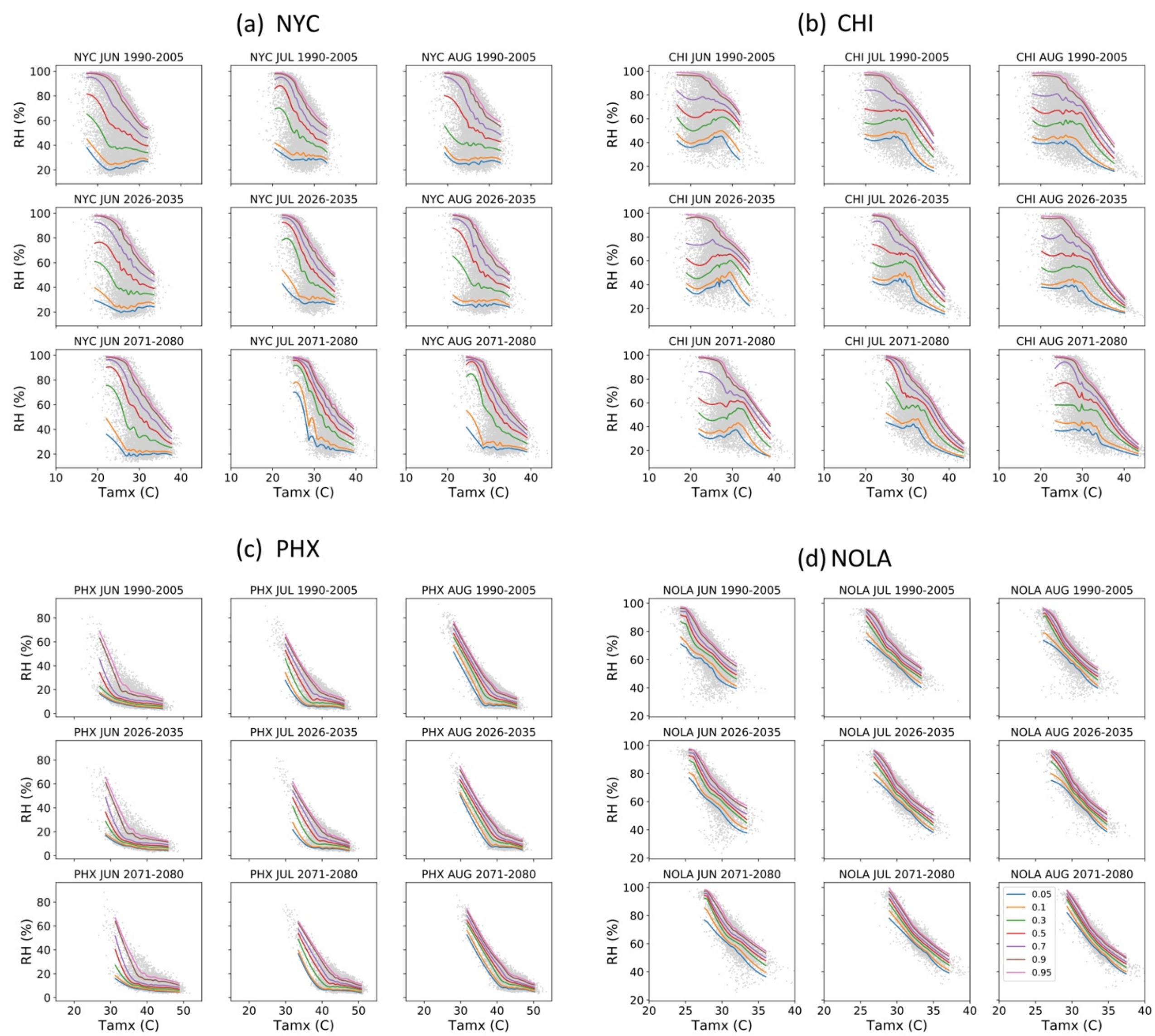
Here, θ is the intercept; γ, η_j are coefficients of basis functions; $K_j(Tmax)$ represents a cubic spline basis function; and *m* is the number of cubic spline basis functions. The metrics for selecting the *m* and T_0 for each city are described in the Appendix. The conditional quantile regression model estimates $\widehat{RH}_\tau$ at specific quantile τ, so that the τ^{th} fraction of the residual between estimated RH ($\widehat{RH}_\tau$) and observations RH (RH_i) in LENS is positive, while a fraction of $1-\tau$ of the residual is negative. Mathematically, the quantile regression obtains the best estimates of coefficients through solving a minimization problem as follows (Koenker & Bassett, 1978):

$$\tau \sum_{i=1}^{n} \left(RH_i - \widehat{RH}_\tau(Tmax_i)\right)_+ + (1-\tau) \sum_{i=1}^{n} \left(\widehat{RH}_\tau(Tmax_i) - RH_i\right)_+ \tag{8}$$

where $Tmax_i$ and RH_i indicate the observed value of a quantity on day i. $\widehat{RH}_\tau(Tmax_i)$ is the estimated quantile on day *i*, obtained by replacing $\theta, \gamma, \eta_{1,2,\ldots,m}$ by estimates that minimize equation (7).

The estimated quantiles produced by the quantile regression models closely match the variation of RH | *Tmax* in the simulated data from LENS for each month, period and city (Figure 3). These quantiles show the differences in RH as *Tmax* varies within a month, and the feature of kink/cutoff when it is available. Using the estimated quantiles of RH | *Tmax*, we can flexibly

Figure 3. Estimated quantiles of RH | *Tmax* based on CESM LENS simulations in June-August of three periods at (a) NYC, (b) CHI, (c) PHX, (d) NOLA. Gray dots are observations of *Tmax* and RH from CESM LENS simulations. Lines represent the estimated quantiles at 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95.



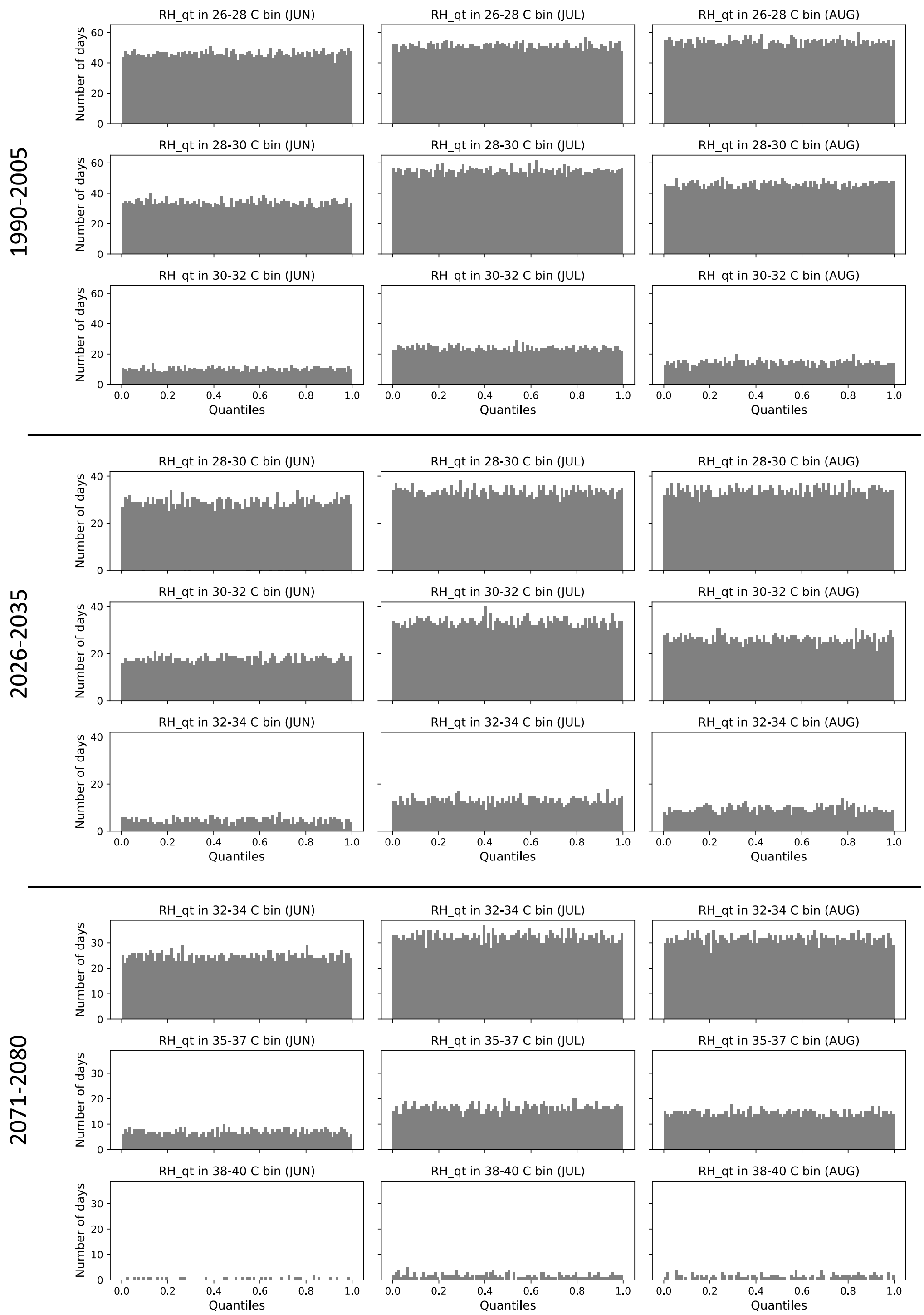
estimate distributions of multiple metrics that describe heat extremes, such as dew point and heat index. There are some spikes shown in the lines of low quantiles in some months in some cities, e.g. 0.025 – 0.7 quantiles in August of 2071-2080 in Chicago. These spikes are due to the overfitting by the kink function. As the spikes only appear in low quantiles that are not important for studies of extremes, we still use the kink function in the model as it is important for estimating the kink in high quantiles, but recognize that these model specifications may not be optimal for all quantiles in all cities in all months.

