Present day greenhouse gases could cause more frequent and longer Dust Bowl heatwaves

Citation for published version:
Cowan, T, Undorf, S, Hegerl, G, Harrington, LJ & Otto, FEL 2020, 'Present day greenhouse gases could cause more frequent and longer Dust Bowl heatwaves', *Nature Climate Change*.
https://doi.org/10.1038/s41558-020-0771-7

Digital Object Identifier (DOI):
10.1038/s41558-020-0771-7

Published In:
Nature Climate Change

Present-day greenhouse gases could cause more frequent and longer Dust Bowl heatwaves

Tim Cowan[1,2,3], Sabine Undorf[3,4], Gabriele C. Hegerl[3], Luke J. Harrington[5], and Friederike E. L. Otto[5]

[1]University of Southern Queensland, Toowoomba, Queensland, Australia

[2]Bureau of Meteorology, Melbourne, Victoria, Australia

[3]School of GeoSciences, The Kings Building, University of Edinburgh, Edinburgh, United Kingdom.

[4]Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

[5]Environmental Change Institute, University of Oxford, Oxford, United Kingdom.

Corresponding author: Tim Cowan (tim.cowan@bom.gov.au)

Substantial warming occurred across North America, Europe and the Arctic over the early twentieth century[1], including an increase in global drought[2], and was partially forced by rising greenhouse gases[3]. The period included the 1930s Dust Bowl drought[4–7] across North America's Great Plains that caused widespread crop failures[4,8], large dust storms[9] and considerable out-migration[10]. This coincided with the central United States experiencing its hottest summers of the twentieth century[11,12] in 1934 and 1936, with over 40 heatwave days and maximum temperatures surpassing 44°C at some locations[13,14]. Here we use a large-

ensemble regional modelling framework to show that greenhouse gas increases slightly enhanced heatwave activity over the eastern US during 1934 and 1936. Instead of asking how a present-day event would behave in a world without climate warming, we ask how these 1930s heatwaves would behave with present-day greenhouse gases. Heatwave activity in similarly rare events would be much larger under today's atmospheric greenhouse gas forcing, and the return period of a 1-in-100-year heatwave summer (as observed in 1936) would be reduced to about 1-in 40 years. A key driver of the increasing heatwave activity and intensity is reduced evaporative cooling and increased sensible heating during dry springs and summers.

The hottest continental US summer (June-August) on record was 1936, with 1934 the fourth hottest[15], up to and including 2019. During the record-breaking summer of 1936, Kansas and Oklahoma experienced more than a month of heatwave days, with individual events exceeding two weeks and maximum temperatures above 44°C (Fig. 1). The extreme heat and drought were compounded by the widespread removal of the native prairie vegetation in the 1920s[16], and with the Great Depression[4], led to substantial out-migration from the central plains[10]. Observational and modelling evidence suggests that warm North Atlantic and cool tropical Pacific sea surface temperature anomalies (SSTAs) forced a distinctive upper-level ridge over the continental US[9,14], and a weakening of moisture advection from the Gulf of Mexico[6,17] that contributed to the Dust Bowl conditions. These extremes further occurred during a period of multidecadal warming[1], with early twentieth century global-scale drought likely amplified by greenhouse gases (GHGs)[2].

With evidence suggesting a human-induced influence on global heat extremes emerged in the 1930s[18] we investigate whether GHG levels contributed to the Dust Bowl heatwaves. Unlike many event attribution studies setting out to determine what a present-day event would be like in a counterfactual world without present-day GHGs[19], we ask how the 1930s heatwaves would manifest in the present day, using event attribution methods. We use the weather@home2 (WAH2) attribution framework to evaluate how the probability of the Dust Bowl heatwaves may have changed under increased GHGs. We further estimate how changes in GHGs since the 1930s would impact the heatwaves, with WAH2 simulations that are forced with 1930s SSTs, but include present day GHGs. We derive probability estimates of extreme events using an ensemble of over 1200 regional model experiments[20]. We investigate the 1934 and 1936 Dust Bowl heatwaves, defined as events consisting of consecutive anomalously hot days and warm nights relative to a reference climatology (at least three days and two nights exceeding the 90th percentile of daily maximum and minimum temperatures; see Methods).

Long-lasting heatwave conditions developed over the central US during the Dust Bowl summers. In 1934, the frequency of heatwave days (HWF) exceeded 50 days per summer over a large region spanning Texas, Oklahoma and Kansas, with the most protracted heatwaves surpassing 18 days and maximum temperatures exceeding 42°C (Extended Data Fig. 1). The summer of 1936 saw hotter and longer heatwaves (although fewer heatwave days) in the northern Great Plains[14], with days exceeding 44°C across parts of Oklahoma, Kansas and north into the Dakotas (Fig. 1; record-breaking years are outlined in black). Almost 25% of all continental maximum temperature records at 755 observing stations were set[12] in 1936.

