

Anthony G. Barnston, a Bradfield Lyon, b.c Ethan D. Coffel, d.e and Radley M. Horton

^a International Research Institute for Climate and Society, Columbia University, Palisades, New York; ^b Climate Change Institute, University of Maine, Orono, Maine; ^c School of Earth and Climate Sciences, University of Maine, Orono, Maine; ^d Department of Geography, Dartmouth College, Hanover, New Hampshire; ^e Neukom Institute for Computational Science, Dartmouth College, Hanover, New Hampshire; ^f Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

(Manuscript received 5 December 2019, in final form 14 July 2020)

ABSTRACT: The frequency of heat waves (defined as daily temperature exceeding the local 90th percentile for at least three consecutive days) during summer in the United States is examined for daily maximum and minimum temperature and maximum apparent temperature, in recent observations and in 10 CMIP5 models for recent past and future. The annual average percentage of days participating in a heat wave varied between approximately 2% and 10% in observations and in the model's historical simulations during 1979-2005. Applying today's temperature thresholds to future projections, heatwave frequencies rise to more than 20% by 2035-40. However, given the models' slight overestimation of frequencies and positive trend rates during 1979–2005, these projected heat-wave frequencies should be regarded cautiously. The models' overestimations may be associated with their higher daily autocorrelation than is found in observations. Heat-wave frequencies defined using apparent temperature, reflecting both temperature and atmospheric moisture, are projected to increase at a slightly (and statistically significantly) faster rate than for temperature alone. Analyses show little or no changes in the day-to-day variability or persistence (autocorrelation) of extreme temperature between recent past and future, indicating that the future heat-wave frequency will be due predominantly to increases in standardized (using historical period statistics) mean temperature and moisture content, adjusted by the local climatological daily autocorrelation. Using nonparametric methods, the average level and spatial pattern of future heat-wave frequency is shown to be approximately predictable on the basis of only projected mean temperature increases and local autocorrelation. These model-projected changes, even if only approximate, would impact infrastructure, ecology, and human well-being.

KEYWORDS: Atmosphere; North America; Climate prediction; Climate models; Trends

1. Introduction

In recent decades, much research has focused on the frequency, intensity, and duration of heat waves, especially in the context of a warming climate (e.g., Vose et al. 2017; Hansen et al. 2012). Heat waves in summer in the United States have been found to be related to specific atmospheric wave patterns (e.g., Teng et al. 2013; Grotjahn et al. 2016), land surface conditions (Durre et al. 2000), or oceanic temperature patterns (McKinnon et al. 2016).

Many heat-wave definitions and associated warning systems exist, often reflecting what local officials consider best serves the public (Sheridan and Kalkstein 2004). For a given location and time of year, heat waves are often defined as a set of consecutive days with temperature exceeding a defined percentile threshold relative to a reference climatological distribution (Perkins and Alexander 2013). Examples requiring

Openotes content that is immediately available upon publication as open access.

© Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAMC-D-19-0291.s1.

Corresponding author: Anthony G. Barnston, tonyb@iri.columbia.edu

various thresholds and consecutive day durations are numerous, for example, Oswald and Rood (2014), Anderson and Bell (2009), Smith et al. (2013), and Tan et al. (2007), among others. Some definitions require only a single heat-wave day (e.g., Alexander et al. 2006). Other definitions include the spatial extent of extreme temperature (e.g., McKinnon et al. 2016; Lyon et al. 2019), or the integrated magnitude of an event (e.g., Russo et al. 2014). The variation of the spatial distribution of heat-wave activity across the United States as a function of the percentile threshold and duration requirement was examined in Lyon and Barnston (2017) (hereinafter LB17). A review of the effects of rising temperature on extreme heat events under various definitions is discussed in Horton et al. (2016).

In this study, the role of day-to-day autocorrelation of the temperature—a feature that has a stable, climatologically determined spatial pattern across the United States for a given time of year—is highlighted as a key indicator of the expected frequency of heat-wave occurrence. Further, the autocorrelation becomes a modulating factor in the prediction of future heat-wave rates, given projected increases in temperature. In this introduction, we now discuss some preliminary analyses to set the stage for the study's main questions and the analyses to address them.

¹ Alexander et al. (2006) also considered other heat-wave definitions.

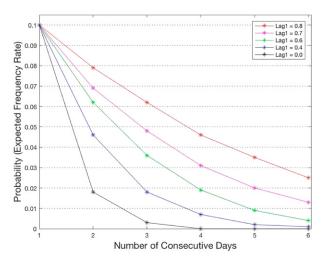


FIG. 1. Probability of 1–6 consecutive days of 90th-percentile temperature, for values of daily autocorrelation of 0, 0.4, 0.6, 0.7 and 0.8. The lines show probabilities for the counting method when each day in an n-day heat wave is counted as an n-day event, as done in this study. Note that the 0.018 and 0.003 probabilities shown for two and three consecutive days, respectively, for zero autocorrelation would be only 0.01 and 0.001 (i.e., $[1-0.90]^2$ and $[1-0.90]^3$) if the string of two (or three) heat-wave-qualifying days were counted as one 2-day (or 3 day) heat-wave occurrence instead of counting each of the two (or three) days as an occurrence as done here.

When the definition of a heat wave requires two or more consecutive qualifying days, the occurrence of heat waves is governed partly by the climatological level of persistence of weather patterns—a characteristic that varies spatially across the United States. The day-to-day (hereafter, "daily") autocorrelation, statistically representing this tendency for persistence, is governed partly by the speed with which upper-atmospheric waves move through a region, and by the persistence of land surface anomalies—particularly soil moisture (Berg et al. 2014). If each day's temperature anomaly were independent of that on adjacent days, the heat-wave rate would be far lower than observed. If each day in a string of n consecutive heat-wave days is counted as an occurrence of an n-day heat wave (as done in this study), the frequency of occurrence of heat waves of varying durations for different levels of autocorrelation can be determined nonparametrically. Here, using a 90th-percentile requirement, 50000 simulations are conducted using a Gaussian random number generator, in which the autocorrelation at *n*-day lag is the assigned 1-day lag autocorrelation value raised to the nth power (i.e., a first-order Markov model, as used in other extreme daily temperature studies; e.g., Mearns et al. 1984). The resulting heat-wave frequencies are shown in Fig. 1. With a typical 1-day observed lag autocorrelation of approximately 0.7 for daily maximum temperature at individual stations, the frequency of observing at least three consecutive days at the 90th percentile is about 4.8%.

The results shown in Fig. 1 assume that daily maximum or minimum temperature is approximately normally (Gaussian) distributed, as has been assumed in some past studies (e.g.,

Mearns et al. 1984). This assumption is not entirely realistic, however. Deviations from normality in daily maximum temperature over extensive regions were noted in Loikith and Neelin (2015) and Loikith et al. (2018), in the form of a shortened right tail of the distribution. A short right tail, associated with negative skewness, results in a greater increase in the frequency of high percentile temperatures with a given increase in temperature than expected in a normal distribution. In LB17 a tendency for negative skewness was shown in maximum and minimum daily summer U.S. temperatures, particularly in the Rockies and Great Basin regions for maximum temperature. We assume a Gaussian distribution in one of our analyses, but also evaluate the realism of that assumption and the effect of negative skewness on the result.

Assuming a Gaussian distribution and using a baseline daily standard deviation, the increase in the mean daily temperature indicated in climate projections allows for an estimation of the future increase in frequency of single heat-wave-qualifying days, and, when combined with knowledge of the daily autocorrelation, an estimation also of the frequency of strings of multiple consecutive qualifying days. For example, using a 90th-percentile threshold and assuming unchanging daily variability, the probability of a single heat-wave day would increase from the 0.10 baseline to 0.22, then to 0.39, and finally to 0.59 for increases in standardized mean temperature of 0.5, 1.0, and 1.5 standard deviations, respectively. This analysis design was applied to individual U.S. stations in Mearns et al. (1984). The calculation of the frequency of multiday sequences of qualifying days depends also on estimated future values of autocorrelation and daily variability, each of which may remain at the currently observed levels or change in the future.

Mann et al. (2017, 2018) suggest that, as the climate warms, quasi-resonant midlatitude wave patterns may result in more persistent summer temperature extremes (implying greater daily autocorrelation). Increases in daily temperature variability alone would also lead to increases in heat-wave frequency. While Schär et al. (2004) and Yiou et al. (2009) discussed upward trends in observed daily temperature variability in Europe, Rhines and Huybers (2013) found no statistically significant change in observed variability since the mid-twentieth century over North America. Further, Sippel et al. (2015) found a lack of increasing daily variability in various regions of the globe once they accounted for a statistical bias affecting some prior assessments. Huntingford et al. (2013) found a lack of secular trend during approximately 1960-2000 in observed global interannual variability (with notable regional variations, including within the United States in summer), which is a secondary contributor to daily variability. For future projections, Cattiaux et al. (2015) noted increases in daily temperature variability and range in Europe based on CMIP5 models, and Schär et al. (2004) similarly identified increases in variability in Europe based on regional climate model projections. Huntingford et al. (2013) found a lack of model indication of a future global average increase in interannual temperature variability, with a possible decrease suggested for late in the century.

In light of this somewhat ambiguous picture, we here examine the CMIP5 models (Taylor et al. 2012) for changes in

TABLE 1. The 10 CMIP5 models and their sources and references, analyzed here for heat-wave behavior during approximately 1979–2005 for historical baseline and 2005–40 for future projections under the RCP4.5 and RCP8.5 greenhouse gas scenarios. When available, the transient climate response and equilibrium climate sensitivity (after Flato et al. 2013) are given, to be compared with the 30-model average and 90% model uncertainty given in the bottom row.

Model name	Source institution; reference	Transient climate response (°C)	Equilibrium climate sensitivity (°C)
ACCESS1.0	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia; Dix et al. (2013)	2.0	3.8
ACCESS1.3	CSIRO BOM; Dix et al. (2013)	1.7	n/a
BCC_CSM1.1(m)	Beijing Climate Center, China Meteorological Administration, China; Wu et al. (2014)	2.1	2.9
CSIRO Mk3.6.0	CSIRO; Jeffrey et al. (2013)	1.8	4.1
FGOALS-g2	Institute of Atmospheric Physics, Chinese Academy of Sciences and Tsinghua University, China; Li et al. (2013)	1.4	n/a
HadGEM2-CC	Met Office Hadley Centre, United Kingdom; Martin et al. (2011)	n/a	n/a
IPSL-CM5A-MR	L'Institut Pierre-Simon Laplace (IPSL), France; Dufresne et al. (2013)	2.0	n/a
IPSL-CM5B-LR	IPSL; Dufresne et al. (2013)	1.5	2.6
MIROC5	Japan Agency for Marine-Earth Science and Technology, Japan; Watanabe et al. (2010)	1.5	2.7
NorESM1-M	Norwegian Climate Centre, Norway; Iversen et al. (2013)	1.4	2.8
10-model mean		1.7	3.2
30-model mean; 90% uncertainty		1.8 ± 0.6	3.2 ± 1.3

daily autocorrelation and variability during summer in the United States. We examine heat-wave frequencies defined by temperature and apparent temperature, and the factors underpinning the broad features of their geographical distributions across the United States in the present and future, using observations and model output for both historical simulations and future projections. The model projections are examined for near-future decades (to 2040), using 10 CMIP5 models for two greenhouse gas forcing scenarios.

