Projected Increase in the Spatial Extent of Contiguous U.S. Summer Heat Waves and Associated Attributes

Bradfield Lyon, Anthony G. Barnston, Ethan Coffel, and Radley M. Horton

Supplemental Information

1. Model performance characteristics

Table S1 shows the equilibrium climate sensitivity (ECS) and transient climate response (TCR) for the 11 CMIP5 models used in the study based on the data provided by Flato et al. (2013).

Table S1. CMIP5 model characteristics

Model	ECS (°C)	TCR (°C)
ACCESS1-0	3.8	2.0
ACCESS1-3	N/A	1.7
BCC-CSM1-1-M	2.9	2.1
BNU-ESM	4.1	2.6
CANESM-2	3.7	2.4
CNRM-CM5	3.3	2.1
CSIRO-MK3-6-0	4.1	1.8
IPSL-CM5A-MR	N/A	2.0
IPSL-CM5B-LR	2.6	1.5
MIROC5	2.7	1.5
NORESM1-M	2.8	1.4
Average	**3.3**	**1.9**
30-model average from Flato et al (2013)	*3.2*	*1.8*

2. Biases in CMIP5 surface variables

For the May-September season, biases in climatological mean daily maximum and minimum temperature, specific humidity and surface pressure were computed across the 11 CMIP5. These biases were defined as the difference between the climatological mean values for the 1980-2005 period in the CMIP5 models and 1) the University of East Anglia maximum and minimum surface air temperature, and 2) the NCEP-DOE Reanalysis 2m specific humidity and surface pressure (all variables were first re-gridded to a common 2.0° × 2.0° lat./lon. grid). The multi-model mean results are shown below in Figure S1. The most distinctive feature for temperature bias is for daily maximum temperature with a pronounced warm bias in the central U.S. (Fig. S1a). The multi-model mean bias is lower for daily minimum temperature (Fig. S1b), with a positive bias extending across the northern Plains. Biases in surface pressure (Fig. S1c) do

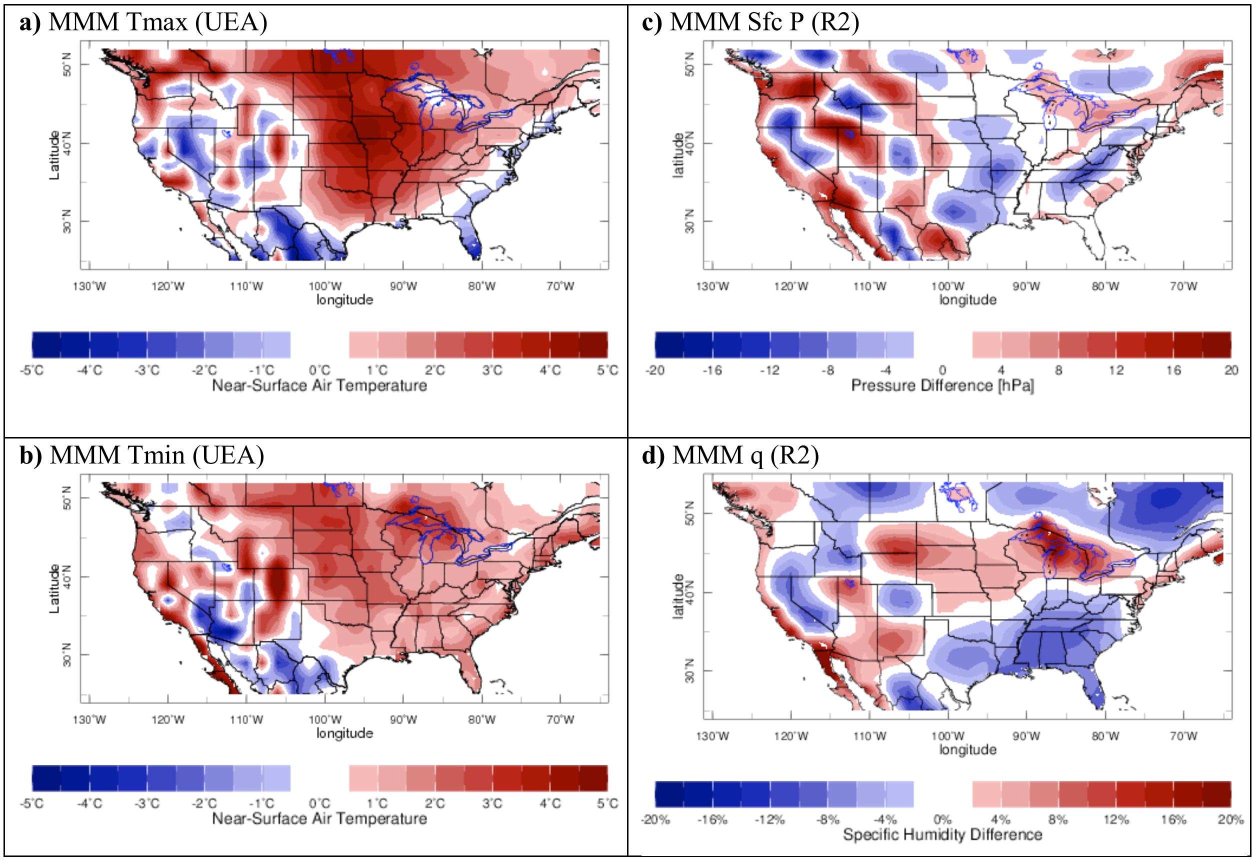

Figure S1. *CMIP5 multi-model mean (MMM) bias in May-September season (a) daily maximum temperature (deg. C) and (b) daily minimum temperature (deg. C), both relative to UEA gridded analyses along with (c) surface pressure (hPa) and (d) specific humidity, computed relative to NCEP Reanalysis 2 data. All biases are computed as differences in climatological values for the 1980-2005 period.*

not exhibit a spatially coherent pattern and are typically less than 1%. Specific humidity biases (Fig. S1d) are typically less than ± 8%, the multi-model mean having slightly too dry conditions in the southeastern U.S. being the most prominent feature. When computing cooling degree days, the monthly mean bias in daily maximum and minimum temperature for each model was first removed from the respective CMIP5 daily temperature data at each grid point. Other heat wave indices are based on percentiles computed relative to a specific model, without bias correction.

