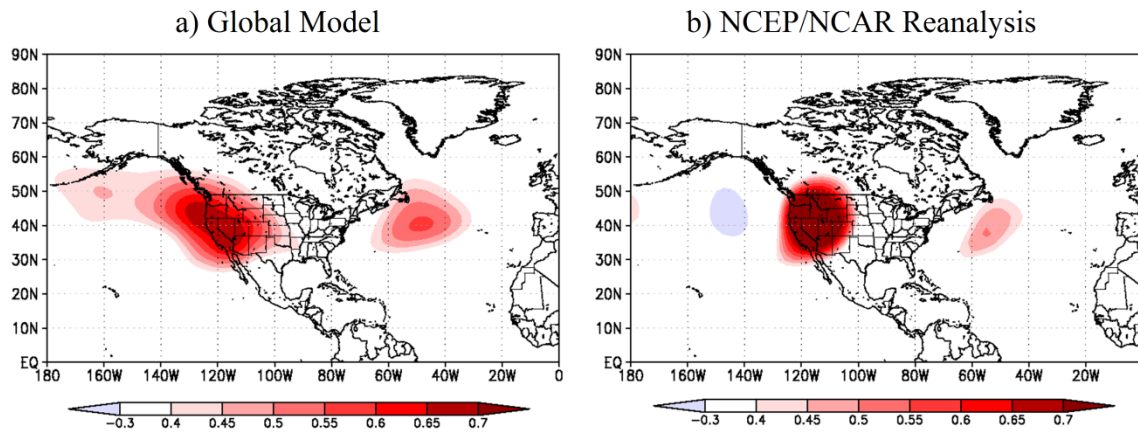**Supplementary materials**

Large Scale Circulation Patterns

The ability of HadAM3P to represent meteorological conditions in the western US is tested through a correlation analysis of geopotential heights over the western US on temperatures in the study region, similar to Sippel and Otto (2014). For this analysis, monthly mean values, averaged over JJA, of 150 global model ensemble members for the 1961-2010 period and National Center for Environmental Prediction / National Center for Atmospheric Research [(NCEP/NCAR) Kalnay et al., 1996] reanalysis data reveals that geopotential heights (m) over the western US both correlate similarly with temperatures (°C) over the same region (Supplementary Fig. 1).

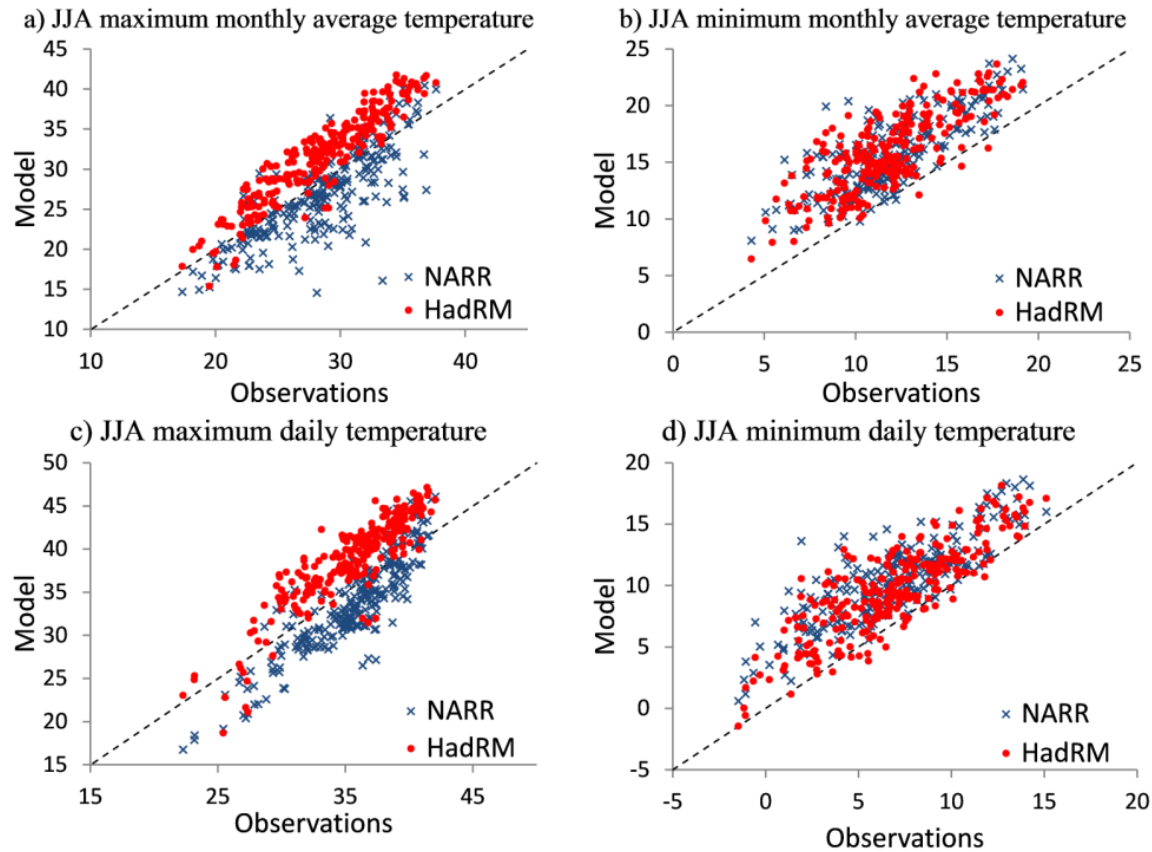a) Global Model   b) NCEP/NCAR Reanalysis