5. Model validation

To evaluate how well the statistical model fits the observations (quality of fit), we construct empirical inverse quantiles of the relative humidity data in various temperature ranges. The empirical inverse quantiles are calculated as follows.

1. Use the quantile regression model to estimate values of $\widehat{RH}_{\tau}(Tmax)$ in 99 quantiles from 0.01 to 0.99 and for the range of *Tmax* present in the data. Use these 99 quantiles as boundaries of 100 bins between 0 and 1.
2. For temperature intervals of two-degree width, assign each day to an interval based on the observed value of *Tmax* for that day.
3. For a particular temperature interval, the RH in these selected days are compared with the values of the 99 $\widehat{RH}_{\tau}(Tmax)$ quantiles, and assigned to the corresponding RH quantile bins.
4. The estimated quantiles of the corresponding RH in the temperature interval are displayed in a histogram to show the number of days falling into 100 bins from 0-0.01 quantile to 0.99-1 quantile (e.g. see Fig. 4). The more uniform the histogram is, the better the quantile model fits the data.

Figure 4. Empirical inverse quantiles of the statistical models emulating the quantiles of simulated data in LENS in New York during (a) 1990-2005; (b) 2026-2035; (c) 2071-2080. The histograms represent the number of RH events falls into 100 bins of estimated quantiles of RH at three given temperature intervals (2 degrees per interval) in June, July, August.



Taking NYC as an example (Fig. 4), the simulated RH data from LENS are evenly distributed in the 100 bins edged by quantiles, which are estimated by the quantile regression models, given randomly picked temperature intervals for all three months in three periods. Similar features are observed in the other three cities (Fig. S4, S5, S6). These results indicate that the models we selected fit the data well for all four cities.

6. Results

To investigate the long-term evolution of heat extremes, we focus upon the conditional median of RH | *Tmax* and upon the conditional 0.95 quantile of RH | *Tmax*, which we take as representative of the tail of the conditional distribution (upper panels of Fig. 5). As noted, the shape of the marginal *Tmax* distribution in each time period (lower panels of Fig. 5) noticeably deviates from a normal distribution. Furthermore, the changes from the period 1990-2005 to the period 2071-2080 include not only increases in mean and changes in standard deviation, but also changes in skewness and kurtosis. For example, in CHI in July, the mean increases from 26.9°C and 32.0°C, while the standard deviation grows from 2.7°C and 3.3°C. The skewness and kurtosis of *Tmax* also increase over time, indicating a more right-skewed and fatter tail distribution in the future climate. The changes in marginal distribution of RH are smaller (Fig. S7). The means of RH decrease in all cities and all months (in a range between -0.4% and -7.3%), while changes in other moments of RH are not consistent across cities and months. The standard deviations of RH slightly increase in CHI, exhibit almost no change in NYC, and decrease in PHX and NOLA.

Two alternative ways of investigating RH | *Tmax* are to look at the change in RH at a fixed quantile of *Tmax* ($Tmax_{\tau}$; Tables 1, 2) and to look at the change in RH at a fixed value of *Tmax* (Tables 3, 4). The former is more representative of the overall shift in the joint distribution, while the latter is relevant to characterizing the impacts of a day of a particular *Tmax*.

Figure 5. Estimated quantiles of relative humidity (RH) given daily maximum temperature (*Tmax*) in the three periods: 1990-2005, 2026-2036, and 2071-2080. Median and 0.95 quantiles are displayed by dashed and solid lines, respectively. Marginal distribution of temperature is also shown on the lower panel. The dashed black lines denote the values of *Tmax* at 0.95 quantile during the period of 1990-2005.