Our main objectives are to 1) demonstrate that daily autocorrelation shapes the spatial distribution of heat-wave frequency in observations and model data; 2) evaluate the realism of CMIP5 models' historical simulations of autocorrelation, heat-wave frequencies, and trends; 3) compare model-projected heat-wave frequencies in apparent temperature with those in temperature alone; and 4) test the hypothesis that future increases in heat-wave frequency are caused predominantly by increases in mean temperature (or mean apparent temperature), modulated by the local climatological daily autocorrelation, with minimal contribution from changes in daily autocorrelation or variability. Following descriptions of the data and methods in sections 2 and 3, respectively, results are presented in section 4, followed by a discussion and some concluding remarks in section 5.

2. Data

This study uses daily station temperature observations, and daily temperature and specific humidity from reanalysis, both spanning 1979–2013. It also uses daily temperature and specific humidity outputs from 10 CMIP5 models (Table 1) for recent past simulations (1979–2005) and near-future climate projections (2006–40). The model projections use the high concentration RCP8.5 (Riahi et al. 2011) and the more moderate RCP4.5 (Thomson et al. 2011) forcing pathways.

Two observational datasets are used. The first is the U.S. Historical Climatology Network (USHCN; Menne et al. 2009, 2015), version 1, providing daily maximum and minimum temperatures for 1218 stations across the continental United States. The USHCN data are postprocessed to maximize homogeneity, based on the homogenized monthly mean data of USHCN, version 2.5 (Menne et al. 2009; Vose et al. 2012). Version 2.5 contains adjustments for temporal discontinuities related to station relocations, change of the time of observation (Vose et al. 2003), and other artifacts. The homogenization method used is that described in Oswald and Rood (2014), where the nonhomogenized daily data are compared with the homogenized monthly mean data. Details about the homogenization process are provided in section S1 of the online supplemental material. Aside from homogenization, some of the daily data in USHCN, version 1, are rejected, based on the data warning flags provided in the version 1 data (see supplemental section S1). Missing data are not filled in. Stations with more than 10% missing data during May-September 1979-2013 are rejected, leaving 857 stations for our analyses. The percentage of days with nonmissing data is taken into account in calculations of the relative (percentage) frequency of heat-wave occurrence.

The second dataset is the North American Regional Reanalysis (NARR; Mesinger et al. 2006), an observational data assimilation system built around the eta forecast model. The NARR contains gridded values at 32-km spatial resolution and provides 3-hourly values of 2-m air temperature, specific humidity, and relative humidity. Daily maximum and minimum temperatures are derived from the 3-hourly NARR values, and the moisture variables enable calculation of the maximum apparent temperature, defined following the multiple linear regression of Steadman (1984) as -1.3 + 0.92T + 2.2e (for shade, with no wind), where T is temperature (°C) and e is vapor pressure (kPa). The apparent

temperature incorporates absolute atmospheric moisture to best reflect human discomfort and is highly correlated with the heat index used by the U.S. National Weather Service. A comparison of the apparent temperature with the heat index is provided in section S2 of the online supplemental material.

Because most of the USHCN stations contain no measurements of atmospheric moisture, the gridded specific humidity data from the NARR at the nearest grid point are used to compute the apparent temperature at the USHCN stations. Use of components from two different datasets introduces some uncertainty into the resulting apparent temperature. However, the viability of this combination was evaluated in LB17 through a comparison with apparent temperatures at 187 stations (NCDC 2011) having direct observations of both temperature and moisture, with favorable results. While the NARR moisture data do have some reported problems, such as a negative bias during high moisture extremes (Raymond et al. 2017), overall, it is considered as representative of observed conditions as most other reanalysis datasets.

The 10 CMIP5 models were selected on the basis of data reliability and processing capacity and were regridded to a common 2° grid. The first individual ensemble member is selected from each model. Table 1 shows the transient climate response and equilibrium climate sensitivity of each of the 10 models (Flato et al. 2013), compared with the full 30-model mean and uncertainty range. The mean of the 10 models is close to that of the full set for both parameters, and the uncertainty range shows that none is near either extreme tail of the 30-model distribution.

Here, just the 2006–40 period is examined, with the rationale that 1) more confidence may be placed in the projections and the greenhouse gas forcing scenarios for the near future, and 2) many users' decision time scales go out only one to two decades. The global average temperature projections under the RCP4.5 and RCP8.5 greenhouse gas scenarios do not diverge substantially until the middle of the twenty-first century (Wuebbles et al. 2017; Sillmann et al. 2013), and are expected to just begin to differ noticeably by 2040. However, the nearterm similarity in mean temperature evolution may not necessarily translate directly to heat-wave frequency (Russo et al. 2015). Examination of both scenarios also provides a secondary benefit of increasing the sample size of projections.

3. Methods

The interannual variability of the frequency of heat waves in the United States is evaluated during May through September, with spatial analyses conducted for the July–August period—the peak of the warm season. A statistical modeling experiment (described below) is carried out to test the hypothesis that future heat-wave increases are caused mainly by increases in mean temperature. A heat wave is defined as at least three consecutive days with daily temperature at or above the 90th percentile threshold, which is defined nonparametrically using ranking (see section S3 of the online supplemental material for further details), with respect to the recent period of 1979–2013 for observations, and 1979–2005 (or 1980–2005) for model simulations. In direct comparisons between observations and models, the 1979–2005 period is used. For the model projections, heat waves are

defined with respect to thresholds identified from each model's own simulations for the historical period for daily maximum and minimum temperature, and maximum apparent temperature. For both the observed and model data, the maximum apparent temperature is calculated using daily average specific humidity. Heat-wave days are tallied by counting each day in a sequence of consecutive qualifying days equally. Thus, for three consecutive qualifying days all three days are counted as an occurrence of a 3-day heat wave, resulting in a count of 3 rather than 1.

Following general results showing time series and spatial maps of historical and projected heat-wave frequencies, the following hypothesis is tested: Model-projected future increases in heat-wave frequency are determined primarily by upward shifts in mean temperature, with little or no contribution from changes in daily variability or autocorrelation. The hypothesis is tested under the assumption of a Gaussian daily temperature distribution, and a first-order Markov daily lag correlation structure in which the 1-day lag correlation is r, the 2-day lag correlation is r^2 , and so on. The theoretically expected future heat-wave frequency is estimated by shifting the models' historical daily temperature distribution by the difference in mean temperature between the models' historical period and their future period, standardized with respect to the mean and standard deviation of the daily historical data. In the analysis, the autocorrelation and daily standard deviation are held fixed to those of the historical data. The model autocorrelation is computed over all years using daily departures from the historical period average for each July day, and therefore contains both intraseasonal and interannual components. (The intraseasonal component alone would use departures only from the average of the days of July for a given single year.) More information about the computation is provided in section S6 of the online supplemental material. A 50 000member synthetic time series with the prescribed autocorrelation is developed from Gaussian random numbers, and the frequency of a single heat-wave qualifying day, and then 3+ heat-wave days, is determined for a specified upward shift in the mean. This calculation is performed separately for each model and each grid point for July, with the resulting expected heat-wave frequencies averaged across all models. Using this nonparametric approach, Fig. 2 shows the frequency of 3+-day heat waves resulting from given increases in standardized mean temperature for given autocorrelation levels, assuming unchanging variability.² These synthetically modeled heat-wave frequencies are compared with direct calculations of heat-wave frequency from the projected model data, and differences are discussed in view of 1) direct determinations of changes in model daily variability and autocorrelation between the historical runs and the future projections and 2) the realism of the Gaussian and Markov assumptions used.

²A plot similar to Fig. 2, but for mean temperature increases for given changes in daily variability assuming an unchanging auto-correlation, is shown in section S4 of the online supplemental material.

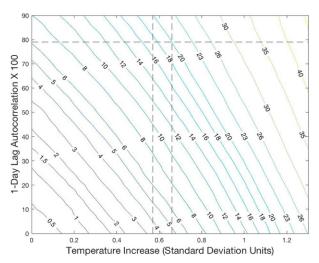


FIG. 2. Probability of occurrence of three or more consecutive heat-wave-qualifying days as a function of standardized temperature increases and daily autocorrelation. A Gaussian distribution of temperature is assumed. A horizontal dashed line is shown at the 0.79 autocorrelation level (as measured for the multimodel July average for the historical model simulation period of 1979–2005), and vertical dashed lines are shown for temperature increases of 0.57 and 0.66 standard deviations (as projected for 2026–40 relative to 1979–2005 for July–August for the RCP4.5 and RCP8.5 scenarios, respectively).

4. Results

a. Heat-wave frequency in observations

The geographical distribution of warm season (May-September) heat-wave frequency over the United States, based on daily maximum temperature during 1979-2013 using the USHCN station data and the NARR gridded data (placed onto a 1° grid; see the caption), is shown in the top panels of Fig. 3. These plots show a relative minimum in frequency in the northeastern and north-central parts of the country where synoptic weather activity, which interrupts persistent regimes, continues through summer. This pattern is similar to that shown in LB17 and is approximately echoed by the patterns for duration and intensity of extreme heat events in Oswald and Rood (2014), also using USHCN data, and less closely matches those in Smith et al. (2013) who used NLDAS data (Xia et al. 2012) under several qualification criteria. The pattern of heat-wave frequency defined using daily minimum temperature (Fig. 3, bottom) shows an additional broad minimum in the Rockies and Great Basin for USHCN data, and overall, fewer heat-wave events than noted for maximum temperature.

The heat-wave frequency patterns show a close correspondence to those of the daily autocorrelation of July³ daily maximum and minimum temperature, shown in Fig. 4, based on the USHCN station data and gridded NARR data (again,

placed onto a 1° grid). It is not surprising that autocorrelation, representing the tendency for anomaly persistence, is associated with heat-wave frequency, given the consecutive day requirement for a heat wave. The autocorrelation in the NARR is higher than in USHCN (Fig. 4) because temperatures are averaged over a grid square in NARR versus a point in USHCN temperatures, the former benefitting from noise filtering. Also, the NARR is based on model dynamics that cannot include all of the smaller-scale processes governing the surface observations, resulting in greater temporal continuity in NARR. In both datasets autocorrelation tends to be higher for maximum than minimum temperature, resulting in higher heat-wave frequencies for maximum temperature (found also in LB17).

The top panel of Fig. 5 shows interannual variations of the percentage frequency of heat waves for the United States as a whole for daily maximum and minimum temperature, averaged over May through September of each year for USHCN and NARR data during 1979-2013. Unlike the frequencies shown in Fig. 3, using heat-wave thresholds based on the 1979– 2013 period, the percentage frequencies here are derived from thresholds based on only the 1979-2005 period (but applied to all of 1979–2013), for comparability to the model simulations (discussed below). For 1979-2013, maximum temperature frequencies in the two datasets are correlated at 0.80, despite the NARR's frequency surge during 2007 and 2009–13, relative to USHCN. The USHCN data show an upward trend in heat waves based on daily maximum temperature (the linear trend slope is 0.10% per year, statistically significant at p < 0.05), while the positive trend in the NARR is greater (0.28% yr⁻¹, significant at p < 0.01). Positive trends in U.S. daily maximum temperature as well as number of days exceeding the 90th percentile were similarly found in Mutiibwa et al. (2015) using NLDAS phase-2 data (derived largely from the NARR data fields). The notable rise in heat-wave frequency based on maximum temperature in the NARR data late in the observed period is found to reflect a similar rise in the maximum temperature. Section S5 of the online supplemental material provides more detail about this surge in NARR maximum temperature.