We investigate the most extreme heatwave summers over the central US as simulated in the WAH2 ensemble suite. We select the top 200 experiments ranked by HWF; these experiments better represent the large-scale mid-tropospheric circulation associated with the heatwaves (Supplementary Fig. 1). The ensembles that best capture the large-scale mid-tropospheric circulation during the hottest heatwave weeks (Supplementary Figs. 2 and 3; based on 500 hPa geopotential height) generally simulate more frequent and longer events than ensembles with a poor representation of the reanalysis circulation (Supplementary Fig. 4, see Methods for analogue description). The spatial representation of HWF is captured reasonably well by the average of the top 200 ranked $WAH2_{1930s}$ simulations, however with values of ~25-30 days (Fig. 2a,d), the ensemble underestimates the observed frequency. Using every member of the $WAH2_{1930s}$ ensemble, instead of the top 200, gives average HWF values of around 11 days, with the longest heatwaves close to one week and the hottest events surpassing 40°C (Supplementary Fig. 5). These underestimates likely arise because the ensemble average includes experiments with weaker and slightly eastward displaced mid-tropospheric ridging (Supplementary Fig. 3) and wet biases, which produce cooler summers with fewer heatwaves.

To test if increased GHG levels amplified the Dust Bowl heatwaves we compare the $WAH2_{1930s}$ top 200 simulation ensemble to another set with the human response removed from SSTs and using pre-industrial GHGs and aerosols ($WAH2_{NAT}$; see Methods for the SST removal process). The anthropogenic GHG forcing ($WAH2_{1930s}$ – $WAH2_{NAT}$) leads to a small increase in HWF of around two extra days over southeast US in 1934 and across the broader eastern US (and a small area of the northern Great Plains) in 1936 (Fig. 2b,e). Over the same regions, the longest heatwaves increase by almost one day while the hottest heatwave days warm by 0.2-0.5°C. A small percentage of the central US shows a significant HWF change due to GHG forcing (3% and 13% in 1934 and 1936, respectively; see Methods for a discussion of its statistical significance) if only considering simulations with strongest HWF. However, average summer HWF in the 1930s shows a clear and significant GHG-induced increase, with more of the central US featuring a significant response as the ensemble size increases (Supplementary Fig. 6). This is more apparent in 1936, with an increase of between 1-2 heatwave days for WAH2 ensembles > 500, whereas smaller ensembles appear strongly influenced by large intra-member variability; this suggests a detectable GHG contribution to HWF by the mid-1930s.

In order to estimate what effect changes in atmospheric composition since the 1930s would have on the Dust Bowl heatwaves, we analyse simulations with present day GHG and aerosol conditions, yet identical SSTs to the $WAH2_{1930s}$ simulations (named $WAH2_{PD}$). This shows that the anthropogenic changes in atmospheric composition compared to 1934 and 1936 alone would have resulted in almost five extra heatwave days at present across the central US for 1934 conditions, increasing to eight extra days for 1936 (Fig. 2c,f; average of the top 200 ranked experiments by HWF). The amplification of heatwave conditions in the $WAH2_{PD}$ ensemble is robust both to ensemble size (Supplementary Fig. 7) and to ranking experiments by their resemblance to the reanalysis mid-tropospheric circulation (not shown); this amplification is driven predominantly by GHGs, and likely moderated by sulfate aerosols[21].

To understand the potential driving mechanism behind the present day amplification of the Dust Bowl heatwave conditions, we consider the influence of spring drought in amplifying heat extremes over the central US[14,17]. We first re-order the simulations based on their spring-time (March-May) precipitation over the central US, driest to wettest, and then see how this affects the subsequent summer precipitation and heatwave behavior (frequency, amplitude and timing) over the central US (Fig. 3). All ensembles show a clear association between spring precipitation and summer conditions (Fig. 3a-d, f-i), with summer deficits tending to follow dry springs, in association with more heatwave days, hotter peak days and earlier events. The differences between $WAH2_{1930s}$ and WAH_{NAT} over the central US for the amplitude and timing are, however, marginal, while a clearer difference in HWF is seen for 1936. Dry springs substantially enhance heatwave conditions in the $WAH2_{PD}$ simulations, on average by around two extra heatwave days in 1934 and ~3-4 days in 1936 (Fig. 3). Across the entire ensemble, the $WAH2_{PD}$ heatwaves are between 0.4 and 0.6°C hotter (Fig. 3c,h) and the first events occur 2-3 days earlier than in the $WAH2_{1930s}$ ensemble (Fig. 3d,i). Yet even the driest 200 simulations (in any WAH2 simulation type) cannot replicate the observed HWF (24-28 days), only explaining between 50 and 66% of the total. An inability to fully replicate the Dust Bowl conditions has been a common feature in SST-forced models[7], and is likely in part due to the under-representation of land use changes[23].

The partitioning of surface heat fluxes, which connect the soil moisture to the atmosphere, drive the hotter and more extreme heatwave conditions in $WAH2_{PD}$ (Supplementary Fig. 8). Drier springs and summers drive a reduction in latent heat fluxes (reduced evaporative cooling) and increased sensible heating leading to lower evaporative fractions (Fig. 3e,j) in $WAH2_{PD}$ relative to $WAH2_{1930s}$ (difference of ~3-5%, but up to 25% over specific central US locations in the hottest months). This amplifies the heatwave conditions in the $WAH2_{PD}$ ensemble. How the land surface determines the partitioning of surface heat fluxes is dependent on precipitation[24], so a wet spring bias over central US[20] could influence the summer conditions in the WAH2 model. Potentially offsetting the wet spring biases is the overestimated spring evaporative fraction in WAH2, which could drive excessive soil moisture depletion, as seen for Europe[20].