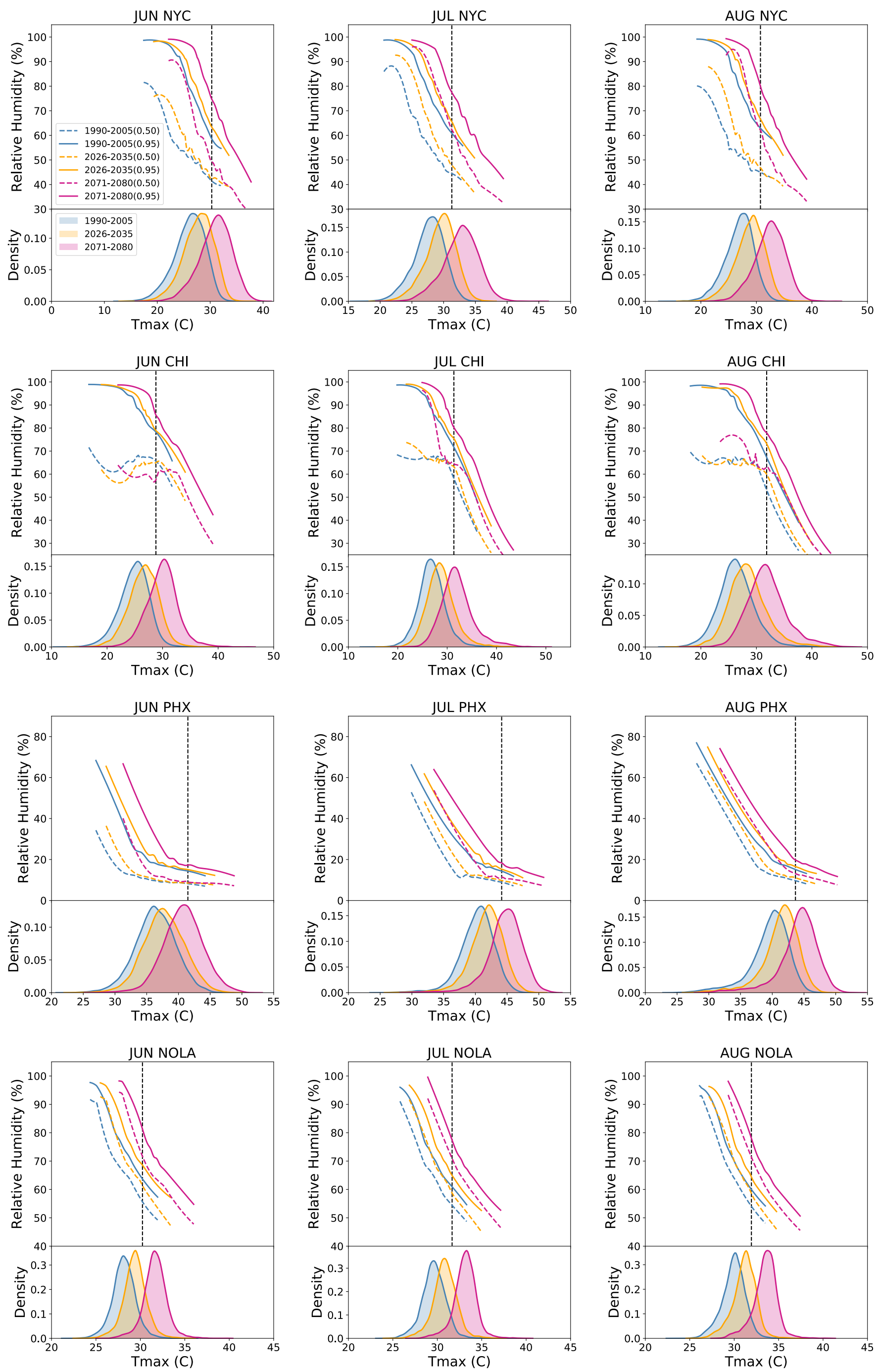


Table 1 Summary of relative humidity (RH), dew point (DP), and heat index (HI) at 0.5 quantile in the period of 1990-2005 (denoted by Hist.) given 0.5 and 0.95 quantiles of *Tmax* in the same period. The changes in the conditional median of RH, DP, and HI from the period of Hist to the period of 2026-2035 (denoted by 2030) and to the period of 2071-2080 (denoted by 2075) given 0.5 and 0.95 quantiles of *Tmax* in the same period are shown in the table, where the 0.5 and 0.95 quantiles of *Tmax* in the period of 2030 and 2075 are also displayed as the deviation from the *Tmax* in the period of Hist. (Units of *Tmax*: °C; units of RH: %; units of dew point: °C; units of HI: °C).

City	Tmax quantiles	Periods	June				July				August			
			Tmax	RH	DP	HI	Tmax	RH	DP	HI	Tmax	RH	DP	HI
NYC	0.5	Hist.	26.3	51.6	15.6	26.9	27.9	55.2	18.1	28.8	27.2	52.3	16.6	27.8
		2030	+1.6	-1.2	+1.1	+1.5	+1.8	-0.3	+1.6	+2.5	+1.9	-3.9	+0.5	+1.8
		2075	+4.9	-6.0	+2.5	+5.3	+5.0	-3.4	+3.6	+7.8	+5.4	-2.5	+4.2	+7.8
	0.95	Hist.	30.3	41.4	15.8	30.3	31.3	44.5	17.8	32.1	30.7	45.0	17.4	31.3
		2030	+1.7	-0.4	+1.3	+2.1	+1.8	-2.2	+0.8	+2.3	+2.0	-2.2	+1.1	+2.8
		2075	+5.3	-8.4	+1.2	+6.1	+5.6	-6.9	+2.3	+8.1	+5.9	-6.6	+2.8	+8.6
CHI	0.5	Hist.	25.1	66.8	18.5	26.1	26.8	67.5	20.3	28.4	26.3	67.0	19.7	27.6
		2030	+1.6	-1.7	+1.1	+1.9	+1.7	-0.8	+1.4	+2.8	+2.0	-2.0	+1.3	+2.8
		2075	+5.0	-5.8	+3.2	+7.1	+4.9	-3.4	+3.8	+9.3	+5.4	-4.3	+4.0	+9.5
	0.95	Hist.	29	65.6	21.7	31.6	31.4	58.3	22.2	35.1	31.9	53.4	21.2	34.8
		2030	+2.0	-3.5	+1.0	+3.3	+2.3	-9.3	-0.6	+2.4	+2.6	-9.1	-0.6	+2.6
		2075	+5.6	-14.4	+1.1	+8.0	+6.8	-21.4	-1.3	+7.2	+6.6	-17.4	-0.5	+7.8
PHX	0.5	Hist.	36.4	9.6	-0.6	33.6	40.5	11.1	4.5	37.7	40.1	12.5	5.8	37.5
		2030	+1.3	-0.1	+0.8	+1.2	+1.5	+0.0	+1.2	+1.5	+1.7	-0.0	+1.3	+1.8
		2075	+4.4	-0.4	+2.8	+4.1	+4.4	-0.5	+2.6	+4.5	+4.5	-0.1	+3.4	+5.1
	0.95	Hist.	41.5	8.2	1.0	38.1	44.2	9.0	4.1	40.8	43.7	9.7	4.9	40.6
		2030	+1.5	+0.3	+1.6	+1.5	+1.5	-0.1	+1.0	+1.5	+1.5	+0.1	+1.2	+1.5
		2075	+4.4	+0.2	+3.5	+4.2	+4.6	-0.3	+2.9	+4.5	+4.7	-0.3	+3.0	+4.8
NOLA	0.5	Hist.	28.2	66.0	21.2	30.4	29.6	64.7	22.3	33.0	30.0	63.9	22.4	33.7
		2030	+1.2	-0.8	+1.0	+2.2	+1.2	-1.3	+0.9	+2.4	+1.2	-1.5	+0.8	+2.4
		2075	+3.5	-2.1	+2.8	+7.1	+3.6	-2.5	+2.8	+8.0	+3.6	-3.5	+2.5	+7.7
	0.95	Hist.	30.2	55.6	20.4	32.4	31.7	54.3	21.3	34.9	31.9	54.2	21.5	35.2
		2030	+1.2	-0.2	+1.1	+2.3	+1.2	-0.5	+0.9	+2.3	+1.1	-1.3	+0.7	+2.1
		2075	+3.5	+0.3	+3.4	+7.6	+3.3	+0.7	+3.2	+7.7	+3.4	-1.3	+2.8	+7.5

As expected, both $Tmax_{0.5}$ and $Tmax_{0.95}$ increase over time (Table 1). The conditional median of RH at the increasing $Tmax_{0.5}$ decreases over time in all cities, except for PHX between July 1990-2005 and 2026-2035, a period over which the conditional median of RH is effectively constant. The same pattern holds most cities at $Tmax_{0.95}$, with the conditional median of RH decreasing over time for NYC, CHI, and August of NOLA, while changing only slightly (absolute changes <1%) in PHX, as well as in June and July for NOLA (Table 1). The conditional 0.95 quantile of RH | $Tmax_{0.5}$ decreases consistently over time in all four cities and all three months (Table 2). The same pattern holds for most cities for RH | $Tmax_{0.95},$ except for June of PHX, in which 0.95 quantile of RH | $Tmax_{0.95}$ shows a slightly increase (+0.24% between 1990-2005 and 2071-2080) (Table 2). The projected decrease in the conditional median quantile of RH at a fixed quantile of *Tmax* indicates that the growth in saturation vapor pressure due to increased temperature is larger than the growth in the vapor pressure of the air. The result is consistent with findings in previous studies (Byrne & O'Gorman, 2018; Flato et al., 2013; Joshi, Gregory, Webb, Sexton, & Johns, 2008; O'Gorman & Muller, 2010), which show a decrease in RH over land in response to a warming climate, as well as with the marginal decrease in mean RH noted early (Figure S7).

Given the increase in $Tmax_{0.5}$ and $Tmax_{0.95}$, the conditional median and 0.95 quantile of heat index both increase in 2026-2035 and 2071-2080 relative to that in 1990-2005 (Table 1 and 2, Fig. 6). The increase in heat index in the two future periods is faster than the increase in $Tmax_{\tau}$ in most cities, except for PHX where the increase in heat index is close to the increase in $Tmax_{\tau}$. The conditional median and 0.95 quantiles of dew point at both quantiles of *Tmax* increase over time except for July and August in CHI, where the decrease in RH is large enough that there is a minimal change in conditional dew point (Table 1 and 2, Fig. 7). This result suggests that,

Table 2 Same as Table 1 except for relative humidity (RH), dew point (DP), and heat index (HI) at the conditional 0.95 quantile.