3. Estimated errors in computing daily maximum apparent temperature

As indicated in the main text, the computation of apparent temperature (T_a) is based on the work of Steadman (1984) and may be written as $T_a = -1.3 + 0.92T + 2.2e$, where T is the 2m air temperature (°C) and e is the vapor pressure (kPa). The vapor pressure is computed as $e \cong 1.608\times10^{-3}\, p \cdot q$, where p is the surface pressure (Pa) and q is the specific humidity (g/kg). This formulation tends to slightly underestimate (by < 1%) the vapor pressure for typical values of pressure and specific humidity found over the U.S. in summer, which translates to a decrease in apparent temperature of roughly 0.1-0.2 °C. This approximate formulation for vapor pressure was used rather than an exact relationship, as a necessary input variable to compute the latter was not available.

A slightly larger error is introduced in the daily maximum T_a when the vapor pressure is computed based on daily average values of specific humidity rather than the specific humidity at the time of maximum temperature. Only daily average values of specific humidity are available for the CMIP5 models, for example. To estimate the magnitude of this error, two specific humidity files were generated based on the 3-hourly NARR data. The first file contains daily values of the specific humidity at the time of maximum surface air temperature, the second file contains daily mean values of specific humidity. The year 2001 was chosen at random to compute the mean absolute error (MAE) represented as the absolute difference of these two files at each grid point for each day. The average of the difference in these two files was for the period 1 May – 30 September. The results are shown below in Figure S2. The MAE averaged across the entire domain in Figure S is 7.9%. For typical values of surface pressure and specific humidity, an 8% absolute error in the latter variable corresponds to a change in the apparent temperature of about ± 0.2-0.3 °C, which is not expected to substantially effect the overall results presented.

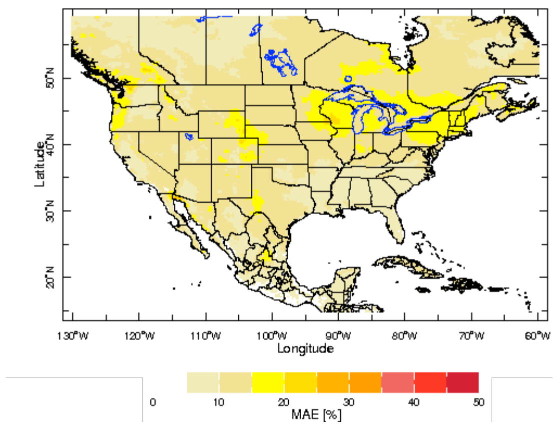

Figure S2. *The mean absolute error (in percent) computed as the difference between the daily near surface specific humidity at the time of maximum temperature and the daily average specific humidity, both based on 3-hourly data from NARR. The average is computed for daily values running from 1 May – 30 September 2001.*

4. Comparison of apparent temperature and the Heat Index of the U.S. National Weather Service

Apparent temperature is used in this study as it represents the perceived temperature felt by humans given the combination of heat and humidity and is thus a measure of "thermal comfort." Its computation involves the use of surface air temperature and vapor pressure in a multiple linear regression fit to apparent temperature values obtained using the more complex formulation by Steadman (1984) under the assumptions of shady outdoor conditions and no wind. The multiple linear regression used here was developed by Steadman (1984) who labeled its use as an "Indoors" indicator (his Table 5) because it did not explicitly incorporate wind and solar radiation. However, in this study the regression formula is used to compute an outdoor measure of apparent temperature, just under the assumption of being in the shade and where the wind does not vary. Steadman (1984) reports that this multiple linear regression approximation of apparent temperature has a residual standard deviation of 0.32 °C.

Another measure of human thermal comfort that is in widespread use in the U.S. is the Heat Index of the National Weather Service (Rothfusz 1990). However, the Heat Index is also computed as a statistical fit to apparent temperature, in this case using surface air temperature and relative humidity in a non-linear multiple regression to approximate the published apparent temperature values of Steadman (1979) (his Table 2) based on a more complex formulation and additional inputs. The apparent temperature values that the Heat Index is fit to are derived under the assumptions of being in the shade with no variability in wind speed (Steadman 1979), similar to the "Indoors" multiple linear regression fit of Steadman (1984) used in this study. As reported by Rothfusz (1990), the Heat Index approximation to full apparent temperature calculation of Steadman (1979) has an error of ±1.3 °F (0.72 °C), roughly twice the reported error when approximating the apparent temperature using the multiple linear regression formula in this study.