The $WAH2_{PD}$ ensemble does not account for a SST warming since the 1930s, which could further amplify heatwave conditions. An ensemble of simulations that account for a SST warming with 2015 SSTs (Supplementary Fig. 9) and present-day GHG levels ($WAH2_{2015}$) produces summers exceeding 40 heatwave days (top 200 HWF summers; Extended Data Fig. 2c), more akin to what was observed in the 1930s. While 2015 values of the Pacific decadal variations match well to 1930s values, 2015 was an El Niño year, which did not occur in 1934 or 1936, but developed in 1931. El Niño is typically associated with cooler and wetter conditions over the southern US and Midwest[25], and reduced heat extremes across continental US[26]. When the same 1921-1948 reference period is used for the $WAH2_{2015}$ ensemble, the resultant impact of higher air temperatures and warmer SSTAs produce more heatwave days (approx. 5-15 days) than in 1931 (Extended Data Fig. 2; compared to 5-10 days for $WAH2_{PD}$ for 1931). This suggests the heatwaves are linked to higher air temperatures, against the backdrop of warmer mean-state SSTs. For 2015, the higher GHG levels and warmer SSTAs, particularly the warming along the coastal US, likely invigorated the turbulent heat fluxes, triggering more summer heatwaves days. An overestimation of sensible heat fluxes in the land surface model[20], a common problem for climate models[28], may have contributed some to this extra warming.

We finally quantify the impact of present-day GHGs on the likelihood of Dust Bowl-type heatwaves by calculating return periods (RP) of maximum HWF over the central US for each experiment (Fig. 4, Extended Data Fig. 3), with uncertainty estimates defined as the one-standard deviation from a randomly selected sub-sample of 1000 simulations, bootstrapped 2000 times. For 1934, an approximate 1 in 100 (93 to 122)-year event in $WAH2_{1930s}$ becomes a 1 in 39 (37 to 41)-year event in $WAH2_{PD}$, and thus shows a more than twofold increase in likelihood due to changes in atmospheric composition of GHGs since 1934 (risk ratio: $RP(WAH2_{1930s}) / RP(WAH2_{PD}) = 2.56$ (2.51 to 2.98)). With $WAH2_{2015}$ the RP reduces further to a 1 in 12 (11.4 to 12.5)-year event, although we note that this result may also be influenced by differences between the SST patterns and resulting atmospheric response, not just overall warmer ocean temperatures. A clear increase in likelihood in WAH2$_{2015}$ (RP: 1 in 32 (30 to 37) years) compared to WAH2$_{1930s}$ and WAH2$_{PD}$ (RPs: ~1 in 250 years (209 to 263)), is found using the heatwave metrics from observations, however observed events are exceptionally rare and their risk changes are thus more affected by uncertainty. The RP for the summer of 1936 is reduced by a similar factor from a 1 in 100 (83 to 105)-year event in WAH2$_{1930s}$ to a 1 in 29 (28 to 30)-year event in WAH2$_{PD}$ (risk ratio = 3.45 (2.96 to 3.5)).

The probability of summers with a HWF similar to the hottest Dust Bowl summers was explored under present-day conditions, both in terms of atmospheric composition changes and in combination with SST warming. One caveat worth noting is that differences in SST anomalies between the mid-1930s and 2015 likely account for part of the varying heatwave responses simulated by WAH2$_{PD}$ and WAH2$_{2015}$. Other caveats include irrigation and dynamic vegetation, important components not featured in the WAH2 model. With an observed cooling of summer temperatures across the central US during the twentieth century attributed to intensive cropping and irrigation[29,30,31], the lack of irrigation in the WAH2 model hampers its ability to capture the likely dampening effect on present-day heat extremes leading to overamplification[28]. Similarly, without dynamic vegetation, the model only has fixed historical bare soil fractions across the central US, making it difficult to assess land-surface feedbacks in the response to rapid land clearing. Modeling studies have shown that the Dust Bowl conditions are amplified by rapidly increasing levels of bare soil and imposed dust[16,32], via surface energy fluxes accelerating the drought[17]; the human-induced contribution to the heatwaves is therefore likely to be under-estimated here. That is reason why the focus of the present study is the direct impact of greenhouse gases on the historical heatwaves under comparable conditions.

The 1930s Dust Bowl heatwaves had devastating impacts[9,10,17] that led to widespread changes to how the US Great Plains was to be managed[4]. This study has shown that as early as the mid-1930s, GHGs likely increased the frequency of summer heatwave days relative to a pre-industrial climate, and demonstrated how the risk of similar events in the present has further increased more than twofold since then. This has wide implications for land management across the central US, given warmer temperature overall could lead to large crop losses on par with the Dust Bowl[8]. This effect may be mitigated at present by irrigation, but if groundwater depletion in the southern central US[33] occurs in the future, heatwaves may amplify strongly. With summer heat extremes expected to intensify over the US throughout the twenty-first century[34], it is likely that the 1930s records will be broken in the near-future even if there is action to mitigate emissions.

References

1. Hegerl, G. C., Brönnimann, S., Schurer, A. & Cowan, T. The early 20th century warming: Anomalies, causes, and consequences. *Wiley Interdiscip. Rev. Clim. Chang.* e522 (2018). doi:10.1002/wcc.522

2. Marvel, K. *et al.* Twentieth-century hydroclimate changes consistent with human influence. *Nature* **569**, 59–65 (2019).