City	Tmax quantiles	Periods	June				July				August			
			Tmax	RH	DP	HI	Tmax	RH	DP	HI	Tmax	RH	DP	HI
NYC	0.5	Hist.	26.3	78.4	22.3	28.2	27.9	76.1	23.3	31.3	27.2	78.2	23.1	30.0
		2030	+1.6	-2.4	+1.0	+3.1	+1.8	-3.5	+0.9	+3.5	+1.9	-4.3	+0.9	+3.7
		2075	+4.9	-8.3	+2.9	+10.0	+5.0	-8.3	+2.9	+11.0	+5.4	-9.8	+3.0	+11.7
	0.95	Hist.	30.3	58.5	21.3	33.2	31.3	60.8	22.8	35.7	30.7	63.0	22.8	35.0
		2030	+1.7	-0.9	+1.2	+3.0	+1.8	-3.8	+0.6	+2.9	+2.0	-3.9	+0.9	+3.8
		2075	+5.3	-8.8	+2.2	+8.9	+5.6	-10.7	+2.0	+9.9	+5.9	-12.7	+1.8	+10.0
CHI	0.5	Hist.	25.1	90.3	23.4	25.9	26.8	88.5	24.8	30.1	26.3	91.0	24.7	28.9
		2030	+1.6	-2.6	+1.0	+3.7	+1.7	-3.4	+1.0	+4.3	+2.0	-5.7	+0.8	+4.7
		2075	+5.0	-10.4	+2.8	+11.9	+4.9	-9.5	+2.9	+12.8	+5.4	-12.7	+2.7	+13.6
	0.95	Hist.	28.8	78.2	24.6	33.9	31.4	71.9	25.6	39.1	31.9	67.5	25.0	39.0
		2030	+2.0	-4.6	+0.9	+4.2	+2.3	-8.5	+0.1	+3.7	+2.6	-10.4	-0.4	+3.1
		2075	+5.6	-13.8	+2.1	+11.4	+6.8	-22.6	-0.0	+9.4	+6.6	-22.6	-0.7	+8.0
PHX	0.5	Hist.	36.4	18.1	8.4	34.4	40.5	18.3	11.8	39.6	40.1	19.6	12.5	39.4
		2030	+1.3	-0.1	+1.0	+1.5	+1.5	-0.2	+1.0	+2.0	+1.7	-0.1	+1.3	+2.4
		2075	+4.4	-0.8	+2.9	+5.3	+4.4	-1.7	+2.0	+5.5	+4.5	-1.3	+2.6	+6.2
	0.95	Hist.	41.5	14.4	9.0	39.6	44.2	14.4	11.0	43.0	43.7	15.1	11.4	42.7
		2030	+1.5	-0.3	+0.8	+1.8	+1.5	-0.7	+0.4	+1.7	+1.5	-0.3	+0.8	+1.8
		2075	+4.4	+0.2	+3.7	+5.8	+4.6	-1.1	+2.3	+5.6	+4.7	-1.3	+2.3	+5.8
NOLA	0.5	Hist.	28.2	74.3	23.1	31.6	29.6	71.0	23.8	34.3	30.0	70.0	23.9	35.0
		2030	+1.2	-2.1	+0.7	+2.5	+1.2	-1.3	+0.9	+2.8	+1.2	-1.9	+0.7	+2.7
		2075	+3.5	-3.2	+2.7	+8.4	+3.6	-3.4	+2.6	+8.8	+3.6	-4.0	+2.5	+8.6
	0.95	Hist.	30.2	63.7	22.6	34.1	31.7	60.8	23.2	36.6	31.9	60.1	23.2	36.9
		2030	+1.2	-0.6	+1.0	+2.7	+1.2	-1.3	+0.7	+2.4	+1.1	-1.3	+0.7	+2.4

despite a modest decrease in mean RH, heat stress impacts in a warming climate will increase faster than temperatures alone would indicate in many locations.

At a fixed value of *Tmax* (e.g., the 0.95 quantile of *Tmax* during 1990-2005), the conditional 0.95 quantiles of RH increase over in all three month and all cities (upper panels of Fig. 5, and Table 3). Changes in the conditional median quantiles of RH show similar pattern, except for June in CHI, where the conditional medians of RH decrease (Table 4).

Table 3 Summary of relative humidity (RH), dew point (DP), and heat index (HI) at the conditional 0.95 quantile in the period of 1990-2005 (denoted by Hist.), conditional upon the 0.95 quantile of *Tmax* during the Hist period. Changes in RH at 0.95 quantile in the period of 2026-2030 (denoted by 2030) and the period of 2071-2080 (denoted by 2075) deviating from the Hist period, as well as changes in 0.95 quantile of DP and HI converted by the RH and *Tmax* are shown in the corresponding rows (Units of *Tmax*: °C; units of changes in RH: %; units of changes in dew point: °C; units of changes in HI: °C).

City	Period	June				July				August			
		Tmax	RH	DP	HI	Tmax	RH	DP	HI	Tmax	RH	DP	HI
	Hist		58.5	21.3	33.2		60.8	22.8	35.7		63.0	22.8	35.0
NYC	2030	30.29	+4.8	+1.3	+1.0	31.29	+4.9	+1.2	+1.3	30.73	+3.8	+1.0	+1.0
	2075		+16.1	+4.0	+4.0		+16.6	+4.1	+5.3		+16.8	+4.0	+5.0
	Hist		78.2	24.6	33.9		71.9	25.6	39.1		67.5	25.0	39.0
CHI	2030	28.8	+0.7	+0.1	+0.0	31.37	+3.7	+0.8	+1.2	31.85	+5.8	+1.4	+2.1
	2075		+7.6	+1.5	+1.4		+8.2	+1.9	+3.2		+10.4	+2.5	+4.0
	Hist		14.4	9.0	39.6		14.4	11.0	43.0		15.1	11.4	42.7
PHX	2030	41.48	+0.7	+0.7	+0.2	44.15	+0.9	+1.0	+0.5	43.65	+1.5	+1.4	+0.6
	2075		+2.9	+2.8	+0.9		+4.0	+3.8	+2.0		+4.7	+4.1	+2.2
	Hist		63.7	22.6	34.1		60.8	23.2	36.6		60.1	23.2	36.9
NOLA	2030	30.23	+5.1	+1.2	+1.1	31.67	+4.2	+1.1	+1.2	31.94	+4.1	+1.2	+1.4
	2075		+18.0	+4.1	+4.7		+16.9	+4.1	+5.5		+18.0	+4.4	+6.5

High conditional quantiles of heat index and dew point similarly display large increases over time in the four cities (Table 3, Fig. 6 and Fig. 7). Conditional medians of dew point and heat index given *Tmax* at the historic 0.95 quantile increase over time except for June of Chicago, where both conditional dew point and heat index are decrease due to the large reduction of RH (Table 4, Fig. 6 and 7). Many previous impact analysis on heat waves (Carleton et al., 2019; Dosio et al., 2018; e.g. Mazdiyasni et al., 2017) only considered increases in extreme temperature. Our results suggest that, at a fixed extreme temperature, increase in both median and high quantiles of RH due to warming climate will increase the health impacts of heat extremes in future days. In other words, a day of a given temperature will be more impactful in a warmer climate than a day of the same temperature in the current climate.

Table 4 Same as Table 3 except for relative humidity (RH), dew point (DP), and heat index (HI) at the conditional median quantile.