For daily minimum temperature (dotted lines in top panel of Fig. 5), lower heat-wave frequencies are apparent, and USHCN and NARR frequencies are correlated at 0.89, with statistically significant (p < 0.05) positive trends slopes of 0.12% and 0.08% yr⁻¹, respectively. A larger upward trend in U.S. heat-wave frequency for minimum vis-à-vis maximum temperature was found in LB17, Oswald (2018), and in Peterson et al. (2008) for high temperature extremes in many North American subregions. The trend difference seen here for the USHCN data is in the same sense, but more modest and not statistically significant.

The marked difference in the maximum temperature trend in USHCN versus NARR observations may raise questions about the reality of the NARR data at the surface, assuming

³The autocorrelation patterns for May–September and July–August (not shown) are similar to that of July.

⁴The "1-sided correlation" (Van Den Dool 1987), in which the contribution to the autocorrelation comes only from positive temperature anomalies (for greater relevance to heat waves), is found to be highly similar to the full-sample autocorrelation.

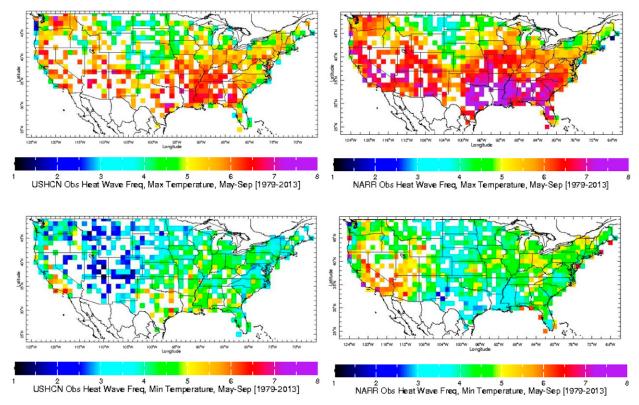


Fig. 3. Observed heat-wave frequency during May–September of 1979–2013, interpolated to a 1° grid, for (left) the USHCN observations and (right) the NARR observations, based on (top) daily maximum temperature and (bottom) minimum temperature. Heat-wave thresholds are derived using the full 35-yr observation period. When more than one station contributes to a 1° cell, the frequencies are averaged for the USHCN data, and, for comparability, the NARR gridded values closest to each station are likewise averaged for the 1° cells. For all maps, if a grid square has no stations within it, it is left blank.

that the homogenized USHCN station observations better reflect ground truth.

The observed heat-wave frequency time series for daily maximum apparent temperature is shown on the bottom panel of Fig. 5, compared with the observed maximum temperature (dotted lines). Frequencies of the two variables generally track one another, correlating at 0.88, and the upward trend in maximum apparent temperature (0.08% yr⁻¹) is significant (p < 0.05). This feature agrees with findings in Gaffen and Ross (1998, 1999), although they found more strongly significant upward trends in apparent temperature in much of the United States for somewhat earlier periods. We also note that, using NLDAS phase-2 data, Smith et al. (2013) showed statistically significant upward trends in heat waves in daily maximum apparent temperature over a similar analysis period as ours, using multiple qualification criteria. As noted above, NARR (and NLDAS) show greater upward trends in temperature and in heat-wave frequency than USHCN, due to NARR's strong late-period increases.

b. Comparison of model simulations with observations

The multimodel mean of the heat-wave frequency for the historical simulations (1979–2005) are shown on the left side of the top and bottom panels of Fig. 5 (purple line), with a closer view of just the historical period of the top panel provided in the middle panel. While simulated and observed frequencies

are at comparable levels in the first decade, simulated frequencies tend to exceed those for the USHCN and NARR observations from the early 1990s through 2005, and thus show a slightly stronger upward trend and a higher period mean. While the NARR frequency rises rapidly from 2007 to 2013, trend consistency during the common period of 1979–2005 is desired. The difference in the linear trend between the USHCN observations and the multimodel mean during 1979–2005 (0.06% yr⁻¹ vs 0.19% yr⁻¹ for maximum temperature-based heat waves) is weakly significant at p=0.07. The trend of the NARR during the period is 0.10% yr⁻¹, statistically indistinguishable from the model trend. These comparisons are important because, if the credibility of the trend in the model simulations is in question, one might distrust the models' future heat-wave trend rates and resulting frequencies.

The left side of the panels of Fig. 6 shows the range in simulated heat-wave frequency across the single members of the 10 CMIP5 models and the multimodel mean, based on daily maximum temperature (Fig. 6, top) and apparent temperature (Fig. 6, bottom). Figure 6 shows that despite the greater upward trend slope in the models than in the observations, the frequencies of both USHCN and NARR observations fall within the range of the model frequencies during most years. On the other hand, the first column of data in Table 2 shows, for each model, the mean heat-wave frequency and the linear

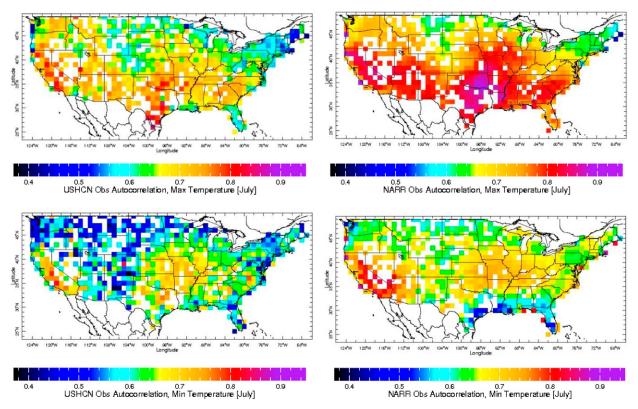


FIG. 4. Adjacent-day lag autocorrelation of July temperature in (left) USHCN-based station observations and (right) NARR reanalysis observations, both for 1979–2013, and interpolated to a 1°-grid, for (top) daily maximum temperature and (bottom) daily minimum.

trend rates for maximum temperature in the model simulations and the observations. Despite substantial differences in mean frequency level and trends among the models (due to both model differences and differing phasing of multidecadal internal variability), upward trends exceed that of the USHCN observations in all 10 models.

Other studies have shown closer correspondence between CMIP5 model and observed trends in actual temperatures than we find here for heat-wave frequencies, for longer or earlier periods and/or larger regions. Kumar et al. (2013) showed a realistic positive trend rate in CMIP modelsimulated annual mean temperature averaged over global land areas during 1930-2004. More specific to the United States during summer in 1980-2005, Lee et al. (2019) noted reasonably good CMIP5 trend fidelity in mean temperature. except for the southeast region where the simulated temperature trend was positive while the observed trend was negative. However, while not significant, they found CMIP5 trends more strongly positive than the USHCN-observed trend in five of the remaining six U.S. regions during 1980-2005. Similarly, Stegall and Kunkel (2017) found greater upward trends for summer in CMIP5 simulations than in observations using U.S. Climate Division data in most regions of the country during 1980-2005.

To further validate the CMIP5 model performance during the historical period, the spatial distributions of the linear trend slopes, the mean heat-wave frequencies, and the daily autocorrelation within the United States in the CMIP5 multimodel average are compared with those of the USHCN and NARR observations during 1979–2005 (Fig. 7). The left column of Fig. 7 shows, despite an overall overestimation of the upward trend by the models, a reasonably good model reproduction of the observed spatial variation of trends in both USHCN and NARR observations, with maxima in the interior southwestern United States and secondarily in the interior mid-Atlantic/Northeast region. The model pattern correlates at approximately 0.6 with both USHCN and NARR. The model reproduction of the pattern of heat-wave frequencies themselves during the historical period (middle column of Fig. 7) is also satisfactory, although the models generally overestimate heat-wave frequencies, and do not reproduce the relative minimum in the northern Great Plains region. This model omission is likely related to the omission of a minimum in daily autocorrelation in the same region (right column of Fig. 7), while approximately reproducing the main features of the autocorrelation pattern otherwise. The models overestimate the autocorrelation level in much of the country, and this bias helps explain the general overestimation of heat-wave frequencies. The bottom row of Fig. 7 shows the spatial distribution of the difference between the models (Fig. 7, third row) and the USHCN observations (Fig. 7, first row). In all three variables—trend, frequency, and autocorrelation—the models tend to show smaller spatial variability across the United States than seen in the observations, in

agreement with findings in Kumar et al. (2013) for temperature itself over a longer period. The spatial average of the standard deviation of the daily maximum temperature is also compared between models and observations during 1979–2005 and is found to be comparable (3.35°C for USHCN and 3.05°C for NARR vs 3.29°C for CMIP5). The exceedance of the model overall autocorrelation level over that of observations is substantially, but not entirely (not shown), due to the larger (here, 2°) grid squares used for the models than for the observations.

The somewhat greater upward trend in heat-wave frequency in the multimodel ensemble (MME) than in observations during 1979–2005, while concerning, does not necessarily imply overestimation of projected future heat-wave frequency. It is possible that the models are reproducing the observed upward trend as well as expected if they are not able to incorporate some circumstances unique to the brief 27-yr period in the United States in summer—circumstances that may not persist into the forthcoming decades. For example, a relative pause in the global warming trend that started around the turn of the century in global surface observations (e.g., Tollefson 2014; Meehl et al. 2011; Trenberth et al. 2014) has been tied to tropical Pacific decadal variability (Kosaka and Xie 2013) and would not be expected to be reproduced in phase in uninitialized model simulations.⁵ More locally, Hansen et al. (1999) found that the United States lagged the globe in its warming rate, and that the eastern United States showed no warming trend (for a longer period, ending in 1998) likely due to multidecadal fluctuations in North Atlantic Ocean sea surface temperature. Another possible longer-term contributor to a slowdown in extreme summer heat is a half-to-full-centurylong agricultural cropland intensification in a large area centered in the Midwest, causing more evapotranspiration and precipitation, suppressing maximum temperatures (Mueller et al. 2015). Moreover, the use of irrigation has been shown to suppress summer daily maximum temperatures (Thiery et al. 2017; Cook et al. 2011; Adegoke et al. 2003).