3. Hegerl, G. *et al.* Causes of climate change over the historical record. *Environ. Res. Lett.* **14**, (2019).

4. Worster, D. *Dust Bowl: The Southern Plains in the 1930s*. (Oxford University Press, 1979).

5. Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D. & Bacmeister, J. T. Causes of

long-term drought in the U.S. Great Plains. *J. Clim.* **17**, 485–503 (2004).

6. Brönnimann, S. *et al.* Exceptional atmospheric circulation during the 'Dust Bowl'. *Geophys. Res. Lett.* **36**, L08802 (2009).

7. Hoerling, M., Quan, X. & Eischeid, J. Distinct causes for two principal U. S. droughts of the 20th century. *Geophys. Res. Lett.* **36**, (2009).

8. Glotter, M. & Elliott, J. Simulating US agriculture in a modern Dust Bowl drought. *Nat. Plants* **16193**, 1–6 (2016).

9. Cook, B. I., Seager, R. & Smerdon, J. E. The worst North American drought year of the last millennium : 1934. *Geophys. Res. Lett.* **41**, 7298–7305 (2014).

10. Gutmann, M. P. *et al.* Migration in the 1930s : Beyond the Dust Bowl. *Soc. Sci. Hist.* 707–740 (2016). doi:10.1017/ssh.2016.28

11. DeGaetano, A. T. & Allen, R. J. Trends in Twentieth-Century Temperature Extremes across the United States. *J. Clim.* **15**, 3188–3205 (2002).

12. Abatzoglou, J. T. & Barbero, R. Observed and projected changes in absolute temperature records across the contiguous United States. *Geophys. Res. Lett.* **41**, 6501–6508 (2014).

13. Kohler, J. P. WEATHER OF 1936 IN THE UNITED STATES. *Mon. Weather Rev.* **65**, 12–16 (1937).

14. Cowan, T. *et al.* Factors Contributing to Record-Breaking Heat Waves over the Great Plains during the 1930s Dust Bowl. *J. Clim.* **30**, 2437–2461 (2017).

15. NOAA. *State of the Climate: National Climate Report for August 2018*. (2018).

16. Cook, B. I., Miller, R. L. & Seager, R. Amplification of the North American 'Dust Bowl'

drought through human-induced land degradation. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4997–5001 (2009).

17. Donat, M. G. *et al.* Extraordinary heat during the 1930s US Dust Bowl and associated large-scale conditions. *Clim. Dyn.* **46**, 413–426 (2016).

18. King, A. D. *et al.* Emergence of heat extremes attributable to anthropogenic influences. *Geophys. Res. Lett.* **43**, 3438–3443 (2016).

19. Otto, F. E. L. Attribution of Weather and Climate Events. *Annu. Rev. Environ. Resour.* **42**, 627–646 (2017).

20. Guillod, B. P. *et al.* weather@home 2 : validation of an improved global – regional climate modelling system. *Geosci. Model Dev.* **10**, 1849–1872 (2017).

21. Undorf, S., Bollasina, M. A. & Hegerl, G. C. Impacts of the 1900-74 increase in anthropogenic aerosol emissions from North America and Europe on Eurasian summer climate. *J. Clim.* **31**, 8381–8399 (2018).

22. Mueller, B. & Seneviratne, S. I. Hot days induced by precipitation deficits at the global scale. *Proc. Natl. Acad. Sci.* **109**, 12398–12403 (2012).

23. Cook, B. I., Seager, R. & Miller, R. L. Atmospheric circulation anomalies during two persistent north american droughts: 1932-1939 and 1948-1957. *Clim. Dyn.* **36**, 2339–2355 (2011).

24. Donat, M. G., Pitman, A. J. & Seneviratne, S. I. Regional warming of hot extremes accelerated by surface energy fluxes. *Geophys. Res. Lett.* **44**, 7011–7019 (2018).

25. Hu, Z. & Huang, B. Interferential Impact of ENSO and PDO on Dry and Wet Conditions in the U.S. Great Plains. *J. Clim.* **22**, 6047–6065 (2009).

26. Kenyon, J. & Hegerl, G. Influence of Modes of Climate Variability on Global Temperature Extremes. *J. Clim.* **21**, 3872–3889 (2008).

27. Jia, L. *et al.* The roles of radiative forcing, sea surface temperatures, and atmospheric and land initial conditions in U.S. summer warming episodes. *J. Clim.* **29**, 4121–4135 (2016).

28. Ukkola, A. M., Pitman, A. J., Donat, M. G., De Kauwe, M. G. & Angélil, O. Evaluating the Contribution of Land-Atmosphere Coupling to Heat Extremes in CMIP5 Models. *Geophys. Res. Lett.* **45**, 9003–9012 (2018).

29. Mueller, N. D. *et al.* Cooling of US Midwest summer temperature extremes from cropland intensification. *Nat. Clim. Chang.* **6**, 317–322 (2016).

30. Thiery, W. *et al.* Present-day irrigation mitigates heat extremes. *J. Geophys. Res.* **122**, 1403–1422 (2017).

31. Alter, R. E., Douglas, H. C., Winter, J. M. & Eltahir, E. A. B. Twentieth Century Regional Climate Change During the Summer in the Central United States Attributed to Agricultural Intensification. *Geophys. Res. Lett.* **45**, 1586–1594 (2018).

32. Cook, B. I., Miller, R. L. & Seager, R. Dust and sea surface temperature forcing of the 1930s '"Dust Bowl"' drought. *Geophys. Res. Lett.* **35**, L08710 (2008).