City	Period	June				July				August			
		Tmax	RH	DP	HI	Tmax	RH	DP	HI	Tmax	RH	DP	HI
NYC	Hist	30.29	41.4	15.8	30.3	31.29	44.5	17.8	32.1	30.73	44.5	17.8	32.1
	2030		+0.9	+0.3	+0.1		+3.4	+1.1	+0.6		+3.4	+1.1	+0.6
	2075		+7.9	+2.8	+1.2		+17.7	+5.5	+4.1		+17.7	+5.5	+4.1
CHI	Hist	28.8	65.6	21.7	31.6	31.37	58.3	22.2	35.1	31.85	58.3	22.2	35.1
	2030		-0.7	-0.2	-0.2		+5.1	+1.4	+1.3		+5.1	+1.4	+1.3
	2075		-8.2	-2.2	-1.3		+6.2	+1.8	+1.9		+6.2	+1.8	+1.9
PHX	Hist	41.48	8.2	1.0	38.1	44.15	9.0	4.1	40.8	43.65	9.0	4.1	40.8
	2030		+0.3	+0.5	+0.1		+0.8	+1.2	+0.3		+0.8	+1.2	+0.3
	2075		+0.9	+1.4	+0.2		+2.4	+3.5	+1.0		+2.4	+3.5	+1.0
NOLA	Hist	30.23	55.6	20.4	32.4	31.67	54.3	21.3	34.9	31.94	54.3	21.3	34.9
	2030		+6.1	+1.6	+1.1		+4.5	+1.3	+1.1		+4.5	+1.3	+1.1
	2075		+16.3	+4.2	+3.6		+16.9	+4.4	+4.9		+16.9	+4.4	+4.9

7. Discussions and Conclusions

In this study, we use a large ensemble of simulations to investigate the joint distribution of summertime RH and *Tmax* for three time periods (1990-2005, 2026-2035, 2071-2080) in four U.S. cities that represent a range of climates. For each month in a city, the joint distribution changes shape and shifts toward higher dew point values over time. A cutoff in the highest *Tmax* for which RH of 100% can occur, followed by a steep drop in the maximum attainable RH as *Tmax* increases beyond this cutoff is observed in the joint distribution in all months for NYC and CHI, and some months for NOLA (Fig. 1). The cutoff shifts toward higher dew point temperature as climate warms.

Based on the information provided by the joint distribution diagnostics, we developed statistical models to capture the conditional distribution of RH | *Tmax* using quantile regression, where a kink function is used to capture the kink, and a number of cubic spline basis functions are used to describe smooth variation in RH quantiles with *Tmax* within a month.

Figure 6. Estimated quantiles of heat index given daily maximum temperature (*Tmax*) in June, July, and August of three periods at New York (NYC), Chicago (CHI), Phoenix (PHX), and New Orleans (NOLA). Shadings represent the central 95% range of the distribution. Blue denotes the period of 1990-2005, green 2016-2035, red 2071-2080. Heavy lines are median of the distribution. The horizontal dashed lines denote the thresholds for the categories defined by national weather service.

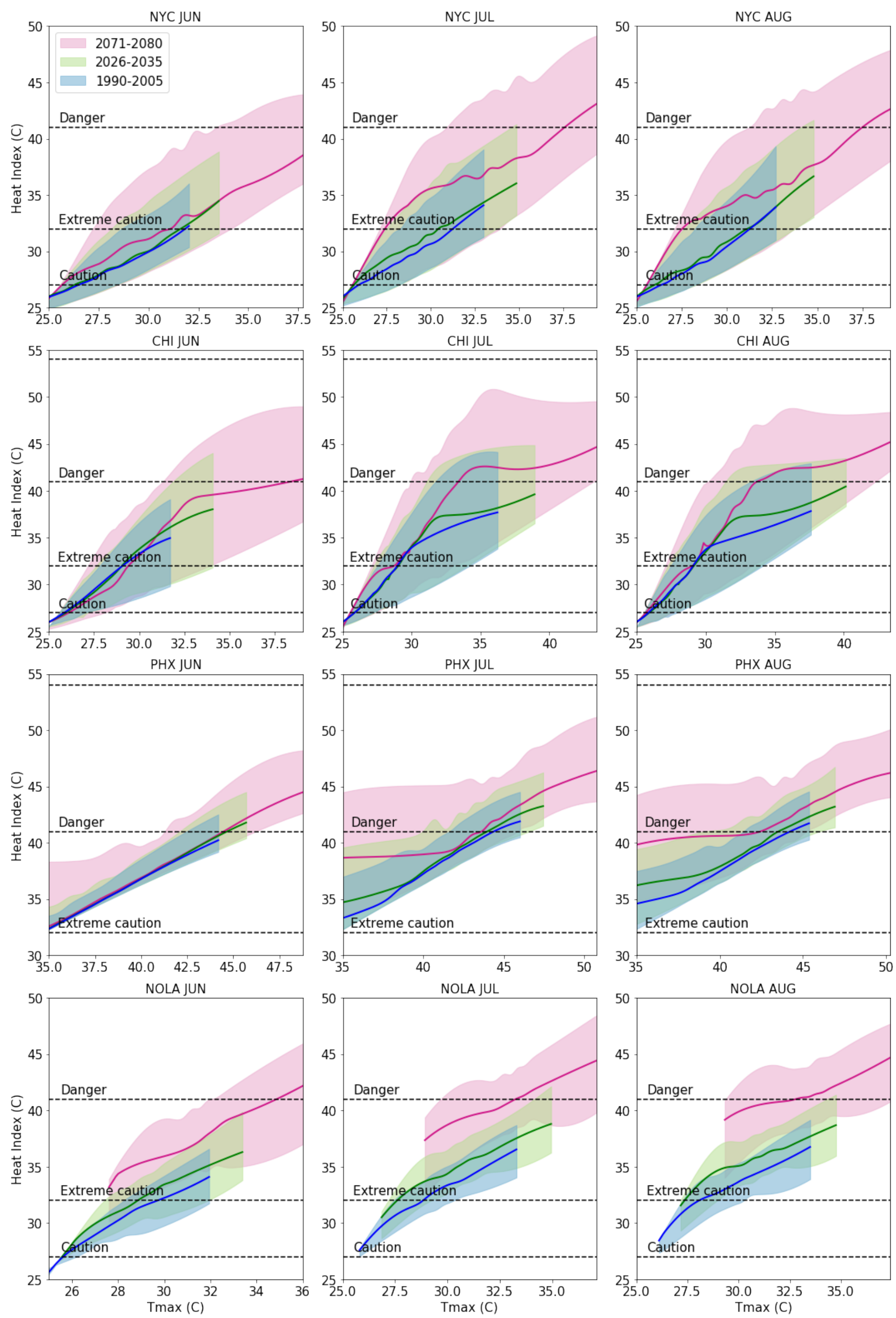
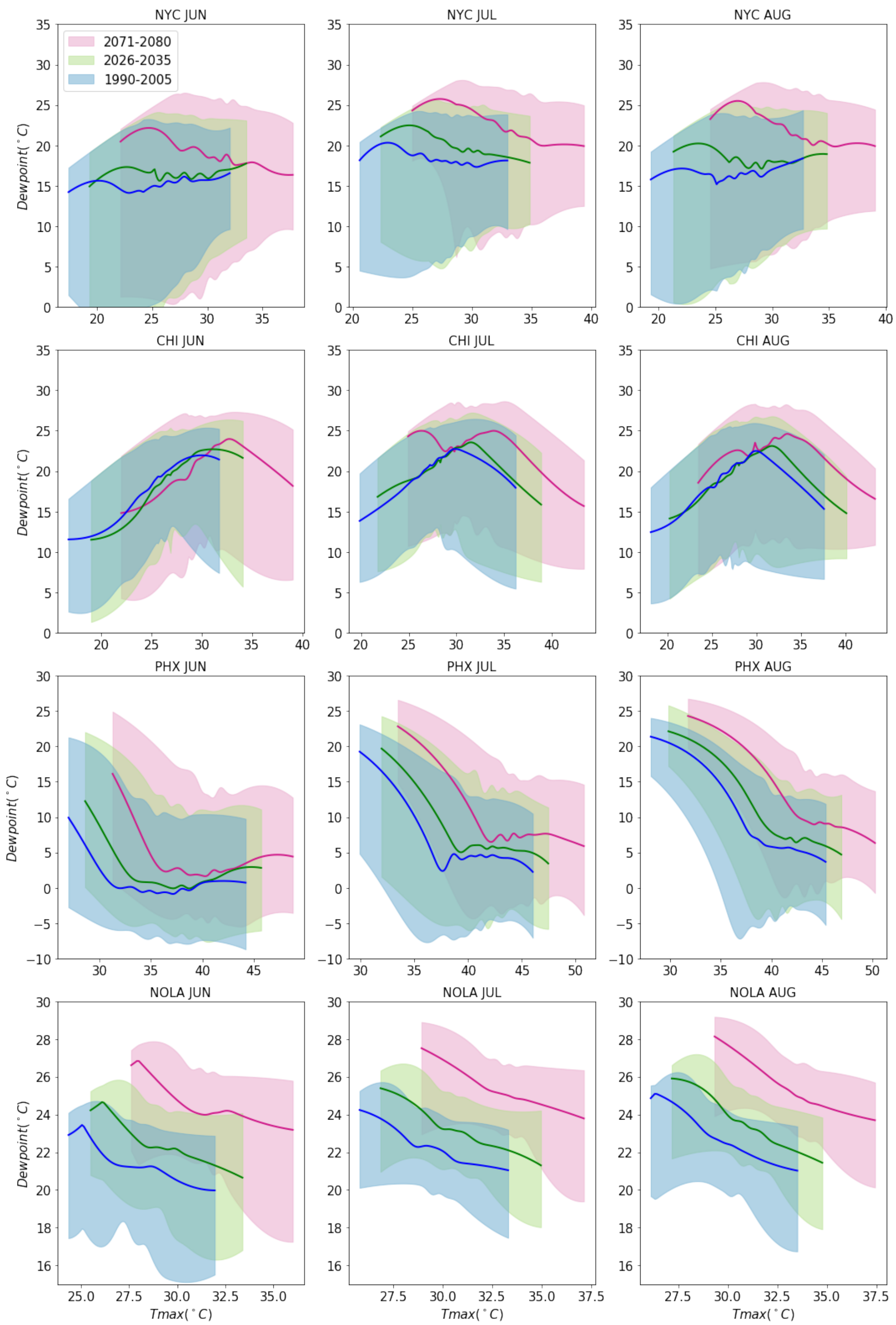