Supplementary Figure 1. Correlation coefficients between June-July-August temperatures in the western United States and geopotential heights in the (a) global model and (b) NCEP/NCAR Reanalysis.

Validation

Both the model and North American Regional Reanalysis (NARR: Mesinger et al. 2006) were compared against station observations for June-July-August (JJA) during the 1979-2010 period to test the superensemble's capability to approximate the observed climate. NARR (32 km resolution) was included in order to show the added skill of HadRM3P's resolution.

In Supplementary Figs. 2a-d, each point in the scatterplots represents the 1979-2010 average for a single month (June, July, August) for 125 locations throughout the western US, which includes the Pacific Northwest, providing 375 points of comparison.

Supplementary Figure 2. HadRM3P (red circles) and NARR (blue marks) versus observed June-July-August (JJA) (a) maximum monthly average temperature, (b) minimum monthly average temperature, (c) maximum daily temperature and (d) minimum daily temperature. Each point represents an individual month for each of the 125 stations for the 1979-2010 period.

Biases for HadRM3P are presented in Supplementary Table 1. The figures are consistent with findings by Duliere et al. (2011), where HadRM3P was generally better correlated than reanalysis to observations yet also exhibited a general warm bias.

Supplementary Table 1. Correlation coefficients and biases for the variables Tmax, Tmin, Tmax_hi, and Tmin_lo between observations and HadRM3P and NARR for June, July and August

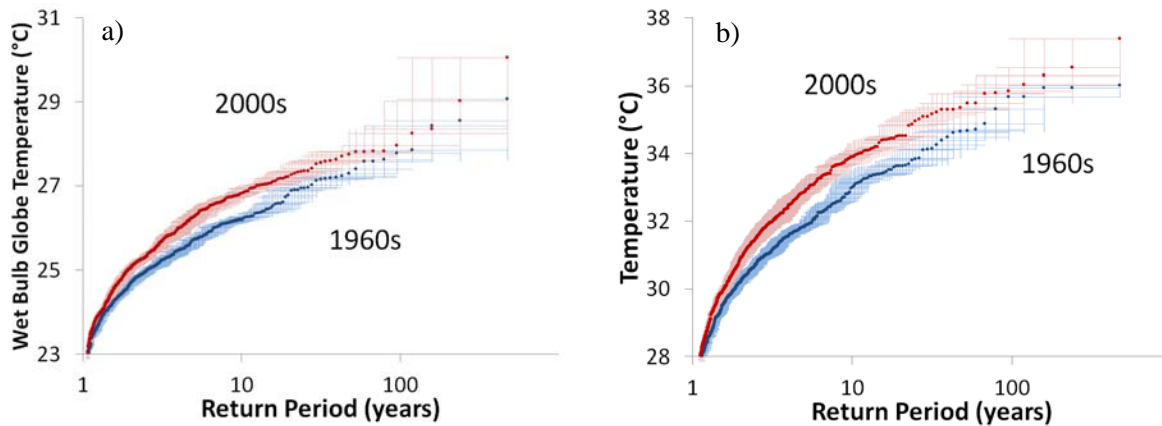|		Tmax	Tmin	Tmax_hi	Tmin_lo
HadRM3P	Correlation	0.85	0.75	0.89	0.84
	Bias	3.37	3.98	3.26	2.71
NARR	Correlation	0.74	0.70	0.86	0.77
	Bias	-2.02	3.05	-2.55	2.86
HadRM3P	Correlation	0.85	0.88	0.90	0.81
v NARR	Bias	5.80	0.92	-0.15	5.40