33. Scanlon, B. R. *et al.* Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proc. Natl. Acad. Sci.* **109**, 9320–9325 (2012).

34. Diffenbaugh, N. S. & Ashfaq, M. Intensification of hot extremes in the United States. *Geophys. Res. Lett.* **37**, 1–5 (2010).

Methods

Heatwave definition We investigate the heatwaves that emerged during the summers (June-August, JJA) of 1934 and 1936, as these were the two most intense and active heatwave summers across the central US (defined as 105°-85°W, 30°-44°N) in the 1930s. A heatwave is defined to occur when the daily maximum and minimum temperatures exceed their daily 90th percentile for at least three consecutive days and two nights, respectively[14]. The percentile approach is based on a centered 15-day window that removes all monthly and seasonal variations[35], and we use a climatological base period of 1920-2012 for observations. Percentile based definitions are widely used across the world to define heatwave conditions[36]. We quantify four main heatwave metrics: the total count or *frequency* of heatwave days (HWF), the longest *duration* summer heatwave (HWD), the hottest heatwave day of the hottest heatwave or the *amplitude* (HWA); and the *timing* of the earliest summer heatwave. We predominantly focus on the HWF. The HWF and HWD are considered relative heatwave metrics, as they are referenced against the climatology of observed data and model simulation respectively, and hence account for temperature biases in the model[37]. Given model warm biases are prominent in the summer over Europe and North America[20], the daily modelled Tmax and Tmin were bias corrected against the 90th percentile observed temperatures. This only made a difference for heatwave intensity metrics such as HWA.

To calculate the observed hottest heatwave week for 1934 and 1936 across the central US (domain shown in Fig. 1a), we determine the start date of the hottest heatwave for each grid cell. We then find the percentage of grid cells that share the same date, performing a 7-day running mean to choose the week centered on the start date with the largest percentage of grid cells over the central US. Hence, the hottest observed summer heatwave weeks, based on daily maximum temperature in gridded observations are 16-22 July 1934 and 3-9 July 1936.

Observations Observed heatwaves are calculated using observed station temperatures from the Global Historical Climatology Network-Daily (GHCN-D) archive[38], and the homogenized daily Berkeley Earth Surface Temperature (BEST) dataset[39]. The BEST dataset is a 1° × 1° gridded 'experimental' product that incorporates over 2000 stations (mostly GHCN-D) in the 1930s decade and is created using the same techniques as the monthly dataset; the GHCN-D network, quality control and station selection are described in Cowan et al.[14]. A direct comparison of the 1934 and 1936 heatwaves in GHCN-D and BEST is shown in Figure 1.

Weather@home2 experiments The weather@home version2 (WAH2) uses a distributed network of home computers across the globe to conduct thousands of model simulations, each with slightly perturbed physics to characterize the spread of uncertainties[20]. The WAH2 experiments are run on the Met Office Hadley Centre N96 Atmospheric Model (HadAM3P; 1.25° × 1.875° resolution), forced with observed SSTs from HadISST2.1. The HadAM3P provides boundary conditions to the 25 km resolution Hadley Centre Regional Model (HadRM3P), which is fixed over the United States, south of 45°N, one of the pre-defined WAH2 regions. This region experienced the most intense heat observed during the 1930s[14], although for analysing the atmospheric circulation from HadAM3P, we extend our focus to 60°N. The most extreme summer heatwave years, 1934 and 1936, in terms of HWF, HWD, and HWA[14], were chosen for the WAH2 simulations. For these two years, three sets of atmospheric model simulations driven by observed SSTs were performed over 390 days, from the previous years' December through to the end of December of the year in question, with a small perturbation added to the initial potential temperature field. These simulation types include:

1) 1934/1936 observed SSTs and 1934/1936 prescribed greenhouse gases (GHGs) + aerosols ($WAH2_{1930s}$: ensemble size of 1585 and 1576 experiments for 1934 and 1936, respectively). It should be noted that anthropogenic aerosol emissions do not include those associated with land degradation; in fact, rapid land use change in the 1930s is typically not considered in SST-forced atmospheric model experiments, leading to their general failure to replicate the magnitude of the Dust Bowl drought and heatwaves[7,16].

2) 1934/1936 observed SSTs with human-induced warming removed and pre-industrial GHGs + aerosols ($WAH2_{NAT}$: ensemble size same as $WAH2_{1930s}$). The $WAH2_{NAT}$ simulations are considered counterfactual as they provide an estimate of Dust Bowl heatwave activity across the central US in a world without anthropogenic changes in atmospheric composition. Note that land-cover does not change relative to the 1930s in these simulations, as HadRM3P does not have dynamic vegetation. To obtain the SST pattern of change, 11 Coupled Model Intercomparison Project Phase 5 (CMIP5) models that have three or more ensemble members for their Historical and HistoricalNatural simulations are used[40]. For each CMIP5 model, the human-induced SST signal is taken as the difference between its available Historical and HistoricalNAT simulations, with 10-year running mean estimates of ΔSST determined for each month and averaged over the three or more ensemble members per simulation, centered on the year of interest. For each of the 11 CMIP5 models, this ΔSST is removed from the observed SSTs to obtain 11 estimates of 'naturalized' SSTs, which are used to force $WAH2_{NAT.}$ Thus the $WAH2_{NAT}$ ensemble captures the CMIP5 model uncertainty in the removal of the anthropogenic warming from the observed SSTs at least to some extent.