Although strongly related, comparisons of daily values of apparent temperature from the multiple linear regression and the Heat Index were made for the arbitrarily selected month of July 1980. The comparisons were made for 9 stations (Table S2) included in the U.S. Historical Climatology Network (USHCN), one station selected randomly from each of the 9 U.S. climate

Table S2. USHCN stations used for Heat Index and apparent temperature comparisons

Station	Elevation	Climate Region
Brewton, AL	25m	Southeast
Brinkley, AR	64m	South
Tucson, AZ	728m	Southwest
Colfax, CA	739m	West
New Hampton, IA	353m	East North Central
Oakley, ID	1393m	Northwest
Goshen, IN	244m	Central
Portland, ME	19m	Northeast
Aberdeen, SD	396m	West North Central

regions identified by Karl and Koss (1984). As in Lyon and Barnston (2017), daily maximum temperature values at these 9 stations were combined with daily humidity data taken from the NARR (at the nearest grid point to the station) to compute both apparent temperature and the Heat Index. The results are shown in Figure S3, where the Heat Index is plotted against the apparent temperature values as calculated in this study. A linear regression fit is shown for each station in Fig. S3, with the associated linear correlation coefficient shown in the upper-left of each panel, along with the standard error of estimate in computing the Heat Index from apparent temperature. Given the strong connection between the Heat Index and apparent temperature, the results reveal linear correlation coefficients ≥ 0.98 across all 9 stations, which provide a wide range of temperature and humidity conditions. Averaged over the 9 stations, the standard error of estimate is 0.42 °C, considerably smaller than the error of 0.72 °C found by Rothfusz (1990) in approximating the apparent temperature from the Heat Index.

Overall, these comparisons indicate that the maximum daily apparent temperature (as computed in this study) is strongly related to the Heat Index, as expected. The use of apparent temperature in this study has the advantage of avoiding the need to make "adjustments" when computing the Heat Index based on the absolute temperature and humidity thresholds identified by Rothfusz (1990). Using these absolute threshold values in model data would first require making bias adjustments to the model output before computing the Heat Index. As no similar adjustments are required in our computation of apparent temperature and since percentiles (not absolute thresholds) are used in defining heat waves, the need for bias correction in the humidity and temperature fields is avoided. For researchers analyzing specific impacts of heat waves (e.g., mortality), use of the Heat Index should be considered over our simplified formulation of apparent

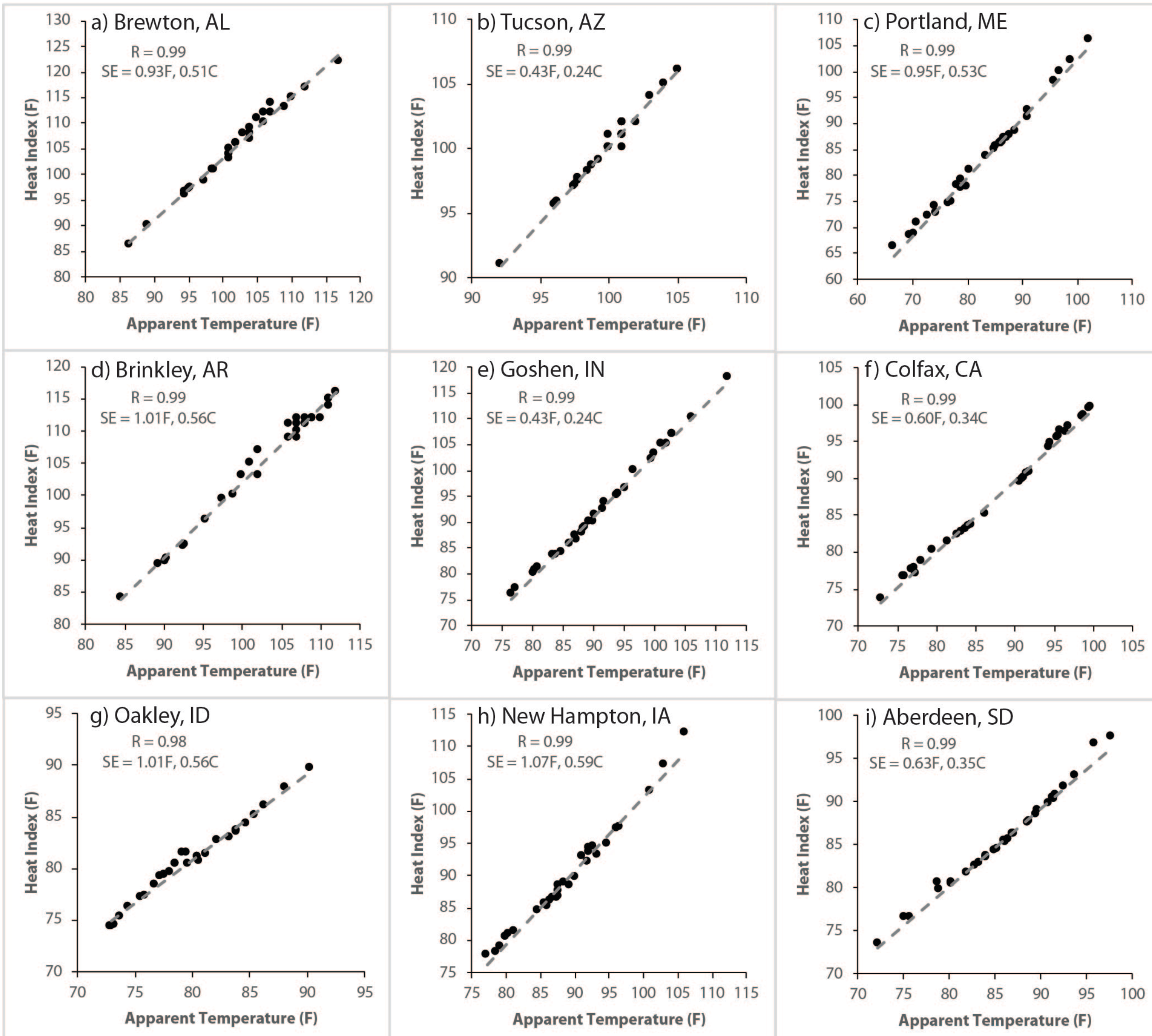

Figure S3. *Daily values (℉) of the Heat Index (vertical-axis) vs. apparent temperature (horizontal axis) for 1-31 July 1980 at nine stations, one randomly selected from each of the 9 climate regions identified by the U.S. Climate Prediction Center. Dashed lines indicate linear regression fits for each station, with correlations and standard error of estimate of the Heat Index from apparent temperature indicated in the upper-left of each panel (in both ℉ and ℃).*