Impact of Moisture

We investigate the role of moisture in recent heat waves through the wet-bulb globe temperature (*WBGT*), a measure utilized by weather services to issue heat-health warnings:

$$\text{WBGT} = 0.567T + 0.393e + 3.94 \qquad (1)$$

where *T* is the air temperature and *e* denotes water vapor pressure (Fischer and Knutti 2012). PEA was applied to heat stress through the warmest 5-day mean *WBGT*, calculated for the summer months over the Central Valley using simulations from AF1.

According to output from the NARR, the warmest five-day period of the 2006 heat wave had a remarkable *WBGT* value of 29.3°C. Supplementary Fig. 3a suggests that anthropogenic greenhouse gas emissions influence the higher heat stress levels experienced during the 2000s when compared to a time period with lower GHG concentrations in the atmosphere (1960s) and that it is not limited to extremes. A once-in-a-decade event (10 year return period), for example, has a *WBGT* value that is about 0.6°C higher than the 1960s, with a corresponding FAR of 0.5.

In comparison with the *WBGT*, the warmest 5-day mean temperatures used to calculate the *WBGT* (Supplementary Fig. 3b), a 10-year event increased by 0.89°C. This difference in magnitude of change in the temperature and *WBGT* variables is partially due to the usually-dry environment in the Central Valley during the summer. The lower humidity levels partly counterbalance the effects of increasing 5-day temperatures on the human body. Unfortunately, there was not enough daily output to further scrutinize the humidity-heat link found in these humid heat waves, but a good proxy for the increased impact of recent heat waves is the strong nighttime component found in the simulations.

Supplementary Figure 3. Central Valley return periods of heat stress in WBGT (a) and warmest 5-day mean temperatures in the summer (b). Hatchings show bootstrapped inner 95% percentile uncertainty ranges.

Internal Variability
Internal variability of the model simulations is tested in order to see its contribution to the difference between the decades of the 1960s and 2000s. For this purpose, both the PDO and ENSO are examined in more detail for simulations occurring during the 2000s.

Supplementary Fig. 4 shows that the state of the PDO has very limited impact for either variable. Here, 4 different years were considered as PDO positive (2003, 2004, 2005, 2007) and 4 as PDO negative (2001, 2002, 2009, 2010). There was strong PDO positive during the middle of the decade and strong PDO negative at the beginning (2001) and end of the decade (2009). PDO positive was strong in 2003 and 2005.

In Supplementary Fig. 5, the impact of ENSO on Tmin_lo and Tmax_hi is measured. Three years were picked for El Niño (2002, 2004, 2009) and La Niña (2007, 2008, 2010). Note that the phase of ENSO in the analysis has the Niño 3.4 region anomalies concurrent with the summer season (JJA) being analyzed.

The ENSO phase has a larger effect on Tmin_lo than PDO. In fact, this effect is comparable to the overall change in return periods when comparing all years in the 1960s to the 2000s (Fig. 2 in main text). Here, a 100-year event in CA/NV (Supplementary Fig. 5a) is 1°C warmer for La Nina years than El Nino years. As with the PDO, Tmax_hi shows no difference between ENSO states (Supplementary Fig. 5b).