3) 1934/1936 observed SST and land surface, and present-day (2015) prescribed GHGs + aerosols ($WAH2_{PD}$): ensemble size of 1258 and 1222 experiments for 1934 and 1936, respectively.

The short model spin-up period is sufficient for allowing water to penetrate the four soil layers (0-0.1 m, 0.1-0.4 m, 0.4-1 m, 1-2 m) for the central US, although a longer spin-up would likely reduce the warm summer bias[20]. We also conducted an experiment with present-day (2015) observed SSTs and prescribed GHGs + aerosols ($WAH2_{2015}$; ensemble size of 1276) and a 1931 experiment (both $WAH2_{1930s}$ and $WAH2_{PD}$ versions; sizes 1589 and 1201, respectively). The $WAH2_{2015}$ includes the combined impact of warmer mean-state SSTs and present-day GHGs, in a period where the large-scale SST patterns, particularly the Atlantic and Pacific states are similar to 1934 and 1936, but not identical (we argue that 2015 is the most suitable recent year for that purpose, see Supplementary Fig. 9). The advantage of using SSTs from 2015, instead of adding a generic warming pattern to 1930s SSTs, is that it avoids the uncertainty surrounding the perturbed SST warming pattern. This is at the expense of possibly not fully capturing 1930s atmospheric conditions forced by perturbed 1930s SST patterns. Yet 2015 was an El Niño year, so we can compare it to the only El Niño year in the Dust Bowl decade, 1931, which coincidently was a strong heatwave year[14]. Yet, we are aware that there are limitations to the $WAH2_{PD}$ experiments in that they would not capture the effect on the heatwaves of a long-term ocean warming superimposed on 1930s interannual SSTs. This cannot be fully replicated in $WAH2_{2015}$ given the difference in interannual SST anomalies to the 1930s. A 1921-1948 climatology experiment was also conducted, from which the heatwave percentile thresholds for each individual WAH2 simulation was determined. The residual differences in heatwave patterns between the $WAH2_{1930s}$ simulations, and $WAH2_{NAT}$ and $WAH2_{PD}$ are tested using the non-parametric Mann-Whitney U test[41]. The null hypothesis tested here is that the heatwaves from the two sets of experiments are drawn from the same distribution. The Mann-Whitney U test determines whether the experiment in question is distinguishable from its partner experiment at the 5% confidence level. Accounting for a false discovery rate[42] of 5%, the null hypothesis cannot be reliably rejected for $WAH2_{NAT}$ and $WAH2_{1930s}$ differences over the central US for small ensemble sizes (< 500), whereas using the whole ensemble suite yields widespread significant differences (see Supplementary Fig. 6e,j). Yet, given clusters of significant points show little variation as ensemble size increases (above 200), we are satisfied that the differences between $WAH2_{NAT}$ and $WAH2_{1930s}$ are not statistical artefacts.

Circulation analogues To assess the anthropogenic influence on the simulated heatwaves given the atmospheric circulation from 1934/1936, we choose the most realistic simulations from each of the ensembles making use of the circulation analogs method[43]. This approach selects 7-day periods that display the greatest similarity between an atmospheric circulation in the Twentieth Century Reanalysis V2c[44] (ensemble average of 56-members) and that in the HadAM3P simulations over the North American domain of [140°-60°W, 20°-60°N]. Here we treat the reanalysis ensemble as our best guess "observed" circulation (Donat et al.[17] showed that the spread between the individual members is small after 1910), noting that synoptic pressures are the only land surface observations assimilated in the reanalysis model. We analyse the start of the hottest observed summer heatwave week over the central US for 1934 and 1936. Analogues are found from each individual model simulation for a circulation state that is most close to that of the first day of the hottest summer heatwave and each of the 3 days before and after (7 days in total). From this, the 5 best ranked analogues for each day are averaged, meaning each experiment consists of 35 analogue patterns. We choose 500 hPa geopotential height (Z500) to diagnose similarity of simulated WAH2 circulation to the circulation in the reanalysis (based on minima in Euclidean distance to the reanalysis), as Z500 it is less affected by surface heat low variations than sea level pressure[43]. Choosing a smaller number of analogues (~5) has also been shown to better capture observed conditions[43]. The WAH2$_{2015}$ experiments are less skillful at capturing the reanalysis circulation states from 1934 and 1936 (Supplementary Fig. 2), presumably because the 2015 SSTA pattern is not identical to the 1934 and 1936 SSTA patterns, and hence triggers a different atmospheric response. This ranking by similarity to the reanalysis circulation during the hottest heatwaves is important, as summer heatwave metrics are typically larger for the experiments that exhibit more realistic circulation states, as shown in the WAH2$_{1930s}$ simulations (Supplementary Fig. 4).

Return period analysis To evaluate return periods for our observed heatwave metrics we use the Weibull interval formula ($r/(n+1)$) for estimating probabilities of exceedance in our WAH2 simulations, based on ranking (r) the heatwave metrics - in our case, the maximum HWF over the central US - across the whole ensemble (n). The return period, which is the reciprocal of the exceedance probability, describes the time one would on average have to wait for an event of the same or more extreme magnitude to reoccur. We treat each model simulation per experiment type (e.g., WAH2$_{PD}$ or WAH2$_{1930s}$) as one independent year, hence our return periods are based on 1000+ model (repeated) years. The risk ratio (or increase in likelihood of particular heatwave metric value) can be calculated from the ratio of the return periods for two different experiments (e.g., WAH$_{1930s}$ and WAH$_{PD}$). Uncertainty estimates (error bars) for the return periods (Fig. 4) and risk ratios are determined from 1000 members, sub-sampled from each WAH2 ensemble and bootstrapped 2000 times. We also use two estimates of the observed HWF from BEST (in Fig.