Figure 7. Estimated quantiles of dew point given daily maximum temperature (*Tmax*) in June, July, and August of three periods at New York (NYC), Chicago (CHI), Phoenix (PHX), and New Orleans (NOLA). Shadings represent the central 95% range of the conditional distribution. Blue denotes the period of 1990-2005, green 2016-2035, red 2071-2080. Heavy lines are median of the distribution.



The quality of fit diagnostic indicates the distribution of RH | *Tmax* estimated by these quantile regression models fit the data of LENS well (Fig. S4-S6). In addition, the quantile regression models could estimate the distribution and tails of conditional distribution of RH | *Tmax* without any parametric assumptions.

The conditional quantiles of RH | *Tmax* allow us to investigate the changes in heat extremes in multiple ways. First, we investigate the changes in RH, heat index and dew point given a fixed quantile of *Tmax* during any of the three periods. As expected, both $Tmax_{0.5}$ and $Tmax_{0.95}$ increase over time. At the $Tmax_{0.5}$ and $Tmax_{0.95}$ during future periods, the conditional quantiles of heat index or dew point are generally higher than that during the historical period, even though the conditional quantiles of RH are lower (Table 1 and 2). These results suggest that, despite a modest decrease in relative humidity, heat stress impacts in a warming climate will tend to increase faster than temperatures alone would indicate.

Second, we investigate the changes in RH, heat index and dew point given a fixed *Tmax*. Consider *Tmax* at its historical 0.95 quantile, for instance, the conditional 0.95 quantiles of RH, dew point and heat index increase from 1990-2005 to 2071-2080 in all four cities (Table 3). The increase pattern holds in the conditional median quantiles of these three variables given the fixed *Tmax*, except for June in CHI (Table 4). Our results indicate that, in a warming climate, a future day will tend to have higher RH than a day of the same temperature under the historic climate. Therefore, even at the same temperature, the increase in RH will increase the impact of heat extremes in the future day. Ignoring this conditional increase in RH, as many previous studies have done in assessing the heat impact (Carleton et al., 2019; Dosio et al., 2018; e.g. Mazdiyasni et al., 2017), may lead to underestimating impact of heat waves in a warmer world.

This study gives us confidence about applying quantile regression models to quantify the conditional distribution of RH | *Tmax* in different climate background. For a specific city, although these statistical models capture the variability of RH given *Tmax* for each summer month and each time period, respectively, we see a need for a uniform model for all days of the year and all the time period that could capture the seasonal variability and long-term evolution with the same set of parameters. To reach this end, we have to include more terms in the statistical model to capture the variability of RH with time (e.g. days of the year, and years) and the interaction between the *Tmax* and time. These statistical models developed based on quantile regression approaches could be eventually applied to estimating the future climate projections that require less computing resources than the climate models (Haugen et al., 2019).

Appendix

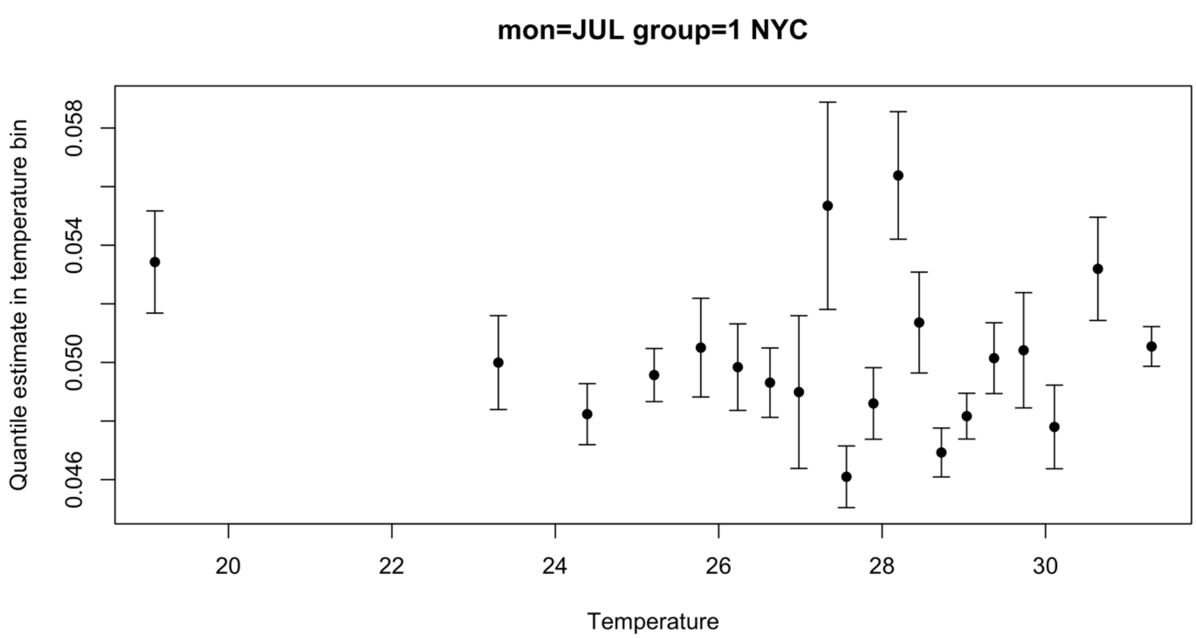
a. Selecting the temperature at the kink

As discussed in the Section 3, the cutoff indicates the moisture in the atmosphere is largely constrained by the dew point. Therefore, we select the T_0 value through examining the density distribution of dew point converted by corresponding RH and *Tmax* (Fig. 2). When there is a kink of the joint distribution, the distribution of dew point displays a cliff-like shape at the high end of temperature (e.g. NYC, CHI, and some months of NOLA). There is no such feature observed in the density distribution of dew point in PHX. Therefore, we use the following rule for deciding whether to include a kink function: First, we check if the differences between 0.99 quantile and 0.9 quantile of dew point is < 9% of differences between maximum and minimum values of dew point. If it is, then we include a kink function, and so must obtain a value for T_0, which we set to the empirical 0.999 quantile of dew point for that month and period. Note that this method is only feasible because of the use of a large ensemble of simulations. The selected T_0 values are listed in Table S1 in supplement information.

b. Selecting number of basis functions

The number of basis functions (m) in the model is an important parameter that influences how well the model fits the observations. Increasing m improves the quality of fit of the model. However, if m is too large, the model runs the risk of overfitting the data, which can lead to diminished performance in assessments with out of sample data. Here we use a cross-validation metric to select the simplest model that provides overall good estimation of quantiles and prevent overfitting the data.