In order to further test whether anthropogenic forcing might have changed the risk of a warm Tmin_lo occurring, simulations from 2 La Niña years are chosen from both the 2000s and 1960s and compared in return period plots for both domains (Supplementary Fig. 5). A clear signal emerges for the CA/NV domain, which has warmer (>1°C) Tmin_lo values in the 2000s for return periods higher than 30 years. This implies a significant shortening of the diurnal range, which can be dangerous during major heat waves. It also supports the attribution of changes in Tmin_lo during the 2000s to increasing greenhouse gas concentrations.
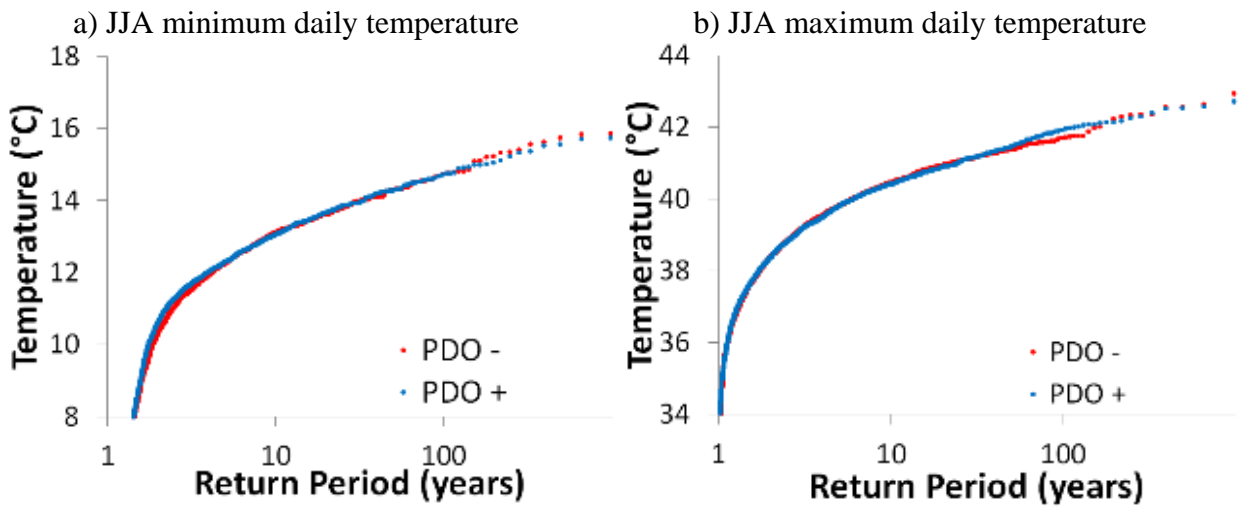
One additional test of internal variability is to compare sigma for all groups of simulations (Supplementary Table 2). Most of the simulations have fairly close standard deviations, suggesting generally the same variability for all model runs. However, in the case of Tmax_hi, the groups of simulations point out interesting differences. For instance, AF1 2000s (2.37°C) is higher than AF1 1960s (2.28°C), suggesting marginally higher variability for the warmer decade of the 2000s. The opposite is found when comparing Natural 2000s (2.47°C) versus AF2 2000s (2.26°C), where the warmer AF2 has lower variability than the Natural. This is a much larger difference and would suggest more extremes in the Natural simulations. Also worth noting is that the observations had a lower standard deviation at 2.11°C, suggesting generally more variability in model output.

For Tmin_lo there is almost no difference between AF1 1960s (2.16°C) and AF1 2000s (2.12°C). However, the standard deviation is twice as large as the observations (1.08°C). Tmin_hi, however, is also fairly close between AF2 2000s (2.32°C) and Natural 2000s (2.44°C), but Natural has slightly higher numbers that echo the results with Tmax_hi. It hints at SST patterns possibly playing a role. The observations are also significantly higher (3.01°C), where extremes were much more prevalent during the 2000s, arguably due to heat waves in 2003 and 2006. This latter point