If the above factors are not regarded as possible extenuating circumstances, one might consider the CMIP5 model output used here as unrealistic given their positive bias in heat-wave trend rates. However, even if heat-wave trends in the model projections are overestimated, this bias does not imply likely failure to reproduce the approximate spatial distribution of heat-wave trends (Fig. 7) and thus future frequency patterns, relative changes in daily variability or autocorrelation, and differences in rates of increase of heat waves using apparent temperature versus temperature, or daily maximum versus minimum temperature.

c. Model projections of future heat-wave frequency

The right side of Fig. 5 (top panel) shows the interannual variations of the projected multimodel average heat-wave

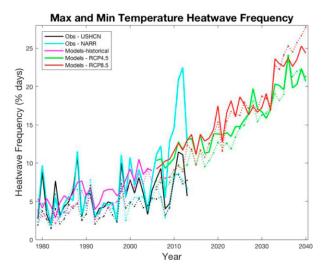
percentage frequency averaged over the United States for the May–September season for daily maximum and minimum temperature for the RCP4.5 and RCP8.5 scenarios. Frequencies for the two scenarios begin at near 10% and reach 20%–25% by 2035–40, with RCP8.5 becoming higher than RCP4.5 as expected given RCP8.5's greater rate of increase in radiative forcing, consistent with findings in Wuebbles et al. (2014) and Sillmann et al. (2013) for annual mean U.S. temperature.

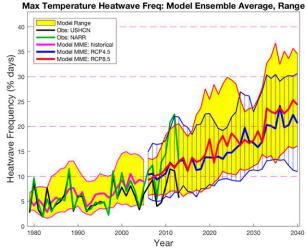
Figure 5 shows heat-wave frequencies maintaining a higher level for maximum than minimum temperature during the historical and much of the future period, as expected assuming that the higher daily autocorrelation for maximum than minimum temperature in observations (Fig. 4) continues into the future. However, the projected difference in frequency between daily maximum and minimum decreases with time, and virtually disappears by 2035-40. This outcome continues the greater upward trend rate in past observed daily minimum than maximum temperature across the United States, with an associated decrease in diurnal temperature range (Karl et al. 1984, 1993; LB17). The higher rate of warming in minimum temperature, also noted in Sillmann et al. (2013) in CMIP5 models, offsets the effect of the lower daily autocorrelation by 2040. The difference in linear trend slope between heat-wave frequency for minimum and maximum temperature, statistically significant (p < 0.05) for both RCP4.5 and RCP8.5 scenarios, is associated with a generally increasing atmospheric moisture accompanying the warming trend, placing higher limits on daily minimum temperatures. Additionally, a higher greenhouse gas concentration decreases the rate of net nighttime radiative cooling. Third, summer minimum temperature has smaller daily variability than maximum temperature (e.g., Barnston 1993), making heat-wave thresholds more readily exceeded for minimum than maximum temperature with comparable increases in mean temperature.

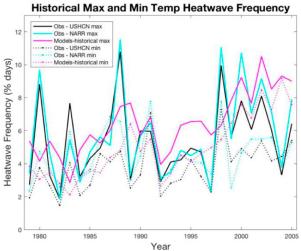
During 2006–40, modeled future frequencies of heat waves in maximum apparent temperature follow those of maximum temperature closely for both RCP4.5 and RCP8.5 scenarios until about 2020, when frequencies for apparent temperature begin exceeding those for temperature by increasing margins (Fig. 5, bottom). The difference between the trend slopes of maximum temperature and apparent temperature is weakly statistically significant (p = 0.07) for the RCP4.5 scenario and significant (p = 0.05) for RCP8.5.

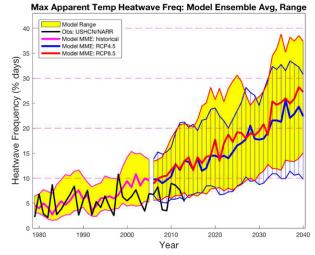
Heat-wave occurrences depend greatly on the persistence of weather patterns that underpin multiday sequences of high temperatures (Grotjahn et al. 2016), reflected in the daily autocorrelation. Autocorrelations could increase if extremes in atmospheric circulation patterns become longer lasting or slower moving, or soil moisture feedbacks intensify. Increases in the strength of circulation patterns might occur if circulation became more meridionally amplified, as conditionally suggested by some (e.g., Coumou et al. 2015, 2017; Mann et al. 2017, 2018). Such amplification could increase the range and the variability of daily maximum temperature, increasing the likelihood of meeting the heat-wave criteria defined from the historical

⁵ However, on a global basis, observed global upward trends in the frequency of *extreme* high temperature continued during this warming pause in some studies (Seneviratne et al. 2014); perhaps the United States is an exception during summer.









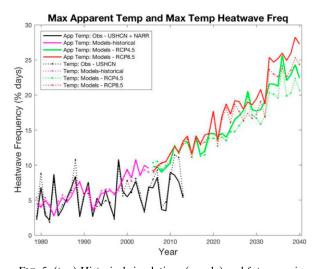


FIG. 6. Multimodel mean and individual model range for (top) maximum temperature and (bottom) maximum apparent temperature for historical simulations for 1979–2005 (the left part of the curves) and projections for the RCP4.5 and RCP8.5 scenarios for 2006–40 (the right part of the curves). Range for historical simulations and for RCP8.5 projections is shaded yellow; range for RCP4.5 is denoted by vertical hatching. Observations are also shown on the left side, where heat-wave thresholds are derived using 1979–2005 data.

FIG. 5. (top) Historical simulations (purple) and future projections (average of the set of 10 CMIP5 models) for 2006–40, along with the USHCN (black) and NARR (cyan) observations for 1979–2013, all for heat-wave percentage frequency for daily maximum temperature (solid lines) and minimum temperature (dotted lines).

period further than expected from an increase in mean temperature alone.

In anticipation of the spatial patterns of model-projected future heat-wave frequency, Fig. 8 shows the pattern of the

 \leftarrow

For the observations, heat-wave thresholds here are derived using 1979–2005 data. For projections, the RCP4.5 and RCP8.5 greenhouse gas scenarios are shown by green and red lines, respectively. (middle) As in the top panel, but for a close-up view of only the 1979–2005 period of model simulations. (bottom) As in the top panel, but for daily maximum apparent temperature (solid lines) and maximum temperature (dotted lines).

TABLE 2. Mean heat-wave percentage frequency over the United States for May–September season and, in parentheses, the linear trend (percentage per year) for daily maximum temperature, for the individual models during the 1979–2005 historical period of historical simulations and the 2006–40 projection period for the RCP4.5 and RCP8.5 scenarios. The standardized anomaly of the maximum temperature itself is shown for the 2026–40 period, with respect to the historical period. The multimodel mean is shown in boldface type. The means and trend slopes for the two observational datasets are shown for comparison, both based on 90th-percentile heat-wave thresholds derived from the 1979–2005 period.

Model	Historical mean percentage heat wave (and linear trend; % yr ⁻¹) 1979–2005	RCP4.5 mean percentage heat wave (and linear trend; % yr ⁻¹) 2006–40	RCP8.5 mean percentage heat wave (and linear trend; % yr ⁻¹) 2006–40	Temperature change from the period of 1979–2005 to 2026–40 (std dev units) RCP4.5/8.5
ACCESS1.0	6.9 (0.25)	19.3 (0.26)	19.3 (0.55)	0.58/0.75
ACCESS1.3	6.9 (0.13)	14.0 (0.22)	15.7 (0.44)	0.43/0.54
BCC_CSM1.1(m)	6.5 (0.13)	14.1 (0.33)	14.9 (0.32)	0.49/0.66
CSIRO Mk3.6.0	6.6 (0.27)	15.6 (0.34)	17.0 (0.22)	0.60/0.58
FGOALS-g2	6.5 (0.20)	12.3 (0.16)	13.1 (0.28)	0.57/0.66
HadGEM2-CC	6.9 (0.07)	15.6 (0.45)	20.1 (0.59)	0.64/0.78
IPSL-CM5A-MR	6.4 (0.28)	18.3 (0.40)	20.8 (0.51)	0.49/0.63
IPSL-CM5B-LR	6.4 (0.16)	12.0 (0.44)	13.4 (0.46)	0.60/0.62
MIROC5	6.0 (0.22)	16.9 (0.52)	15.8 (0.53)	0.75/0.79
NorESM1-M	5.8 (0.20)	14.3 (0.45)	14.8 (0.47)	0.52/0.59
MME Avg	6.5 (0.19)	15.2 (0.36)	16.5 (0.44)	0.567/0.660
USHCN obs				
1979-2005	5.39 (0.06)			
1979-2013	5.93 (0.10)			
NARR obs	, ,			
1979-2005	4.93 (0.10)			
1979-2013	6.56 (0.28)			

multimodel mean 1-day lag autocorrelation for July daily maximum temperature for the 2026-40 periods for projections.⁶ It is important to note that the model-projected autocorrelation patterns in Fig. 8 show little change from the historical period⁷ (right panel of the bottom-middle row of Fig. 7), with ratios of the average autocorrelation for RCP4.5 or RCP8.5 to the historical autocorrelation both 1.00 with small geographical variation. The multimodel average autocorrelation (see Table S1 in the online supplemental material) averages 0.08-0.09 higher for daily maximum than minimum temperature, helping to explain the relatively lower heat-wave frequency for minimum temperature in the simulations (consistent with observations) and early portion of the future projections (Fig. 5, top), before the effect of the greater warming rate in minimum than maximum temperature begins outweighing that of the autocorrelation.

The upper-left panel of Fig. 9 shows the spatial distribution of the multimodel average percentage frequency of heat waves in daily maximum temperature for the historical simulations

(1979-2005) for July-August. Frequencies are consistent with those of the USHCN observations during May-September 1979-2005 (Fig. 7, top-center panel), except that they do not reproduce the relatively low frequencies in the north-central portions of the country—a feature reflected also in the models' historical autocorrelation pattern (Fig. 7, right panel of the bottom-middle row). The autocorrelations of the observed data (Fig. 4 and supplementary Table S1) average lower than those of the historical model simulations, partly because of spatial resolution differences where CMIP5 models are most coarse (2° grid), NARR at an intermediate coarseness (32 km grid), and stations at a point; coarser resolution permits more noise filtering. Even taking the resolution differences into account, however, the model autocorrelations are slightly higher than those of the observations, likely due to their necessarily inadequate representation of some subgrid-scale physical processes. This overestimation of autocorrelation may underlie the models' slight overestimation of heat-wave levels and trend rates in the historical period, and therefore possible overestimation of both of these in the projections.

The second and third panels of Fig. 9 show, respectively, heatwave frequency during 2026–40 for the RCP4.5 and RCP8.5 scenarios, using the 90th-percentile level defined from the models' historical runs. While magnitudes are slightly higher for RCP8.5, the patterns are similar, with highest frequencies in the Southeast, Great Basin, and Southwest, and lowest in the Great Lakes and north-central region. The projected overall spatial average heatwave frequency of 20%–25% is 3–4 times the average frequencies shown over most of the historical period, but the models' trend rate may be overestimated, as noted above. The lower-right panel

⁶Further details about the computation of the autocorrelation, and the autocorrelations of the individual models during 1979–2005 and 2026–40, are provided in section S6 of the online supplemental material.

⁷ Note that for the model projections heat-wave frequencies can be high without also having high autocorrelations, because the autocorrelation is computed with respect to means for the projection period (2026–40) whereas heat-wave frequencies are computed using thresholds from the historical period (1979–2005).