temperature given that the former shows a greater sensitivity to extreme temperature and humidity conditions.

5. Comparison of two apparent temperature estimates

Steadman (1984) performed different multiple regressions to estimate his full apparent temperature calculation from a limited set of inputs. The multiple linear regression formulation used in this study is $T_a = -1.3 + 0.92T + 2.2e$, where Ta is the apparent temperature (deg. C), T the air temperature (deg. C) and e the vapor pressure (kPa). As noted previously, this formulation is labeled as "Indoors" in Steadman (1984) but is used here for outdoor conditions under the assumption of shady conditions with no wind. Steadman (1984) included other multiple regression estimates of apparent temperature. One estimate is labeled "Shade" and is given by the equation $T_a = -2.7 + 1.04T + 2.0e – 0.65v$, where the variables are as defined previously with the additional variable, v, being the wind speed (m/s) at 10 meters above the ground. The residual standard deviation (or error) of this estimate to the full apparent temperature calculation is reported as 0.44 C. Under our assumption of no wind, the variable v could be set to zero in the above equation leaving the alternative estimation of apparent temperature: $T_a = -2.7 + 1.04T + 2.0e$. Given the labeling used by Steadman (1984) of his regression equations, the use of this formula could be argued to be more appropriate to this study.

We thus compared apparent temperature estimates obtained from the "Indoors" and "Shade" equations for a randomly selected set of 25 pairs of temperature and humidity values which covered a wide range of respective conditions (approximately 25-50 deg. C and 10-80% relative humidity). The results are shown in Figure S4, where a least squares linear fit is also

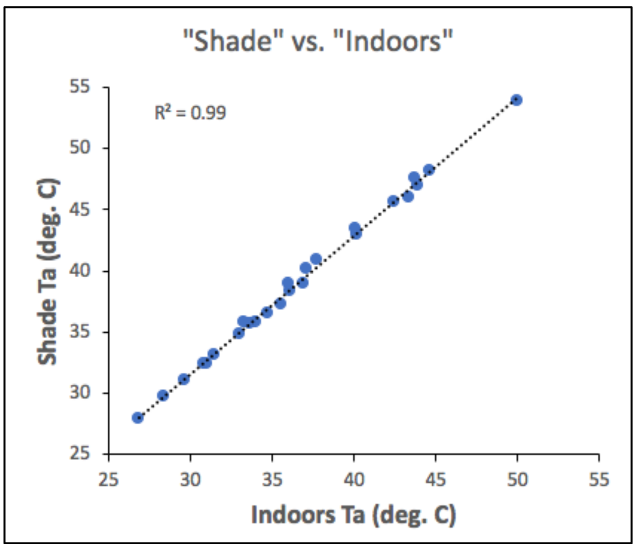

Figure S4. *Apparent temperature estimates (℃) for a randomly selected set of temperature and humidity (vapor pressure) conditions. A least squares linear fit is also shown by the dashed line.*

shown along with the associated explained variance. The linear regression fit shows a near perfect correspondence between the two estimates ($R^2 = 0.996$). We also computed the Heat Index for the above conditions and found that the mean absolute error in approximating the Heat Index with the Indoors formulation is 2.4 C, while for the Shade formulation the error is 2.2 C. These results indicate that using the Shade formulation for apparent temperature in our analysis would have yielded virtually identical results to those obtained using the Indoors formulation. We note, however, that for this study heat waves are defined using percentiles. Researchers interested in analyzing absolute values of apparent temperature should again consider using the Heat Index.

6. Projected mean temperature change and normalized change in CMIP5 models

Maps of the change in mean daily surface air temperature for the May-September season (2031-2055 minus 1980-2005) for the RCP8.5 and RCP4.5 scenarios are shown in the left-hand side of Figure S5. The right-hand side of the figure shows this temperature change divided by the multi-model mean, average monthly standard deviation of May-September temperatures in the current climate (1980-2005) to indicate the magnitude of the projected change relative to the interannual variability of the current climate.