Figure A1. $\widehat{\boldsymbol{S}}_{\boldsymbol{test}}(\boldsymbol{a_j}, \boldsymbol{\tau})$ in July of New York during period of 1990-2005 for the quantile $\boldsymbol{\tau} = \mathbf{0.95}$ estimation. 20 temperature bins are edged by equally spaced quantiles (0.05) from 0 to 1 of the *Tmax* distribution in July. The black dot and error bars indicate the mean and standard deviation in a temperature bin.



To estimate the appropriate number of basis functions used in a quantile-regression model, we apply a cross validation on the model. We extract samples by randomly selecting 34 members from the 35 members of LENS and drop one member without replacement. By this way, we obtained 35 samples. Then, we apply quantile regression on each sample and calculate the fraction of RH events that exceeds a particular quantile τ given *Tmax* bins

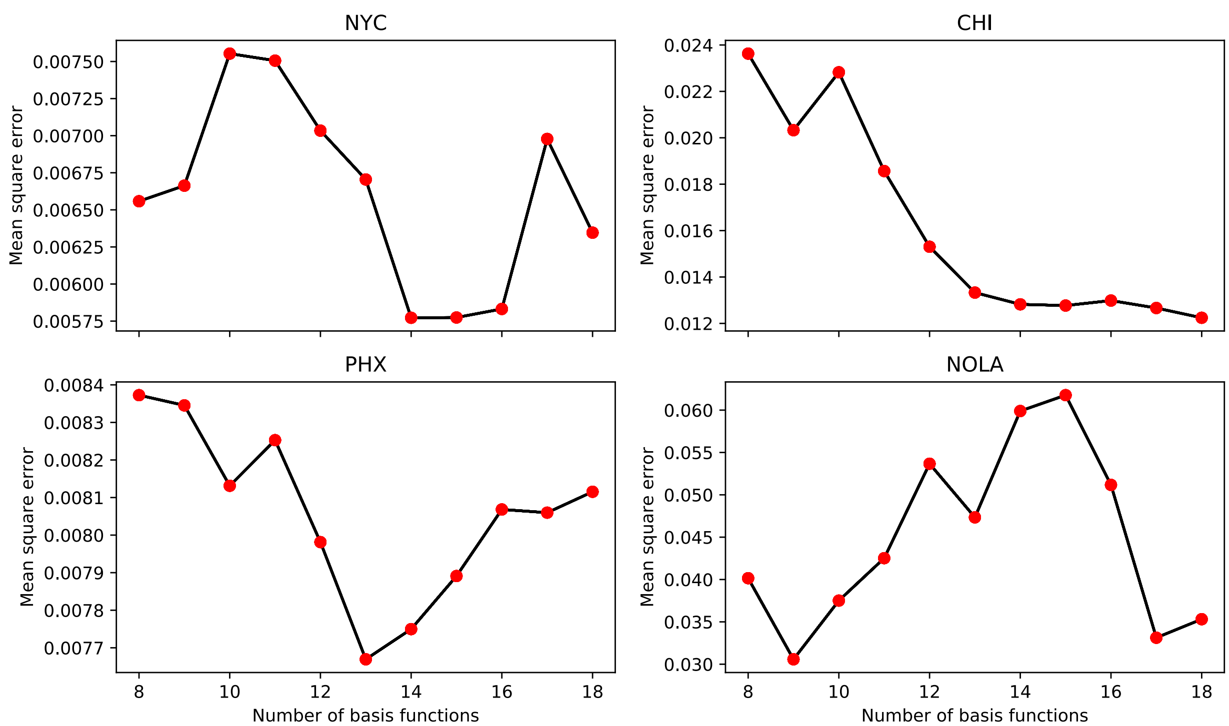
$$\hat{S}_{test}(a_j, \tau) \ = \frac{1}{n}\sum_{i=1}^{n} I[(RH_i(a_j) \ - \widehat{RH}_{i,\tau}(a_j)) > 0] \tag{9}$$

where n = 35 is the number of samples. a_j represent temperature bins whose boundaries are defined by equally spaced quantiles (0.05) from 0 to 1 of the *Tmax* distribution in a month. The RH_i represents the observed values from model output, and $\widehat{RH}_{i,\tau}$ is the estimated value at quantile τ. *I* is the indicator function. An appropriate model which is fit to the data requires the estimated quantiles to contain approximately the desired fraction of positive and negative residuals. Therefore, we seek an appropriate model to satisfy $\hat{S}_{test}(a_j, \tau) \approx (1 - \tau)$. Figure A1 show an example of $\hat{S}_{test}(a_j, \tau)$ in July of NYC during period of 1990-2005 for the 0.95 quantile estimation. As we expected the $\hat{S}_{test}(a_j, \tau)$ at each temperature bin is generally close to 0.05. The mean square error between $\hat{S}_{test}(a_j, \tau)$ and 0.05 is used to measure the variability of $\hat{S}_{test}(a_j, \tau)$. As the model complexity increases with the growth of *m*, the variability of $\hat{S}_{test}(a_j, \tau)$ should decrease until reaching a minimum when *m* reaches an optimal number of basis functions. Once *m* exceeds this point, the model starts to overfit the data and the mean square error between $\hat{S}_{test}(a_j, \tau)$ and 0.05 will grow. For each city, we sum up the mean square error across three months (June, July, and August) and three time periods:

$$CV \ = \frac{1}{3}\sum_{g=1}^{3}\frac{1}{5}\sum_{mon=1}^{5}\frac{1}{k}\sum_{a_j=1}^{k}(\hat{S}_{test}(a_j, \tau)_{g,mon} - 0.05)^2 \tag{11}$$

where g represents 3 time periods, and *mon* represents 3 months. CV is used to quantify the averaged variability of $\hat{S}_{test}(a_j)$ for a city (Fig. A2). Based on Fig. A2, the selected number of cubic-spline basis functions are 14 for NYC, 13 for CHI, 13 for PHX, and 9 for NOLA.

Figure A2. Averaged mean square error between $\boldsymbol{\hat{S}_{test}(a_j, \tau)}$ and 0.05 at estimated quantile $\boldsymbol{\tau = 0.95}$ across 3 months and 3 periods for New York (NYC), Chicago (CHI), Phoenix (PHX), and New Orleans (NOLA). X-axis is the number of cubic-spline basis functions used in the statistic model.



Acknowledgement

We would like to thank Anthony J. Broccoli, who provided inspiring discussion with Jiacan Yuan. This work was initiated from a class project. We would like to thank Joshua Couper, Nazia Arbab, Jessica August, and Sadiya Bab Tijjani for participating this project in the class of

Statistics in Earth Science. The authors used the code developed by Joshua Couper to extract *Tmax* and corresponding RH from the LENS data. We acknowledge CESM Large Ensemble Community Project for providing us the LENS data. We also thank the Met Office Hadley Centre for providing HadISD data. Jiacan Yuan and Robert E. Kopp were supported by Rhodium Group as part of the Climate Impact Lab consortium. Robert E. Kopp has in the past served as a consultant to Rhodium Group. Michael L. Stein was supported by U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR) under Contract DE-AC02-06CH11357. The result data of this study are available in the repository https://doi.org/10.6084/m9.figshare.11283233.v1

Reference

Alduchov, O. A., & Eskridge, R. E. (1996). Improved Magnus Form Approximation of Saturation Vapor Pressure. *Journal of Applied Meteorology*, *35*(4), 601–609. https://doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., … Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, *259*(4), 660–684. https://doi.org/10.1016/j.foreco.2009.09.001

Basu, R., & Samet, J. M. (2002). Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence. *Epidemiologic Reviews*, *24*(2), 190–202. https://doi.org/10.1093/epirev/mxf007

Byrne, M. P., & O'Gorman, P. A. (2018). Trends in continental temperature and humidity directly linked to ocean warming. *Proceedings of the National Academy of Sciences*, *115*(19), 4863–4868. https://doi.org/10.1073/pnas.1722312115

Carleton, T., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., … Zhang, A. T. (2019). *Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits* (SSRN Scholarly Paper No. ID 3224365). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=3224365