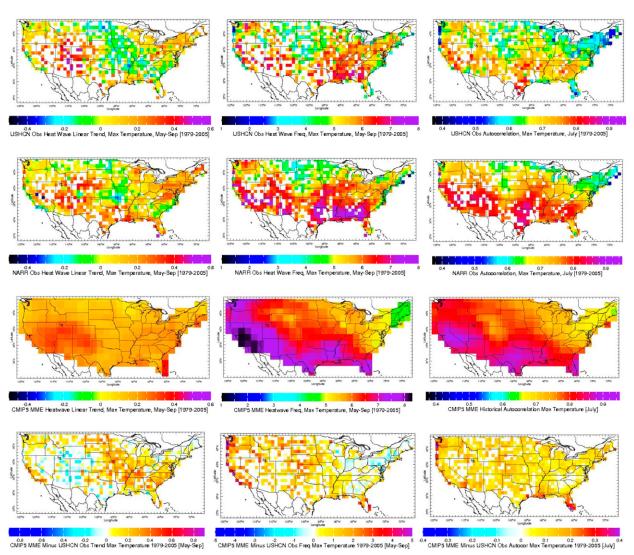


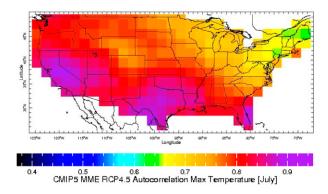
FIG. 7. Comparison of (bottom middle) CMIP5 multimodel ensemble with the (top) USHCN and (top middle) NARR observations during the 1979–2005 period for (left) linear trend slope (percentage change in heat-wave frequency per year), (center) period-mean percentage heat-wave frequency, and (right) adjacent-day lag correlation. Linear trend and mean frequency are for the May–September period, and the autocorrelation is for July. (bottom) Differences between the models and the USHCN observations. Trends and frequencies are derived from heat-wave thresholds based on the 1979–2005 period, and the observed frequency results shown in Fig. 3 use thresholds based on 1979–2013.

of Fig. 9 shows frequency of heat waves in daily maximum apparent temperature for the RCP8.5 scenario, with a pattern similar to that for maximum temperature but with generally higher frequencies.

The spatial patterns of future heat-wave frequency shown in Fig. 9 differ somewhat from those for the model's historical simulations, as for example a shift of the maximum in the Southwest to the Great Basin for the projections, and higher values in the mid-Atlantic region.⁸ This change in pattern is

largely a function of the projected pattern of increase in mean maximum temperature but constrained by the autocorrelation pattern such that regions of higher autocorrelation are able to have larger heat-wave frequency increases. The top row of Fig. 10 shows the pattern of the multimodel-projected increase in maximum temperature itself, from the period of 1979–2005 to 2026–40 for the RCP4.5 and RCP8.5 scenarios. In the second row, this increase is standardized, using statistics of the historical period. The standardized and nonstandardized patterns differ considerably. Since the 90th percentile heat-wave threshold marks a relative position in the historical distribution (as would a standardized anomaly), it is the standardized future mean temperature increase rather than the temperature increase itself that relates to the change in future heat-wave frequency from the historical frequency; all other factors—for

⁸The notable increase along the coasts of the Atlantic and Gulf of Mexico is due to the low historical daily variability near large water bodies that makes the heat-wave threshold easy to exceed for a given rise in mean temperature.



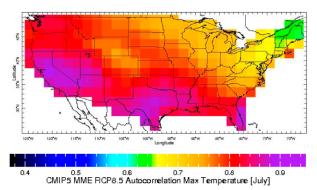


FIG. 8. Multimodel average 1-day lag autocorrelation of July daily maximum temperature during the projected 2026–40 period for the (top) RCP4.5 and (bottom) RCP8.5 scenarios.

example, daily autocorrelation and daily variability—remaining approximately fixed. As noted above, the climatological (even if unchanging) spatially varying values of the autocorrelation modulate the extent to which the standardized temperature increase can increase the heat-wave frequencies, determining their future spatial pattern. The relative maxima in future frequencies in the Great Basin and along the east coast of Florida, for example, come about due to the intersection of high autocorrelation and high projected standardized temperature increase.

Increases in both temperature and absolute moisture content contribute to changes in heat-wave frequency defined by the apparent temperature. The projected spatial distribution of frequency for maximum apparent temperature for July-August 2026-40 for the RCP8.5 scenario was shown in the last panel of Fig. 9. The projected change in maximum apparent temperature itself, from the historical simulations to the 2026-40 period for July-August for each RCP scenario is shown in the third row of Fig. 10, while the bottom row shows the change in maximum apparent temperature minus that of temperature. For both scenarios, the increase in apparent temperature is comparable to that in temperature in much of the arid western part of the United States, while east of the Continental Divide the apparent temperature increases by greater amounts than temperature, particularly in the Southeast and mid-Atlantic seaboard. In section S7 of the online supplemental material, model-projected changes in two measures of absolute atmospheric moisture—specific humidity and dewpoint—are

illustrated, and their contributions to the pattern of change in apparent temperature are discussed.

d. Hypothesis: Future heat-wave frequency change is due mainly to mean temperature change

Here we discuss the results of the hypothesis test described at the end of section 3 for statistically predicting future heatwave frequencies on the basis of the predicted change in mean temperature alone. If projected increases in future heat-wave frequency were due solely to projected increases in mean temperature without contributions from changes in autocorrelation or daily variability, and if daily temperature is approximately normally distributed (which we have noted is not entirely realistic) and the autocorrelation profile is approximately that of a first-order Markov process, then heat-wave frequency modeled using synthetic Gaussian time series with an assigned autocorrelation and an appropriately upwardshifted mean (see section 3 and Fig. 2) should match the heat-wave frequency determined directly from the model projections. The top two rows of Fig. 11 show a comparison of the heat-wave frequencies for July 2026–40 predicted using the synthetic data, with those directly computed from the RCP4.5 and RCP8.5 model projections. While the spatial patterns are similar, the directly computed frequencies average higher than those synthetically modeled by a factor of 1.10 (averaging 20.2% vs 18.3%) for RCP4.5 scenario, and 1.11 (22.1% vs 20.0%) for RCP8.5. Little or no change in the levels of autocorrelation is noted from historical to future periods (Fig. 8 and supplementary Table S1). This outcome may therefore be due to higher daily variability in the model projections than in the historical simulations, systematic departures from normality in the daily temperature data, or deviations of the autocorrelation profiles from a first-order Markov process. We next examine each of these possibilities.

Table 3 shows a comparison of the spatial average of the daily standard deviation in the models' historical simulations and projections. The daily variability is just slightly (about 2%) and statistically insignificantly larger in projections than in the historical simulations. The geographical distribution of the ratio of July daily standard deviation during 2026–40 to that in the historical simulations is shown in section S8 of the online supplemental material. Following the relationship between a change in standard deviation and heat-wave frequency (Fig. S2 of the online supplemental material), for a mean temperature increase of 0.6 standard deviations, a 2% increase in variability implies also an additional approximate 2% increase in frequency, making a minor contribution to the approximate 10% difference between the synthetically modeled and directly computed future frequencies.

Another possible cause of the difference is a lack of a Gaussian daily maximum temperature distribution, such as a shortened right tail (e.g., Loikith and Neelin 2015; Loikith et al. 2018). When the Kolmogorov–Smirnov test is applied to the models' historical simulations of maximum temperature for 80 July days spanning 1980–89 separated by 4 days (to minimize autocorrelation), the null hypothesis of normality cannot be rejected at the 0.05 significance level in more than 5% of the grid squares for any individual model. Inspection of the

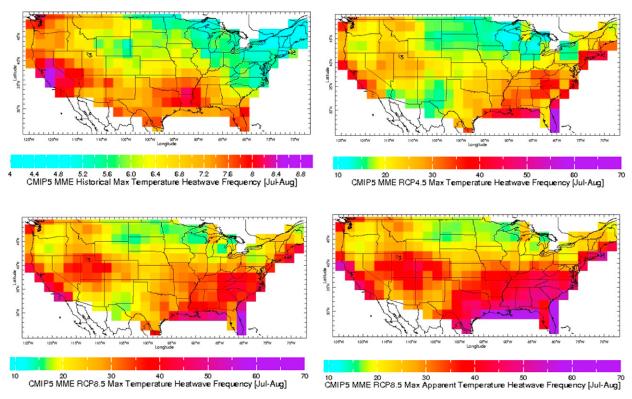


FIG. 9. Multimodel average of the percentage frequency of 3-day (or more) heat-wave occurrence (≥90th percentile), for daily maximum temperature averaged over July–August for the (top left) 1979–2005 historical simulation period, the (top right) RCP4.5 and (bottom left) RCP8.5 greenhouse gas scenarios during 2026–40, and the (bottom right) daily maximum apparent temperature under the RCP8.5 scenario. Note the different scale in the top-left panel.

rejected cases frequently shows no notable geographical preference. Nonetheless, when skewness is computed for July daily maximum temperature for the USHCN station observations and the multimodel ensemble of the CMIP models (Fig. 12), mild to moderate negative skewness appears in the western part of the country—particularly in the Rockies and part of the Great Basin, in agreement with findings in LB17. While the model skewness is slightly less than that observed, the spatial patterns match reasonably well. A shortened right tail implies a negative skewness, in which the 90th percentile corresponds to a lower standardized anomaly value than in a normal distribution (i.e., lower than +1.282 standard deviations). Figure 12 (bottom) confirms that the standardized anomaly of the 90th percentile in the multimodel historical data is lower than expected for a Gaussian distribution in approximately the region having negative skewness. While offset somewhat by positive skewness in the Midwest in the multimodel ensemble, the U.S. average skewness is -0.12, and the average standardized anomaly of the 90th percentile is 1.274. It is difficult to estimate the expected correspondence of this deviation from 1.282 to single-day and multiple day heat-wave frequency without knowing more detail about the shape of the distribution. For a normal distribution there would be approximately a 1.5% single-day frequency underestimation and a 2.3% frequency underestimation for three consecutive days. This slight overall tendency toward negative skewness, representing a failure of the assumption of a normal distribution, likely makes a detectable but minor contribution to the 10% higher directly measured heat-wave frequencies than those theoretically expected assuming a Gaussian distribution.

An analysis of the profile of lag correlation of daily temperature in the observations and model output data does not reveal salient deviations from a first-order Markov model beyond sampling variability, as the lag correlations drop off, on average, as r^t where r is the 1-day lag correlation and t is the number of days of lag. Further detail about this analysis is provided in section S9 of the online supplemental material.