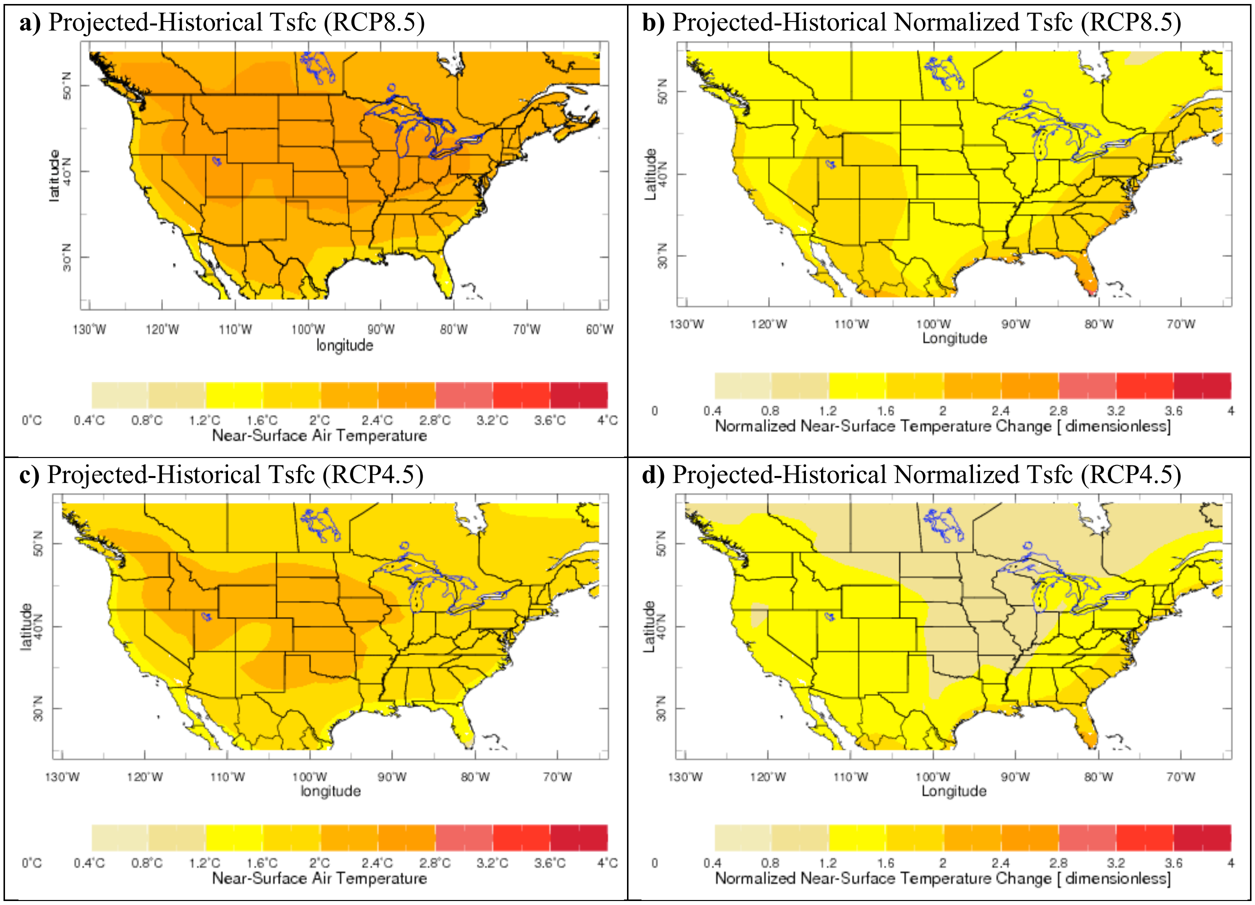

Figure S5. *Multi-model mean temperature change (deg. C) for the May-September season (2031-2055 minus 1980-2005) for the 11 coupled climate models based on the RCP8.5 and RCP4.5 greenhouse gas forcing scenarios (a) and (c). The change in May-September mean temperature divided by the historical (1980-2005) monthly average standard deviation across models for RCP8.5 (b) and RCP4.5 (d).*

7. Comparisons of heat wave days in NARR and CMIP5 historical runs

The number of heat wave days (> 95th percentile for 3 or more days) in contiguous regions was compared for the NARR data (1979-2009) and the multi-model mean CMIP5 historical runs (1980-2005). Three variables were considered: the daily maximum, daily maximum apparent temperature and the daily mean temperature. Results are shown in Figure S6 and show a generally good correspondence for daily maximum and daily maximum apparent temperatures (the pattern correlation is plotted on the lower-right of model results). The largest difference for daily minimum temperature is in the south-central U.S., centered on Texas. A possible contributing factor is that in observations there were several drought years during this period not expected to be captured in the 11-model average. Drought tends to favor extreme maximum temperature (Mazdiyasni and AghaKouchak 2015; Durre et al 2000) over minimum temperatures in this region.