4), calculated over a short (1921-1948) and long (1920-2012) period, to show the effect of climatology selection on the return periods.

Data availability

Source files for Figure 1 (observed heatwave metrics), Figure 3 (WAH2 time series), Figure 4 (return period) and Extended Data Figure 3 can be obtained from: **https://github.com/tcowan80/Cowan_et_al_2020_DustBowl_GHG**. The Berkley Earth Surface Temperature (BEST) gridded product can be downloaded from http://berkeleyearth.org/. The Global Historical Climatology Network-Daily (GHCN-D) archive can be accessed from ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/by_year/. The WAH2 experiments were coordinated through the Environmental Change Institute at the University of Oxford and can be made available on request.

Code availability

The code to generate the main figures and extended data figures is available at: https://github.com/tcowan80/Cowan_et_al_2020_DustBowl_GHG. The code to calculate weather analogs, including installation, is publicly available from https://github.com/sradanov/castf90. Information on its use is available at **https://flyingpigeon.readthedocs.io/en/latest/processes_des.html.** All supplementary figure code is available on request. Spatial plots are produced using NCAR Command Language (NCL; version 6.4.0; doi:10.5065/D6WD3XH5). Return period 2-D plots are generated using Grace 5.1.25 (http://plasma-gate.weizmann.ac.il/Grace/).

References

35. Perkins, S. E. & Alexander, L. V. On the Measurement of Heat Waves. *J. Clim.* **26**, 4500–

4517 (2012).

36. Grotjahn, R. *et al.* North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends. *Clim. Dyn.* **46**, 1151–1184 (2016).

37. Gross, M. H., Alexander, L. V, Macadam, I., Green, D. & Evans, J. P. The representation of health-relevant heatwave characteristics in a Regional Climate Model ensemble for New South Wales and the Australian Capital Territory, Australia. *Int. J. Climatol.* **37**, 1195–1210 (2017).

38. Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An overview of the global historical climatology network-daily database. *J. Atmos. Ocean. Technol.* **29**, 897–910 (2012).

39. Rohde, R. *et al.* A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinformatics Geostatistics An Overv.* **1**, (2013).

40. Haustein, K. *et al.* Real-time extreme weather event attribution with forecast seasonal SSTs. *Environ. Res. Lett.* **11**, (2016).

41. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947).

42. Wilks, D. S. On ‘field significance’ and the false discovery rate. *J. Appl. Meteorol. Climatol.* **45**, 1181–1189 (2006).

43. Jézéquel, A., Yiou, P. & Radanovics, S. Role of circulation in European heatwaves using flow analogues. *Clim. Dyn.* **50**, 1145–1159 (2018).

44. Compo, G. P. *et al.* The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.* **137**, 1–28 (2011).

Acknowledgments

This study forms part of the Transition into the Anthropocene (TITAN) project, funded by a European Research Council (ERC) Advanced Grant (EC-320691), and was further supported by the EUCLEIA project funded by the European Union’s Seventh Framework Programme (FP7/200713) under Grant Agreement 607085 and the EUPHEME ERA4CS grant 690462 and the Sigrist Foundation. G. H. was also supported by the Wolfson Foundation and the Royal

Society as a Royal Society Wolfson Research Merit Award (WM130060) holder. TC was also supported by the Northern Australian Climate Program (NACP), with funding provided by Meat and Livestock Australia, the Queensland Government, and University of Southern Queensland. SU was also supported by the H2020 project EUCP.

Author contributions

T.C. and G.H. designed the study. F. O. and L. H. designed the model experiments. L. H. conducted the model experiments. T.C. performed the analysis and wrote the first draft. S. U. provided analysis for the Pacific and Atlantic decadal variations. All authors helped in the discussions, writing, editing and revising subsequent drafts.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to T. Cowan. (email: tim.cowan@bom.gov.au).

Figure 1: Observed Dust Bowl heatwave conditions in 1936. A comparison between observations from (left) Global Historical Climatology Network-Daily (GHCN-D) stations, and (right) Berkley Earth Surface Temperature (BEST) for summer heatwave conditions averaged over 1936. These include **a,b** heatwave frequency (HWF), **c,d**, heatwave duration (HWD), and **e,f,** heatwave amplitude (HWA). The heatwave metrics are calculated against a 1920-2012 reference period. The outlined GHCN-D stations are those where 1936 was the year with the most heatwave days, and the longest and hottest events of any year on record (up to present). The conditions for 1934 are shown in Extended Data Figure 1.

Figure 2: Simulated Dust Bowl HWF in 1934 and 1936 for strong heatwave summers. weather@home2 (WAH2) simulations with 1930s forcings ($WAH2_{1930s}$) for **a,** 1934 and **d,** 1936. Each ensemble average is based on 200 experiments that simulate the most heatwave days over the central US (boxed region). **b,e,** difference between WAH_{1930s} and simulations with pre-industrial GHGs and SST warming removed ($WAH2_{NAT}$). Significant differences at the 5% level are stippled, based on the non-parametric Mann-Whitney U-test[41] (note ensemble shows overall significant increase, Figure S7). **c,f**, difference between the hottest 200 WAH2 simulations with all forcings and present-day GHG levels ($WAH2_{PD}$) and WAH_{1930s}. All differences in **c,f,** that are *not* grey are significant at the 5% level. The percentage of grid points over the central US that indicate a 5% significant difference is shown in the bottom left corner in **b,c,e,f**.