A final, possibly unexpected, factor that can cause the actual heat-wave frequencies to exceed those synthetically modeled is a positive bias in the *actual* computed frequencies caused by applying heat-wave thresholds derived from the historical period to projections outside of the historical period (Sippel et al. 2015). The bias stems from uncertainty in the statistics within the historical reference period due to the finite sample size, causing the frequency distribution in an external period to follow a Student's *t* distribution (with higher variance and thicker-shaped tails than the normal distribution), such that frequencies of right-tail threshold exceedance are greater. For the 90th-percentile threshold and the effective sample size of the historical period used here, the overestimation of heatwave frequencies outside of the historical period relative to

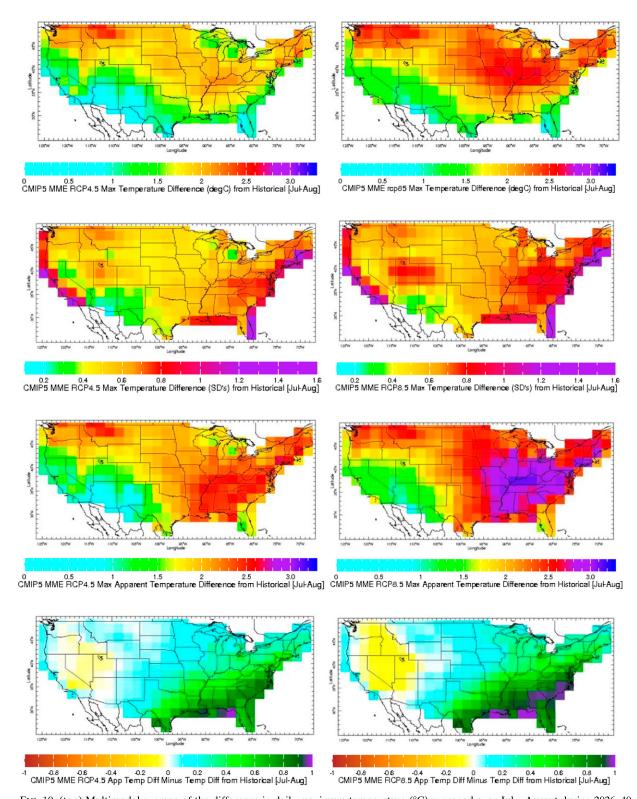


FIG. 10. (top) Multimodel average of the difference in daily maximum temperature (°C) averaged over July–August during 2026–40 period, and that for the historical period of 1979–2005, for the (left) RCP4.5 and (right) RCP8.5 greenhouse gas scenarios. (top middle) As in the top row, except the difference is standardized with respect to the mean and standard deviation of the historical period of 1979–2005. (bottom middle) As in the top row, but for the difference in daily maximum apparent temperature (°C); and (bottom) the difference between the daily maximum apparent temperature (°C) increase from the historical period to the 2026–40 projection (shown in the bottom-middle row) and the increase in daily maximum temperature (as shown in the top row).

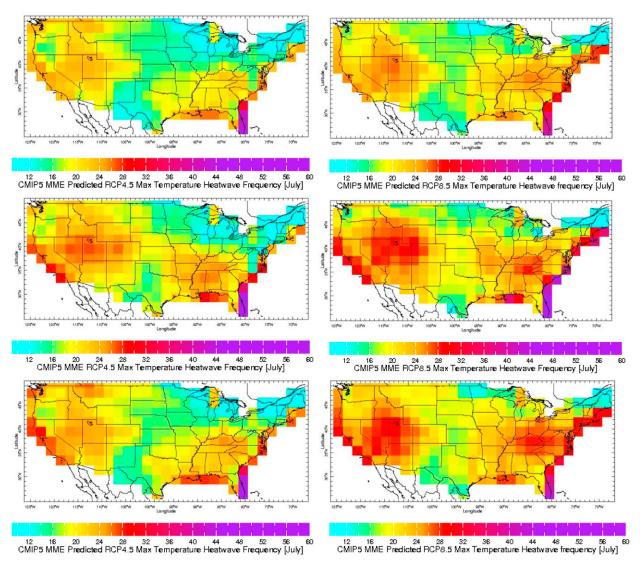


Fig. 11. (top) July heat-wave frequency for the (left) RCP4.5 and (right) RCP8.5 scenarios derived using Gaussian synthetic data with an upward shift in mean daily maximum temperature matching that shown in the 10-model average between the historical period (1979–2005) and the 2026–40 projections, and using historical autocorrelations. (middle) Directly computed 10-model average July heat-wave frequency for 2026–40. (bottom) As in the top row, but adjusted for the ratio of the standard deviation during 2026–40 to that during the historical period and also multiplied by 1.07 to account for the Sippel et al. (2015) statistical bias.

frequencies within the historical period is estimated at 7%. More detail about the calculation of this estimate is provided in section S10 of the online supplemental material. This factor is larger than those stemming from the changes in the autocorrelation, daily variability, and deviations from normality and first-order Markov behavior, and likely explains most of the difference between the actual heat-wave frequencies and the synthetically modeled frequencies. This factor may cause a slight overestimation of future heat-wave frequencies in all of our results.

If the synthetically modeled heat-wave frequencies are adjusted to take into account the geographical pattern of the future-to-historical standard deviation ratio (see Fig. S5 of the online supplemental material), and also multiplied by 1.07 to account for the Sippel et al. (2015) statistical bias, the resulting modeled frequencies come out as shown in the bottom row of Fig. 11. These frequencies more closely match those directly computed, shown in the middle row (with differences of 1% or less when averaged over the United States), despite remaining differences in the details of the spatial patterns that may be associated with varying departures from normal distributions or sampling variability in the frequency of heat waves due to the limited sample size of future cases.

Comparing the synthetically modeled with the directly analyzed heat-wave frequency increases, and directly assessing model changes in autocorrelation and daily variability, it appears that the projected increase in mean temperature,

TABLE 3. Intraseasonal and full standard deviation (°C) of daily July maximum temperature in the historical simulations (1979–2005) of 10 CMIP5 models, and percentage of the standard deviation of the historical simulations found in the future RCP4.5 and RCP8.5 scenarios (2026–40). The multimodel mean is shown in boldface type. For comparison, for 1979–2013, similar to the models' historical period, the intraseasonal and full standard deviations for the NARR and USHCN observations for July maximum temperature are shown at the bottom of the table.

	Intraseasonal std dev		Full std dev			
Model	Historical 1979–2005 (obs to 2013)	RCP4.5 2026–40 as percent of historical	RCP8.5 2026–40 as percent of historical	Historical 1979–2005 (obs to 2013)	RCP4.5 2026–40 as percent of historical	RCP8.5 2026–40 as percent of historical
ACCESS1.0	2.92	104	106	3.30	97	101
ACCESS1.3	3.31	98	102	3.74	97	98
BCC_CSM1.1(m)	3.73	102	99	4.01	105	97
CSIRO Mk3.6.0	3.03	100	100	3.33	105	100
FGOALS-g2	3.02	102	103	3.13	104	108
HadGEM2-CC	2.95	105	105	3.33	103	105
IPSL-CM5A-MR	2.06	106	113	2.28	105	111
IPSL-CM5B-LR	2.28	98	98	2.55	97	101
MIROC5	2.64	97	97	2.90	96	96
NorESM1-M	2.36	110	108	2.59	108	109
MME Avg	2.83	102.2	103.1	3.29	101.4	102.4
NARR obs	2.75			3.05		
USHCN obs	3.01			3.35		

modulated by the autocorrelation, is responsible for the large majority of the projected increase in heat-wave frequency.

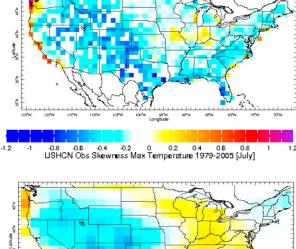
5. Discussion and conclusions

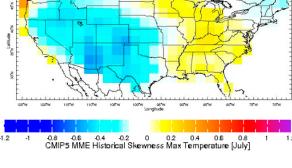
The frequency of heat waves (exceedance of the 90th percentile for 3 + consecutive days) during summer in the United States is analyzed for daily maximum and minimum temperature and maximum apparent temperature in recent observations and in the historical simulations and near-term projections of 10 CMIP5 climate models using RCP4.5 and RCP8.5 greenhouse gas forcing pathways. Over the last several decades, the percentage of days participating in a heat wave was approximately 2%-10% in observed data and in the CMIP5 models' historical simulations. The mean value and the upward trend of the percentage is slightly overestimated in the models' historical simulations relative to the observations, possibly because the models' daily autocorrelations are likewise higher than those observed. Some conditions possibly unique to the historical period in the United States in summer that might exonerate the models were considered; nonetheless, the projections may not necessarily provide quantitatively accurate guidance on future heat-wave frequency. However, based on evaluations of the spatial patterns of the model's heat-wave frequencies and trends, the reliability of the large-scale spatial features of the projected heat-wave frequencies, and differences in overall frequencies for maximum temperature versus maximum apparent temperature, maximum versus minimum temperature, or RCP4.5 versus RCP8.5, may well be realistic.

Keeping the above caveat in mind, when applying the 90th-percentile temperature threshold in the models' historic simulations to their future projections, heat-wave frequencies based on daily maximum temperature rise to 20%–24% under the RCP4.5 scenario and 23%–26% under the RCP8.5 scenario

by 2035–40. These projections vary by up to approximately ±8% among models. While projected heat-wave frequencies based on daily minimum temperature begin lower than those for maximum, their trend rate exceeds that of maximum temperature, as found also in the last 50 years of observations. Heat-wave frequencies based on the daily maximum apparent temperature, reflecting both atmospheric moisture and temperature, increase at a slightly (but statistically significantly) faster rate than temperature itself. This faster increase would have implications for human comfort and health (McGregor 2011), requiring more energy for cooling systems, although some acclimation would be expected (e.g., Sheridan and Dixon 2017; Vicedo-Cabrera et al. 2018; Gasparrini et al. 2015).

CMIP5 model output indicates little or no change in the values or patterns of daily autocorrelation of maximum or minimum temperature between the recent past and the nearfuture decades. These autocorrelation patterns, reflecting the climatological degree of persistence of synoptic weather regimes, substantially contributes to the observed pattern of heat-wave frequency across the United States, and places controls on the increases in frequency in the model projections as the mean temperature rises. The models indicate little change in the daily variability for the future temperature (a statistically insignificant approximate 2% increase, with some geographic variation). Mild to moderate negative skewness in the distribution of daily maximum temperature is seen in part of the western United States, but it is not found to seriously compromise the performance of a statistical prediction of future heat-wave frequency based on the predicted mean temperature increase, that assumes a Gaussian distribution. The lack of significant projected change in autocorrelation and variability suggests that future increases in heat-wave frequency for temperature and apparent temperature are attributable predominantly to increases in mean standardized





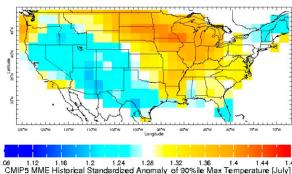


FIG. 12. (top) Skewness in the USHCN observations of July daily maximum temperature during the 1979–2005 period. (middle) As in the top panel, but for the multimodel simulations during the historical period of 1979–2005. (bottom) Standardized anomaly of the 90th-percentile heat-wave threshold for July daily maximum temperature during the historical period. (In a Gaussian distribution, the skewness would be 0 and the standardized anomaly of the 90th percentile would be 1.28.)

(with respect to historical statistics) temperature and absolute moisture. This standardized increase in the mean, modulated by the daily autocorrelation (higher autocorrelation permitting greater increases in heat-wave frequency), leads to the spatial pattern of heat-wave frequency projected by the CMIP models over the coming two decades.