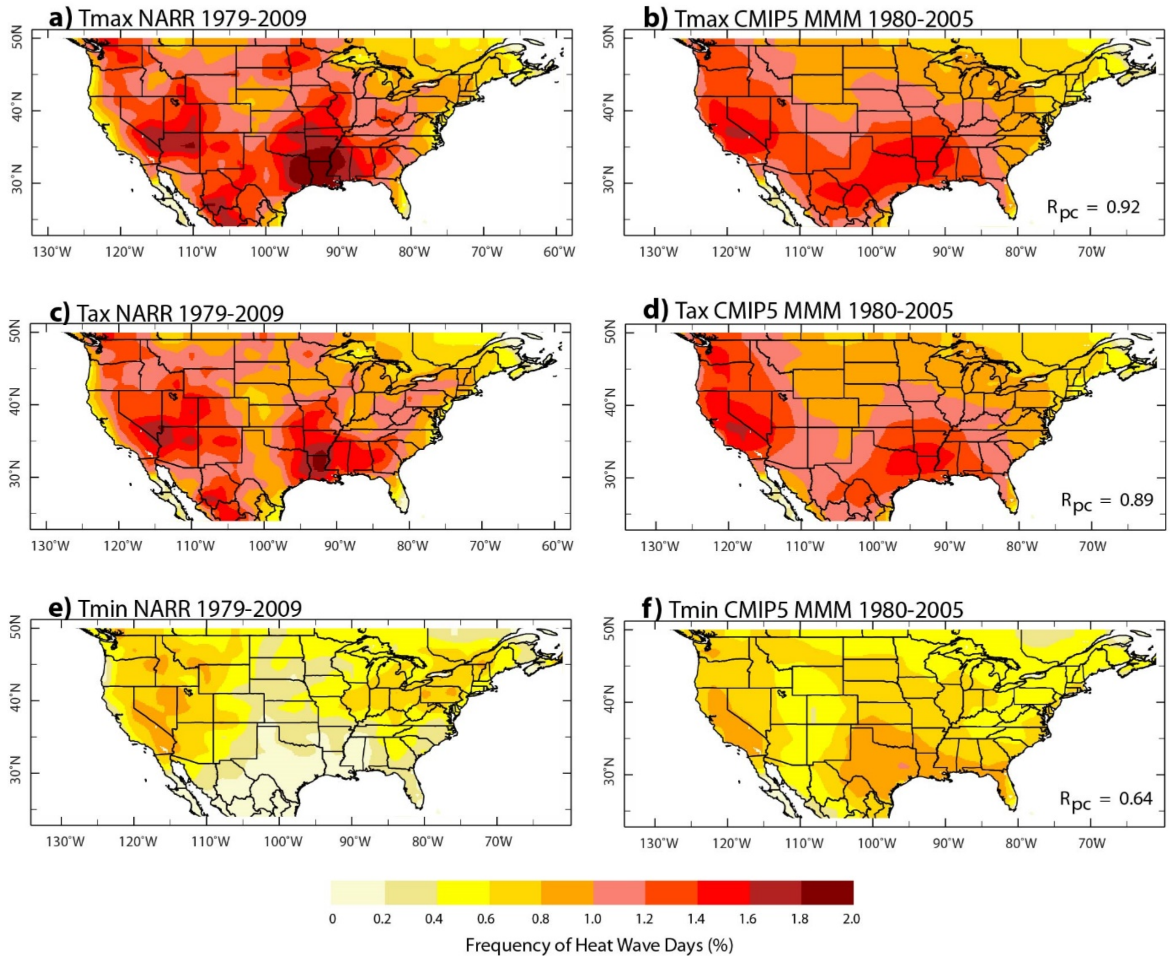

Figure S6. *Percent of warm season (May-September) days in contiguous heat waves (>95th percentile for 3 or more consecutive days) for NARR for 1979-2009 (left-column) and CMIP5 multi-model mean (MMM) historical runs for 1980-2005 (right column). (a) and (b) are for daily maximum temperature, (c) and (d) daily maximum apparent temperature and (e) and (f) daily minimum temperature.*

8. NARR, CMIP5 historical and RCP4.5 heat wave attributes

The heat wave attributes shown in the main text in Figures 3 and 4 were also computed based on output from the CMIP5 models forced with the RCP4.5 scenario. For this scenario, the available daily temperature files for the CANESM2 model were found to contain multiple errors, so this model was excluded from the analysis with results shown for the remaining 10 models. The results are presented in Figure S7. Overall, the pattern of projected attribute changes for the