is noteworthy because it shows that, for minimum temperature variables, the *mean* of the variable is experiencing a greater change than the tails in the distribution.
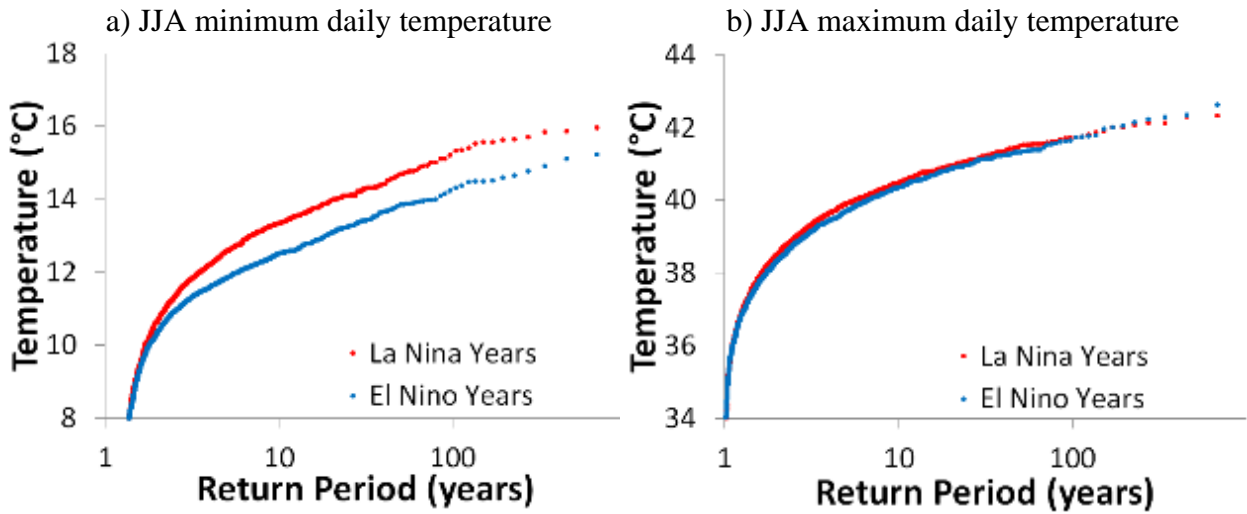
Supplementary Table 2. Sigma for temperature indices in the Central Valley for Tmax_hi, Tmin_lo and Tmin_hi for AF1 1960s, AF1 2000s, AF2 2000s, Natural 2000s, and observations.

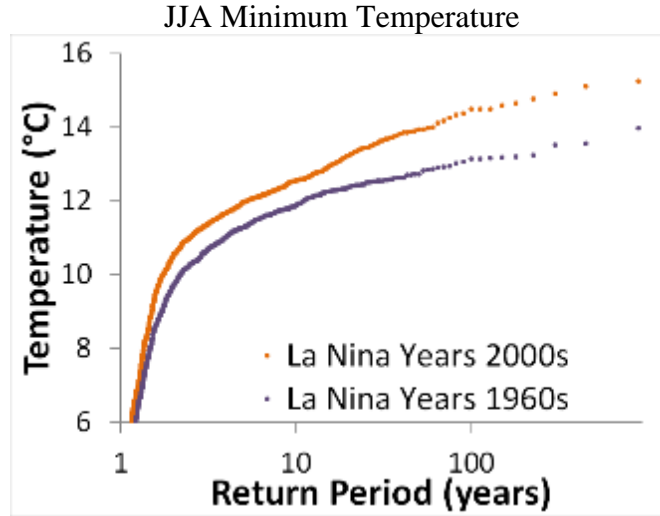|	Tmax_hi	Tmin_lo	Tmin_hi
AF1 1960s	2.28	2.16	
AF1 2000s	2.37	2.12	
AF2 2000s	2.26		2.32
Nat 2000s	2.47		2.44
Observations	2.11	1.08	3.01

Supplementary Fig. 4. Return periods for area averages over California/Nevada for (a) Tmin_lo (°C) and (b) Tmax_hi (°C) during years in the 2000s with positive PDO in blue circles and negative PDO red circles.