For both the RCP4.5 and RCP8.5 greenhouse gas scenarios, the future heat-wave frequency pattern for July-August during 2026–40 features highest frequencies (28%–45%) in the southeastern quarter of the country, the Great Basin and along coastlines, and lowest frequencies (10%–16%) in part

of the northern tier. Increases in temperature show a fairly uniform pattern, similar to the midcentury projected increases shown in Vose et al. (2017), with smallest increases in the southern tier and along coastlines. Projected increases in standardized temperature are largest in parts of the eastern portion and Great Basin, consistent with findings in Stegall and Kunkel (2017) for near-term CMIP5-projected increases in extreme monthly mean standardized temperature. Increases in absolute moisture (e.g., specific humidity and dewpoint) are sufficient for the apparent temperature to increase more than temperature itself everywhere except the Great Basin, with largest increases in the southeast region and along most coastlines. The increased apparent temperature portends greater difficulty for humans to maintain comfort and safety (e.g., Coffel et al. 2018).

Increasing mean temperature is expected to have consequential implications for infrastructure, such as aircraft operations (Coffel et al. 2017), energy requirements (Scott et al. 2008; Petri and Caldeira 2015), and plant ecology (Teskey et al. 2014). Additionally, maintaining human well-being (Peng et al. 2011) depends not just on temperature but on the increases in atmospheric moisture projected over the densely populated eastern United States.

A primary aim in this study is to demonstrate the governing roles of standardized temperature increase and daily autocorrelation on projected heat-wave frequency. Additional determinants of future frequency are *changes* in future daily variability and autocorrelation, but here both of these are found to be insignificant, leading to the conclusion that future heat-wave frequencies are governed predominantly by increases in standardized temperature (or temperature/moisture) modulated by a generally unchanging local climatological daily autocorrelation. Using nonparametric methods with random Gaussian-distributed numbers, levels and spatial patterns of future heat-wave frequency are shown to be predictable with reasonable accuracy despite the fact that the assumption of a Gaussian daily maximum temperature distribution is somewhat violated in some regions of the country.

Acknowledgments. The authors thank the reviewers for their helpful suggestions. This work was supported by Award NA14OAR43110204 from the National Oceanic and Atmospheric Administration.

REFERENCES

Adegoke, J. O., R. A. Pielke Sr., J. Eastman, R. Mahmood, and K. C. Hubbard, 2003: Impact of irrigation on midsummer surface fluxes and temperature under dry synoptic conditions: A regional atmospheric model study of the U.S. High Plains. *Mon. Wea. Rev.*, **131**, 556–564, https://doi.org/10.1175/1520-0493(2003)131<0556:IOIOMS>2.0.CO;2.

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res., 111, D05109, https://doi.org/ 10.1029/2005JD006290.

Anderson, B. G., and M. L. Bell, 2009: Weather-related mortality: How heat, cold, and heatwaves affect mortality in the United States. *Epidemiology*, **20**, 205–213, https://doi.org/10.1097/EDE.0b013e318190ee08.

- Barnston, A. G., 1993: Atlas of Frequency Distribution, Auto-Correlation and Cross-Correlation of Daily Temperature and Precipitation at Stations in the US, 1948-1991 (in Metric Units). NOAA Atlas 11, 440 pp.
- Berg, A., B. R. Lintner, K. L. Findell, S. Malyshev, P. C. Loikith, and P. Gentine, 2014: Impact of soil moisture–atmosphere interactions on surface temperature distribution. *J. Climate*, 27, 7976–7993, https://doi.org/10.1175/JCLI-D-13-00591.1.
- Cattiaux, J., H. Douville, R. Schoetter, S. Parey, and P. Yiou, 2015: Projected increase in diurnal and interdiurnal variations of European summer temperatures. *Geophys. Res. Lett.*, 42, 899–907, https://doi.org/10.1002/2014GL062531.
- Coffel, E. D., T. R. Thompson, and R. M. Horton, 2017: The impacts of rising temperatures on aircraft takeoff performance. Climatic Change, 144, 381–388, https://doi.org/ 10.1007/s10584-017-2018-9.
- ——, R. M. Horton, and A. de Sherbinin, 2018: 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.
- Cook, B. I., M. J. Puma, and N. Y. Krakauer, 2011: Irrigation induced surface cooling in the context of modern and increased greenhouse gas forcing. *Climate Dyn.*, 37, 1587–1600, https://doi.org/10.1007/s00382-010-0932-x.
- Coumou, D., J. Lehmann, and J. Beckmann, 2015: The weakening summer circulation in the Northern Hemisphere midlatitudes. Science, 348, 324–327, https://doi.org/10.1126/ science.1261768.
- ——, K. Kornhuber, J. Lehmann, and V. Petoukhov, 2017: Weakened flow, persistent circulation, and prolonged weather extremes in boreal summer. *Climate Extremes: Patterns and Mechanisms, Geophys. Monogr.*, Vol. 226, Amer. Geophys. Union, 61–73.
- Dix, M., and Coauthors, 2013: The ACCESS coupled model: Documentation of core CMIP5 simulations and initial results. *Aust. Meteor. Oceanogr. J.*, **63**, 83–99, https://doi.org/10.22499/2.6301.006.
- Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 earth system model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, https://doi.org/10.1007/s00382-012-1636-1.
- Durre, I., J. M. Wallace, and D. P. Lettenmaier, 2000: Dependence of extreme daily maximum temperatures on antecedent soil moisture in the contiguous United States during summer. *J. Climate*, 13, 2641–2651, https://doi.org/10.1175/1520-0442(2000) 013<2641:DOEDMT>2.0.CO;2.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.
- Gaffen, D. J., and R. J. Ross, 1998: Increased summertime heat stress in the US. *Nature*, 396, 529–530, https://doi.org/10.1038/ 25030.
- —, and —, 1999: Climatology and trends of U.S. surface humidity and temperature. *J. Climate*, **12**, 811–828, https://doi.org/10.1175/1520-0442(1999)012<0811:CATOUS>2.0.CO;2.
- Gasparrini, A., and Coauthors, 2015: Temporal variation in heat-mortality associations: A multicountry study. *Environ. Health Perspect.*, 123, 1200–1207, https://doi.org/10.1289/ehp.1409070.
- Grotjahn, R., and Coauthors, 2016: North American extreme temperature events and related large scale meteorological patterns: A review of statistical methods, dynamics, modeling, and trends. *Climate Dyn.*, 46, 1151–1184, https://doi.org/ 10.1007/s00382-015-2638-6.

- Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. J. Geophys. Res., 104, 30 997–31 022, https://doi.org/10.1029/1999JD900835.
- ——, M. Sato, and R. Ruedy, 2012: Perception of climate change. Proc. Natl. Acad. Sci. USA, 109, E2415–E2423, https://doi.org/ 10.1073/pnas.1205276109.
- Horton, R. M., J. S. Mankin, C. Lesk, E. Coffel, and C. Raymond, 2016: A review of recent advances in research on extreme heat events. *Curr. Climate Change Rep.*, 2, 242–259, https://doi.org/ 10.1007/s40641-016-0042-x.
- Huntingford, C., P. D. Jones, V. N. Livina, T. M. Lenton, and P. M. Cox, 2013: No increase in global temperature variability despite changing regional patterns. *Nature*, 500, 327–330, https://doi.org/10.1038/nature12310.
- Iversen, T., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1–M—Part 2: Climate response and scenario projections. *Geosci. Model Dev.*, 6, 389–415, https://doi.org/ 10.5194/gmd-6-389-2013.
- Jeffrey, S. J., L. D. Rotstayn, M. A. Collier, S. Dravitzki, C. Hamalainen, C. Moeseneder, K. K. Wong, and J. Syktus, 2013: Australia's CMIP5 submission using the CSIRO Mk3.6 model. Aust. Meteor. Oceanogr. J., 63, 1–13, https://doi.org/ 10.22499/2.6301.001.
- Karl, T. R., G. Kukla, and J. Gavin, 1984: Decreasing diurnal temperature range in the United States and Canada from 1941 through 1980. J. Climate Appl. Meteor., 23, 1489–1504, https://doi.org/10.1175/1520-0450(1984)023<1489:DDTRIT>2.0.CO;2.
- —, and Coauthors, 1993: A new perspective on recent global warming: Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.*, 74, 1007–1023, https:// doi.org/10.1175/1520-0477(1993)074<1007:ANPORG>2.0.CO;2.
- Kosaka, Y., and S. P. Xie, 2013: Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, **501**, 403–407, https://doi.org/10.1038/nature12534.
- Kumar, S., V. Merwade, J. L. Kinter III, and D. Niyogi, 2013: Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations. J. Climate, 26, 4168–4185, https://doi.org/10.1175/ JCLI-D-12-00259.1.
- Lee, J., D. Waliser, H. Lee, P. Loikith, and K. E. Kunkel, 2019: Evaluation of CMIP5 ability to reproduce twentieth century regional trends in surface air temperature and precipitation over CONUS. *Climate Dyn.*, 53, 5459–5480, https://doi.org/ 10.1007/s00382-019-04875-1.
- Li, L., and Coauthors, 2013: The Flexible Global Ocean-Atmosphere-Land System Model, grid-point version 2: FGOALS-g2. *Adv. Atmos. Sci.*, **30**, 543–560, https://doi.org/10.1007/s00376-012-2140-6.
- Loikith, P. C., and J. D. Neelin, 2015: Short-tailed temperature distributions over North America and implications for future changes in extremes. *Geophys. Res. Lett.*, 42, 8577–8585, https://doi.org/10.1002/2015GL065602.
- —, —, J. Meyerson, and J. S. Hunter, 2018: Short warm-side temperature distribution tails drive hot spots of warm temperature extreme increases under near-future warming. *J. Climate*, 31, 9469–9487, https://doi.org/10.1175/JCLI-D-17-0878.1.
- Lyon, B., and A. G. Barnston, 2017: Diverse characteristics of U.S. summer heatwaves. *J. Climate*, **30**, 7827–7845, https://doi.org/10.1175/JCLI-D-17-0098.1.
- —, E. D. Coffel, and R. M. Horton, 2019: Projected increase in the spatial extent of contiguous U.S. summer heat waves and associated attributes. *Environ. Res. Lett.*, **14**, 114029, https://doi.org/10.1088/1748-9326/ab4b41.