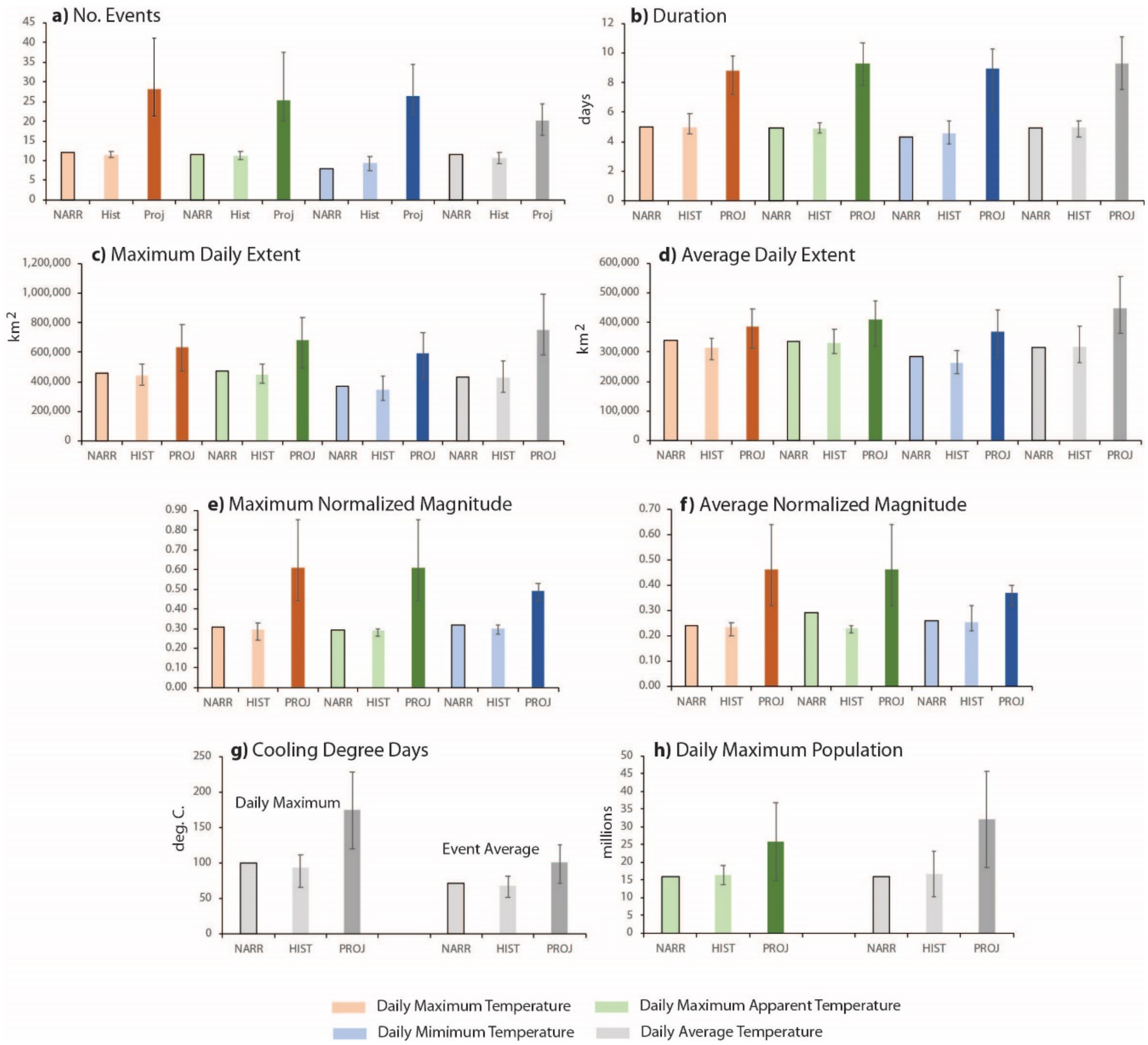

Figure S7. *Heat wave attributes for four temperature variables, with results for NARR, CMIP5 historical runs (light bars) and RCP4.5 projections (darker bars). Whiskers indicate the range in average values across the 11 models in historical runs (1980-2005) and projections (2031-2055). (a) Average number of heat wave events, (b) average duration (days), (c) maximum, and (d) average, daily extent (km^2), (e) maximum, and (f) average, normalized heat wave magnitude, g) daily maximum and daily average cooling degree days, h) daily maximum exposed population.*

RCP4.5 scenario is quite similar to that for the RCP8.5 case, but the magnitude of the changes is generally smaller. It is interesting that for the RCP4.5 scenario, the daily maximum and average normalized magnitude is similar or even slightly larger than for the RCP8.5 scenario. One possible contributor here is that since heat waves are essentially random weather events, in the two sets of model runs they are occurring in somewhat different regions and thus have different normalized values. A few, extreme events may also skew the results towards higher values in the RCP4.5 runs. This may also contribute to the fact that more heat wave events are identified under the RCP4.5 scenario than for RCP8.5. Another possibility for the latter result is that under the higher radiative forcing of the RCP8.5 scenario, heat wave events are found to persist longer than for RCP4.5 and as such, there are fewer “opportunities” for heat waves to develop under RCP8.5 forcing vs. RCP4.5.

9. Average heat wave attributes comparison

Table S2. Observed, Historical and Projected (RCP8.5) Contiguous Heat Wave Region Attributes. Percent difference between multi-model mean CMIP5 historical runs and NARR shown in parentheses with the ratio of projected to historical run values shown in bold.

		T_x		T_{ax}		T_n		T_m
No. Events								
NARR		12.1		11.5		8.0		11.5
Historical	(-5.0%)	11.5	(-2.6%)	11.2	(17.5%)	9.4	(-7.8%)	10.6
Projected		22.3		19.9		22.4		19.1
Ratio (projected/historical)		**1.94**		**1.78**		**2.38**		**1.81**
Duration (days)								
NARR		5.0		4.9		4.3		4.9
Historical	(0.0%)	5.0	(0.0%)	4.9	(7.0%)	4.6	(2.0%)	5.0
Projected		9.0		10.1		9.3		10.1
Ratio (projected/historical)		**1.82**		**2.05**		**2.03**		**2.04**
Maximum Daily Extent (km^2)								
NARR		457,718		473,773		368,648		433,543
Historical	(-3.7%)	440,731	(-4.9%)	450,383	(-5.6%)	348,092	(-1.4%)	427,514
Projected		771,883		838,974		638,550		791,598
Ratio (projected/historical)		**1.75**		**1.86**		**1.83**		**1.85**
Average Daily Extent (km^2)								
NARR		338,171		333,807		285,527		314,628
Historical	(-7.0%)	314,535	(-0.7%)	331,316	(-8.8%)	260,266	(0.7%)	316,784
Projected		431,922		490,358		358,008		464,345
Ratio (projected/historical)		**1.37**		**1.48**		**1.38**		**1.47**
Maximum Normalized Magnitude								
NARR		0.31		0.29		0.32		--
Historical	(-3.2%)	0.30	(0.0%)	0.29	(-6.3%)	0.30		--
Projected		0.54		0.54		0.51		--
Ratio (projected/historical)		**1.83**		**1.86**		**1.68**		--
Average Normalized Magnitude								
NARR		0.24		0.29		0.26		--
Historical	(-4.2%)	0.23	(-20.7%)	0.23	(0.0%)	0.26		--
Projected		0.41		0.40		0.38		--
Ratio (projected/historical)		**1.74**		**1.75**		**1.51**		--
Maximum Cooling Degree Days								
NARR		--		--		--		99.1
Historical		--		--		--	(-5.4%)	93.7
Projected		--		--		--		178.6
Ratio (projected/historical)		--		--		--		**1.91**
Average Cooling Degree Days								
NARR		--		--		--		71.4
Historical		--		--		--	(-4.8%)	68.0
Projected		--		--		--		100.9
Ratio (projected/historical)		--		--		--		**1.48**
Maximum Daily Population (10^6 people)								
NARR		--		16.1		--		15.9
Historical		--	(1.9%)	16.4		--	(5.0%)	16.7
Projected		--		34.8		--		34.6
Ratio (projected/historical)		--		**2.12**		--		**2.07**