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., … Wehner, M. (2013). Long-term climate change: Projections, commitments and irreversibility. In IPCC (Ed.), *Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5*. Retrieved from http://www.climatechange2013.org/images/report/WG1AR5_Chapter12_FINAL.pdf

Dai, A. (2006). Recent Climatology, Variability, and Trends in Global Surface Humidity. *Journal of Climate*, *19*(15), 3589–3606. https://doi.org/10.1175/JCLI3816.1

Dosio, A., Mentaschi, L., Fischer, E. M., & Wyser, K. (2018). Extreme heat waves under 1.5\hspace0.167em°C and 2\hspace0.167em°C global warming. *Environmental Research Letters*, *13*(5), 054006. https://doi.org/10.1088/1748-9326/aab827

Dunn, R. J. H., Willett, K. M., Thorne, P. W., Woolley, E. V., Durre, I., Dai, A., … Vose, R. S. (2012). HadISD: A quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011. *Clim. Past*, *8*(5), 1649–1679. https://doi.org/10.5194/cp-8-1649-2012

Fischer, E. M., & Knutti, R. (2013). Robust projections of combined humidity and temperature extremes. *Nature Climate Change*, *3*(2), 126–130. https://doi.org/10.1038/nclimate1682

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., … Rummukainen, M. (2013). Evaluation of Climate Models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Retrieved from https://www.ipcc.ch/report/ar5/wg1/evaluation-of-climate-models/

Fontana, G., Toreti, A., Ceglar, A., & De Sanctis, G. (2015). Early heat waves over Italy and their impacts on durum wheat yields. *Natural Hazards and Earth System Sciences*, *15*(7), 1631–1637. https://doi.org/10.5194/nhess-15-1631-2015

Gibbins, C. J. (1990). A survey and comparison of relationships for the determination of the saturation vapour pressure over plane surfaces of pure water and of pure ice. | Article Information | J-GLOBAL. *Annales Geophysicae (Berlin)*, *8*(12), 859–885.

Haugen, M. A., Stein, M. L., Moyer, E. J., & Sriver, R. L. (2018). Estimating Changes in Temperature Distributions in a Large Ensemble of Climate Simulations Using Quantile Regression. *Journal of Climate*, *31*(20), 8573–8588. https://doi.org/10.1175/JCLI-D-17-0782.1

Haugen, M. A., Stein, M. L., Sriver, R. L., & Moyer, E. J. (2019). Future climate emulations using quantile regressions on large ensembles. *Advances in Statistical Climatology, Meteorology and Oceanography*, *5*(1), 37–55. https://doi.org/10.5194/ascmo-5-37-2019

Joshi, M. M., Gregory, J. M., Webb, M. J., Sexton, D. M. H., & Johns, T. C. (2008). Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Climate Dynamics*, *30*(5), 455–465. https://doi.org/10.1007/s00382-007-0306-1

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., … Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community

Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, *119*(2), 345–357. https://doi.org/10.1007/s10584-013-0705-8

Kodra, E., & Ganguly, A. R. (2014). Asymmetry of projected increases in extreme temperature distributions. *Scientific Reports*, *4*, 5884. https://doi.org/10.1038/srep05884

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

Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, *15*(4), 143–156. https://doi.org/10.1257/jep.15.4.143

Lawrence, M. G. (2005). *The relationship between relative humidity and the dewpoint temperature in moist air—A simple conversion and applications*. https://doi.org/10.1175/BAMS-86-2-225

Mazdiyasni, O., AghaKouchak, A., Davis, S. J., Madadgar, S., Mehran, A., Ragno, E., … Niknejad, M. (2017). Increasing probability of mortality during Indian heat waves. *Science Advances*, *3*(6), e1700066. https://doi.org/10.1126/sciadv.1700066

Meehl, G. A., & Tebaldi, C. (2004). More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century. *Science*, *305*(5686), 994–997. https://doi.org/10.1126/science.1098704

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., … Trauernicht, C. (2017). Global risk of deadly heat. *Nature Climate Change*, *7*(7), 501–506. https://doi.org/10.1038/nclimate3322

NWS Pueblo, CO. (2011, June 29). Retrieved October 18, 2019, from https://web.archive.org/web/20110629041320/http:/www.crh.noaa.gov/pub/heat.php

O'Gorman, P. A., & Muller, C. J. (2010). How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations? *Environmental Research Letters*, *5*(2), 025207. https://doi.org/10.1088/1748-9326/5/2/025207

Orlowsky, B., & Seneviratne, S. I. (2012). Global changes in extreme events: Regional and seasonal dimension. *Climatic Change*, *110*(3), 669–696. https://doi.org/10.1007/s10584-011-0122-9

Papalexiou, S. M., AghaKouchak, A., Trenberth, K. E., & Foufoula-Georgiou, E. (2018). Global, Regional, and Megacity Trends in the Highest Temperature of the Year: Diagnostics and Evidence for Accelerating Trends. *Earth's Future*, *6*(1), 71–79. https://doi.org/10.1002/2017EF000709

Ramamurthy, P., Li, D., & Bou-Zeid, E. (2017). High-resolution simulation of heatwave events in New York City. *Theoretical and Applied Climatology*, *128*(1), 89–102. https://doi.org/10.1007/s00704-015-1703-8

Rothfusz, L. P. (1990). The Heat Index "Equation" (or, More Than You Ever Wanted to Know About Heat Index). *Tech. Attachment*, *1*, 1–2. https://doi.org/SR/SSD 90-23

Russo, S., Sillmann, J., & Sterl, A. (2017). Humid heat waves at different warming levels. *Scientific Reports*, *7*(1), 1–7. https://doi.org/10.1038/s41598-017-07536-7

Schwartzman, P. D., Michaels, P. J., & Knappenberger, P. C. (1998). Observed changes in the diurnal dewpoint cycles across North America. *Geophysical Research Letters*, *25*(13), 2265–2268. https://doi.org/10.1029/98GL01843

Sherwood, S. C., Roca, R., Weckwerth, T. M., & Andronova, N. G. (2010). Tropospheric water vapor, convection, and climate. *Reviews of Geophysics*, *48*(2). https://doi.org/10.1029/2009RG000301

Sriver, R. L., Forest, C. E., & Keller, K. (2015). Effects of initial conditions uncertainty on regional climate variability: An analysis using a low-resolution CESM ensemble. *Geophysical Research Letters*, *42*(13), 2015GL064546. https://doi.org/10.1002/2015GL064546

Steadman, R. G. (1979). The Assessment of Sultriness. Part I: A Temperature-Humidity Index Based on Human Physiology and Clothing Science. *Journal of Applied Meteorology*, *18*(7), 861–873. https://doi.org/10.1175/1520-0450(1979)018<0861:TAOSPI>2.0.CO;2

Stocker, Thomas. F., Qin, D., Plattner, G.-K., Tignor, M. M. B., Allen, S. K., Boschung, J., … Midgley, P. M. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern- mental Panel on Climate Change*. Retrieved from http://www.ipcc.ch/report/ar5/wg1/

Subin, Z. M., Riley, W. J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems*, *4*, M02001. https://doi.org/10.1029/2011MS000072

Sutton, R. T., Dong, B., & Gregory, J. M. (2007). Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophysical Research Letters*, *34*(2), L02701. https://doi.org/10.1029/2006GL028164

Tebaldi, C., Hayhoe, K., Arblaster, J. M., & Meehl, G. A. (2006). Going to the Extremes. *Climatic Change*, *79*(3–4), 185–211. https://doi.org/10.1007/s10584-006-9051-4 van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., … Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change*, *109*(1), 5. https://doi.org/10.1007/s10584-011-0148-z

Willett, K. M., Dunn, R. J. H., Thorne, P. W., Bell, S., de Podesta, M., Parker, D. E., … Williams Jr., C. N. (2014). HadISDH land surface multi-variable humidity and temperature record for climate monitoring. *Climate of the Past*, *10*(6), 1983–2006. https://doi.org/10.5194/cp-10-1983-2014

Zuo, J., Pullen, S., Palmer, J., Bennetts, H., Chileshe, N., & Ma, T. (2015). Impacts of heat waves and corresponding measures: A review. *Journal of Cleaner Production*, *92*, 1–12. https://doi.org/10.1016/j.jclepro.2014.12.078