- Mann, M. E., S. Rahmstorf, K. Kornhuber, B. A. Steinman, S. K. Miller, and D. Coumou, 2017: Influence of anthropogenic climate change on planetary wave resonance and extreme weather events. *Sci. Rep.*, 7, 45242, https://doi.org/10.1038/srep45242.
- —, —, —, —, S. Petri, and D. Coumou, 2018: Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification. *Sci. Adv.*, **4**, eaat3272, https://doi.org/10.1126/sciadv.aat3272.
- Martin, G. M., and Coauthors, 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geosci. Model Dev.*, **4**, 723–757, https://doi.org/10.5194/gmd-4-723-2011.
- McGregor, G. R., 2011: Human biometeorology. *Prog. Phys. Geogr.*, **36**, 93–109, https://doi.org/10.1177/0309133311417942.
- McKinnon, K. A., A. Rhines, M. P. Tingley, and P. Huybers, 2016: Long-lead predictions of eastern United States hot days from Pacific sea surface temperatures. *Nat. Geosci.*, 9, 389–394, https://doi.org/10.1038/ngeo2687.
- Mearns, L. O., R. W. Katz, and S. H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *J. Climate Appl. Meteor.*, 23, 1601–1613, https://doi.org/10.1175/1520-0450(1984)023<1601: EHTECI>2.0.CO:2.
- Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. Hu, and K. E. Trenberth, 2011: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat. Climate Change*, 1, 360–364, https://doi.org/10.1038/nclimate1229.
- Menne, M. J., C. N. Williams, and R. S. Vose, 2009: The U.S. Historical Climatology Network monthly temperature data, version 2. *Bull. Amer. Meteor. Soc.*, 90, 993–1008, https://doi.org/10.1175/2008BAMS2613.1.
- —, C. N. Williams Jr., and R. S. Vose, 2015: United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, accessed 2 February 2015, https:// cdiac.ess-dive.lbl.gov/epubs/ndp/ushcn/ushcn.html.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.
- Mueller, N. D., E. E. Butler, K. A. McKinnon, A. Rhines, M. Tingley, N. M. Holbrook, and P. Huybers, 2015: Cooling of US Midwest summer temperature extremes from cropland intensification. *Nat. Climate Change*, 6, 317–322, https:// doi.org/10.1038/nclimate2825.
- Mutiibwa, D., S. J. Vavrus, S. A. McAfee, and T. P. Albright, 2015: Recent spatiotemporal patterns in temperature extremes across conterminous United States. *J. Geophys. Res. Atmos.*, **120**, 7378–7392, https://doi.org/10.1002/2015JD023598.
- NCDC, 2011: U.S. heat stress index. Daily maximum and minimum apparent temperature for 187 stations, National Climatic Data Center, accessed 5 February 2015, https://www.ncdc.noaa.gov/societal-impacts/heat-stress/data.
- Oswald, E. M., 2018: An analysis of the prevalence of heat waves in the United States between 1948 and 2015. *J. Climate Appl. Meteor.*, **57**, 1535–1549, https://doi.org/10.1175/JAMC-D-17-0274.1.
- —, and R. B. Rood, 2014: A trend analysis of the 1930–2010 extreme heat events in the continental United States. J. Appl. Meteor. Climatol., 53, 565–582, https://doi.org/10.1175/JAMC-D-13-071.1.
- Peng, R. D., J. F. Bobb, C. Tebaldi, L. McDaniel, M. L. Bell, and F. Dominici, 2011: Toward a quantitative estimate of future heatwave mortality under global climate change. *Environ. Health Perspect.*, 119, 701–706, https://doi.org/10.1289/ehp.1002430.

- Perkins, S. E., and L. V. Alexander, 2013: On the measurement of heatwaves. *J. Climate*, **26**, 4500–4517, https://doi.org/10.1175/ JCLI-D-12-00383.1.
- Peterson, T. C., X. Zhang, M. Brunet-India, and J. L. Vázquez-Aguirre, 2008: Changes in North American extremes derived from daily weather data. *J. Geophys. Res.*, 113, D07113, https://doi.org/10.1029/2007JD009453.
- Petri, Y., and K. Caldeira, 2015: 2015: Impacts of global warming on residential heating and cooling degree-days in the United States. *Sci. Rep.*, **5**, 12427, https://doi.org/10.1038/srep12427.
- Raymond, C., D. Singh, and R. Horton, 2017: Spatiotemporal patterns and synoptics of extreme wet-bulb temperature in the contiguous United States. *J. Geophys. Res. Atmos.*, 122, 13108–13124, https://doi.org/10.1002/2017jd027140.
- Rhines, A., and P. Huybers, 2013: Frequent summer temperature extremes reflect changes in the mean, not the variance. *Proc. Natl. Acad. Sci. USA*, **110**, E546, https://doi.org/10.1073/pnas.1218748110.
- Riahi, K., and Coauthors, 2011: RCP 8.5—Exploring the consequence of high emission trajectories. *Climatic Change*, **109**, 33–57, https://doi.org/10.1007/s10584-011-0149-y.
- Russo, S., and Coauthors, 2014: Magnitude of extreme heatwaves in present climate and their projection in a warming world. *J. Geophys. Res. Atmos.*, **119**, 12 500–12 512, https://doi.org/10.1002/2014jd022098.
- —, J. Sillmann, and E. M. Fischer, 2015: Top ten European heatwaves since 1950 and their occurrence in the coming decades. *Environ. Res. Lett.*, 10, 124003, https://doi.org/10.1088/ 1748-9326/10/12/124003.
- Schär, C., P. L. Vidale, D. Luthi, C. Frei, C. Haberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, 427, 332–336, https://doi.org/10.1038/nature02300.
- Scott, M. J., J. A. Dirks, and K. A. Cort, 2008: The value of energy efficiency programs for US residential and commercial buildings in a warmer world. *Mitig. Adapt. Strategies Global Change*, 13, 307–339, https://doi.org/10.1007/s11027-007-9115-4.
- Seneviratne, S. I., M. G. Donat, B. Mueller, and L. V. Alexander, 2014: No pause in the increase of hot temperature extremes. *Nat. Climate Change*, **4**, 161–163, https://doi.org/10.1038/nclimate2145.
- Sheridan, S. C., and L. S. Kalkstein, 2004: Progress in heat watch-warning system technology. *Bull. Amer. Meteor. Soc.*, 85, 1931–1942, https://doi.org/10.1175/BAMS-85-12-1931.
- —, and P. G. Dixon, 2017: Spatiotemporal trends in human vulnerability and adaptation to heat across the United States. *Anthropocene*, **20**, 61–73, https://doi.org/10.1016/j.ancene.2016.10.001.
- Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part II. Future climate projections. *J. Geophys. Res.* Atmos., 118, 2473–3493, https://doi.org/10.1002/jgrd.50188.
- Sippel, S., J. Zscheischler, M. Heimann, F. E. L. Otto, J. Peters, and M. D. Mahecha, 2015: Quantifying changes in climate variability and extremes: Pitfalls and their overcoming. *Geophys. Res. Lett.*, 42, 9990–9998, https://doi.org/10.1002/2015GL066307.
- Smith, T. T., B. F. Zaitchik, and J. M. Gohlke, 2013: Heatwaves in the United States: Definitions, patterns and trends. *Climatic Change*, **118**, 811–825, https://doi.org/10.1007/s10584-012-0659-2.
- Steadman, R. G., 1984: A universal scale of apparent temperature. *J. Climate Appl. Meteor.*, **23**, 1674–1687, https://doi.org/10.1175/1520-0450(1984)023<1674:AUSOAT>2.0.CO;2.

- Stegall, S. T., and K. E. Kunkel, 2017: Monthly extreme temperature trends in CMIP5 hindcast/prediction simulations, 1981–2010 and 2006–35. *J. Appl. Meteor. Climatol.*, **56**, 1141–1154, https://doi.org/10.1175/JAMC-D-16-0281.1.
- Tan, J. G., Y. F. Zheng, G. X. Song, L. S. Kalkstein, A. J. Kalkstein, and X. Tang, 2007: Heatwave impacts on mortality in Shanghai, 1998 and 2003. *Int. J. Biometeor.*, 51, 193–200, https://doi.org/10.1007/s00484-006-0058-3.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.
- Teng, H., G. Branstator, H. Wang, G. A. Meehl, and W. M. Washington, 2013: Probability of US heat waves affected by a subseasonal planetary wave pattern. *Nat. Geosci.*, 6, 1056–1061, https://doi.org/10.1038/ngeo1988.
- Teskey, R., T. Wertin, I. Bauweraerts, M. Ameye, M. A. McGuire, and K. Steppe, 2014: Responses of tree species to heat waves and extreme heat events. *Plant Cell Environ.*, **38**, 1699–1712, https://doi.org/10.1111/pce.12417.
- Thiery, W., E. L. Davin, D. M. Lawrence, A. L. Hirsch, M. Hauser, and S. I. Seneviratne, 2017: Present-day irrigation mitigates heat extremes. *J. Geophys. Res.*, **122**, 1403–1422, https://doi.org/10.1002/2016JD025740.
- Thomson, A. M., and Coauthors, 2011: RCP4.5: A pathway for stabilization of radiative forcing by 2100. *Climatic Change*, **109**, 77–94, https://doi.org/10.1007/s10584-011-0151-4.
- Tollefson, J., 2014: Climate change: The case of the missing heat. *Nature*, **505**, 276–278, https://doi.org/10.1038/505276a.
- Trenberth, K. E., J. T. Fasullo, G. Branstator, and A. S. Phillips, 2014: Seasonal aspects of the recent pause in surface warming. *Nat. Climate Change*, **4**, 911–916, https://doi.org/10.1038/nclimate2341.
- Van Den Dool, H. M., 1987: An empirical study on the parameterization of precipitation in a model of the time mean atmosphere. J. Atmos. Sci., 44, 224–235, https://doi.org/10.1175/1520-0469(1987)044<0224:AESOTP>2.0.CO;2.
- Vicedo-Cabrera, A. M., and Coauthors, 2018: A multi-country analysis on potential adaptive mechanisms to cold and heat in a changing climate. *Environ. Int.*, 111, 239–246, https:// doi.org/10.1016/j.envint.2017.11.006.

- Vose, R. S., C. N. Williams Jr., T. C. Peterson, T. R. Karl, and D. R. Easterling, 2003: An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophys. Res. Lett.*, 30, 2046, https://doi.org/10.1029/2003GL018111.
- —, S. Applequist, M. J. Menne, C. N. Williams Jr., and P. Thorne, 2012: An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses. *Geophys. Res. Lett.*, 39, L10703, https:// doi.org/10.1029/2012GL051387.
- —, D. R. Easterling, K. E. Kunkel, A. N. LeGrande, and M. F. Wehner, 2017: Temperature changes in the United States. *Climate Science Special Report*, Vol. I, *Fourth National Climate Assessment*, D. J. Wuebbles et al., Eds., U.S. Global Change Research Program, 185–206, https://doi.org/10.7930/J0N29V45.
- Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335, https://doi.org/10.1175/2010JCLI3679.1.
- Wu, T., and Coauthors, 2014: An overview of BCC Climate System Model development and application for climate change studies. J. Meteor. Res., 28, 34–56, https://doi.org/ 10.1007/s13351-014-3041-7.
- Wuebbles, D., and Coauthors, 2014: CMIP5 climate model analyses: Climate extremes in the United States. *Bull. Amer. Meteor.* Soc., 95, 571–583, https://doi.org/10.1175/BAMS-D-12-00172.1.
- Wuebbles, D. J., D.W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock, Eds., 2017: Climate Science Special Report. Vol. I, Fourth National Climate Assessment, U.S. Global Change Research Program, 470 pp., https://doi.org/10.7930/J0J964J6.
- Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, 117, D03109, https://doi.org/10.1029/2011JD016048.
- Yiou, P., D. Dacunha-Castelle, S. Parey, and T. T. Huong Hoang, 2009: Statistical representation of temperature mean and variability in Europe. *Geophys. Res. Lett.*, 36, L04710, https:// doi.org/10.1029/2008GL036836.