10. Top 10% heat wave attributes comparison

We examined more severe heat waves to estimate what might be the upper-bound to projected changes in heat wave attributes. We used the 90th percentile of attribute values to identify such events and for projections we used the RCP8.5 forcing scenario. Figure S8 shows the ratio of the projected (RCP8.5) to historical values for all heat wave attributes above the 90th percentile, with the associated data values given in Table S3. For all four temperature variables, the maximum spatial extent and duration more than double in mid-century projections. This equates to a spatial extent of contiguous heat wave regions that exceeds 2.0 million km^2 for 90th percentile events (for reference, the area of the 48 contiguous U.S. states is about 8.1 million km^2). The daily maximum cooling degree days increase by a factor of 2.3 over the historical 90th percentile values, with the exposed population also doubling over historical 90th percentile values. All other attributes increase by a factor > 1.5 over their historical period values for the high-concentration RCP8.5 scenario.

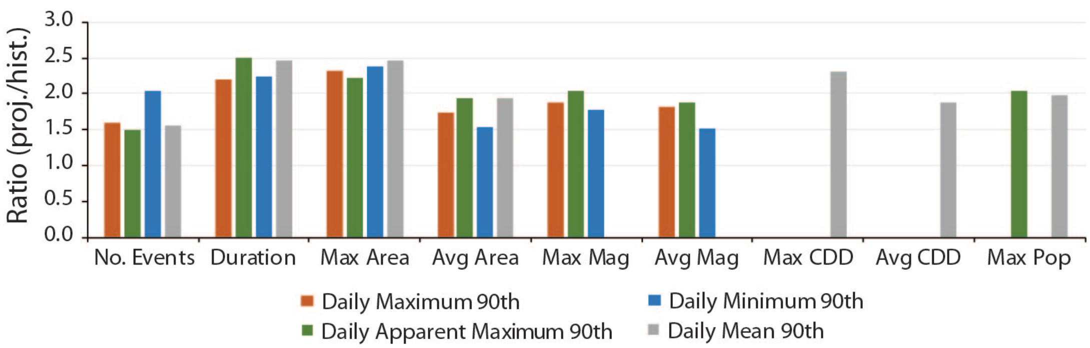

Figure S8. *Ratio of the multi-model mean projected (RCP8.5) 90th percentile heat waver attribute values to their respective 90th percentile historical values in the CMIP5 models.*

11. The 90th percentile heat wave attributes comparison

Table S3. NARR and Multi-model Mean 90th Percentile Heat Wave Attribute Values for Four Temperature Variables and Their Projected Changes (RCP8.5). The ratio of projected to historical run values is shown in bold font.

	Maximum	Apparent Max.	Minimum	Mean
No. Events				
NARR	19.0	18.0	12.0	18.0
Historical	18.1	18.0	16.0	16.7
Projected	29.0	26.7	32.6	26.1
Ratio (proj./hist.)	**1.60**	**1.48**	**2.04**	**1.56**
Duration (days)				
NARR	8.0	8.0	6.0	8.0
Historical	8.4	8.1	7.4	8.4
Projected	18.5	20.2	16.6	20.7
Ratio (proj./hist.)	**2.20**	**2.49**	**2.24**	**2.46**
Maximum Area (km^2)				
NARR	981,500	1,057,000	717,250	943,750
Historical	902,540	942,377	678,470	874,427
Projected	2,081,741	2,083,525	1,610,895	2,160,673
Ratio (proj./hist.)	**2.31**	**2.21**	**2.37**	**2.47**
Average Area (km^2)				
NARR	629,166	549,640	509,625	594,940
Historical	684,379	608,799	542,873	559,212
Projected	1,185,812	1,175,245	835,431	1,087,630
Ratio (proj./hist.)	**1.73**	**1.93**	**1.54**	**1.94**
Maximum Norm. Magnitude				
NARR	0.52	0.47	0.55	--
Historical	0.51	0.48	0.51	--
Projected	0.96	0.97	0.91	--
Ratio (proj./hist.)	**1.88**	**2.02**	**1.78**	--
Average Norm. Magnitude				
NARR	0.37	0.47	0.40	--
Historical	0.36	0.35	0.40	--
Projected	0.66	0.66	0.61	--
Ratio (proj./hist.)	**1.83**	**1.89**	**1.53**	--
Maximum CDD (°C)				
NARR	--	--	--	209.9
Historical	--	--	--	206.3
Projected	--	--	--	479.1
Ratio (proj./hist.)	--	--	--	**2.32**
Average CDD (°C)				
NARR	--	--	--	134.8
Historical	--	--	--	131.6
Projected	--	--	--	245.6
Ratio (proj./hist.)	--	--	--	**1.86**
Maximum Population ($\times 10^6$)				
NARR	--	53.2	--	55.6
Historical	--	53.5	--	55.3
Projected	--	108.6	--	109.3
Ratio (proj./hist.)	--	**2.03**	**--**	**1.98**

Supplemental Information References

Durre I, Wallace JMand LettenmaierDP 2000 Dependence of extreme daily maximum temperatures on antecedent soil moisture in the contiguous United States during summer J. Clim. 13 2641–51

Karl T R and KossWJ 1984 Regional and national monthly, seasonal, and annual temperature weighted by area, 1895–1983 Historical Climatology Series 4-3 (Asheville, NC: National Climatic Data Center) p38.