JJA minimum daily temperature

a) JJA minimum daily temperature   b) JJA maximum daily temperature

Supplementary Fig. 5. Return periods for area averages over California/Nevada for (a) Tmin_lo (°C) and (b) Tmax_hi (°C) during years in the 2000s for ENSO phases where El Niño years are in blue circles and La Niña years are in red.
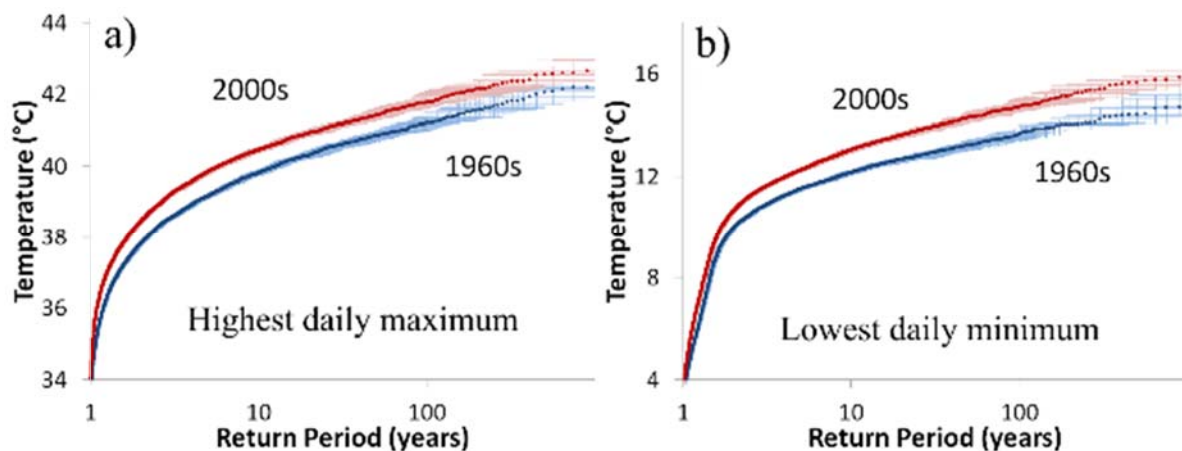
Supplementary Fig. 6. Return periods for Tmin_lo (°C) during La Niña years in the 1960s (purple circles) and the 2000s (orange circles) for California/Nevada.

Trends in temperature for the California/Nevada domain

The trends in the meteorological station observations used for Fig. 1 of the paper merits analysis in order to understand whether the model captures the markedly warmer trends in Tmin_lo compared to Tmax_hi in the larger CA/NV domain. This does not constraint the analysis to only the Central Valley of California. Supplementary Fig. 7a shows return periods for Tmax_hi for area averages over the CA/NV domain for the 2000s and 1960s. Here, Tmax_hi is 0.61°C warmer in the 2000s for 10-100 year return periods.

There is a wider margin in the distribution for Tmin_lo between the two decades analyzed compared to Tmax1 (Supplementary Fig. 7b). For CA/NV the 2000s are, on average, 0.95°C warmer than the 1960s for return periods ranging from 10 to 100 years. Further, a 100-year event in the 1960s is 5 times more likely to occur in the 2000s. This has serious implications for the health sector if cooling centers are to remain open in the evening, as the highest temperatures tend to occur in the late afternoon. Further, the fact that the analysis carried out here includes areas on the coast and mountains, where heat waves happen less often, implies the need for better adaptation as these events grow in frequency and intensity (Giurgis et al. 2014).

Supplementary Fig. 7. Return periods of 2000s versus 1960s comparisons for California/Nevada for (a) Tmax_hi (°C) and (b) Tmin_lo (°C) using AF1 simulations. Hatchings show bootstrapped inner 95% percentile uncertainty ranges.

References

Bumbaco, Karin A, Dello KD, Bond NA (2013) History of Pacific Northwest Heat Waves: Synoptic Pattern and Trends. J Appl Meteor Climatol 52:1618–1631. doi: http://dx.doi.org/10.1175/JAMC-D-12-094.1

Fischer E, Knutti R (2012) Robust projections of combined humidity and temperature extremes. Nature Clim Change 3:126–130

Guirguis K, Gershunov A, Tardy A, and Basu R (2014) The Impact of Recent Heat Waves on Human Health in California.J. Appl. Meteor. Climatol. 53:3–19.
doi: http://dx.doi.org/10.1175/JAMC-D-13-0130.1

Kalnay E., Coauthors (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bull Amer Meteor Soc 77:437–471

Mesinger F, Coauthors (2006) North American Regional Reanalysis. Bull Amer Meteor Soc 87:343360.