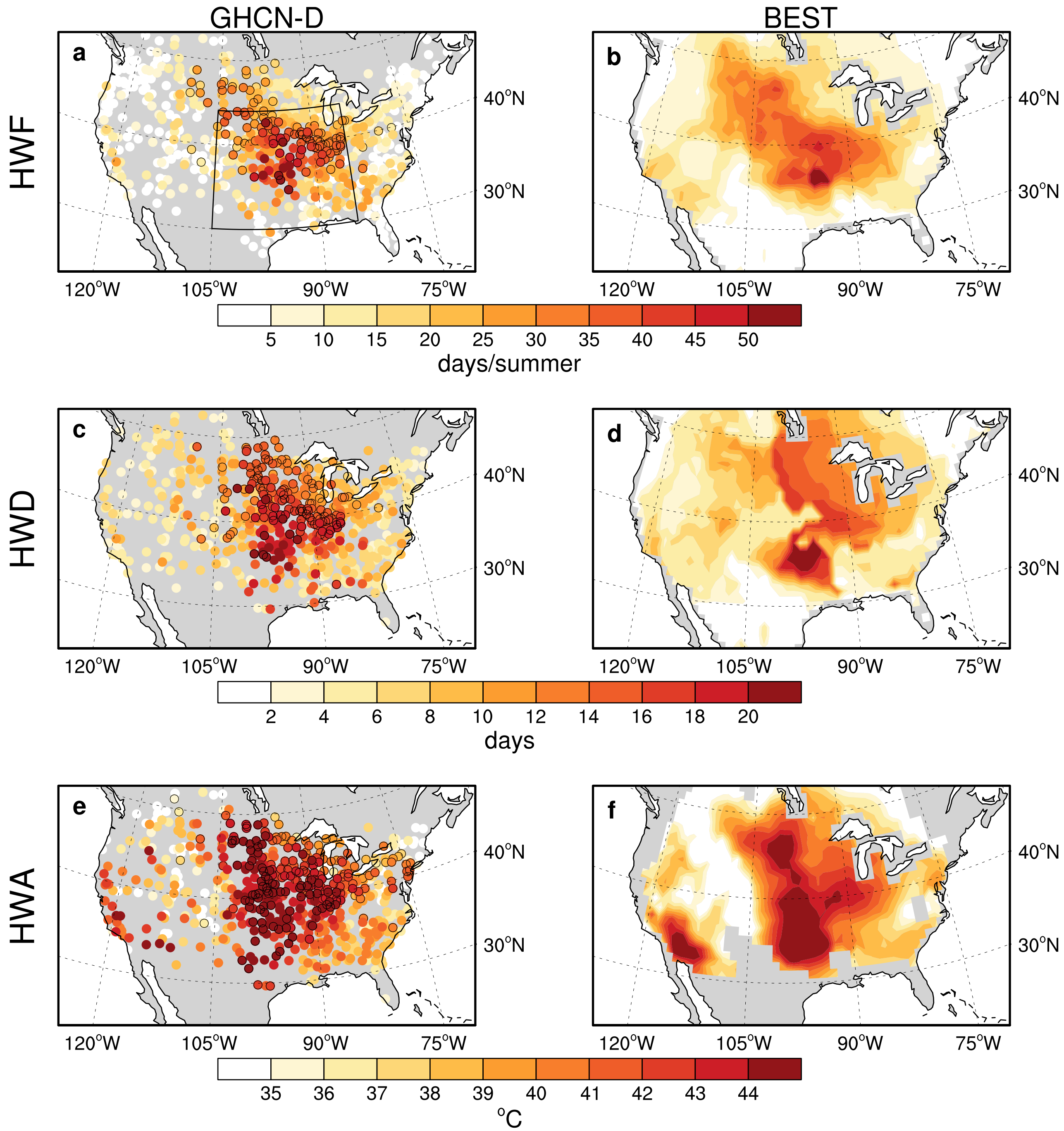
Figure 3: Role of spring precipitation in summer heatwave conditions. Comparison between $WAH2_{NAT}$ (black), $WAH2_{1930s}$ (orange), and $WAH2_{PD}$ (red) of summer **a,f,** precipitation, heatwave **b,g**, frequency, **c,h,** amplitude and **d,i,** timing; and **e,j,** evaporative fraction, with experiments ranked by the preceding spring-time (March, April, May) precipitation over central United States for **a-e,** 1934 and **f-j,** 1936. A 200-member running average is applied to the simulations. The error bars signify the 95% confidence interval based on a t-test of each n = 200 sample.

Figure 4: Return period HWF for central US. Return period of maximum summer HWF over central US (see boxed region in Fig. 1a) for **a**, 1934 and **b,** 1936, for $WAH2_{1930s}$ (orange), $WAH2_{PD}$ (red), and $WAH2_{2015}$ (black). Green horizontal lines indicate the observed estimate range from BEST based on HWF calculated against 1921-1948 (lower line) and 1920-2012 (upper line) climatologies. Error bars reflect the one-standard deviation of a 1000-member sub-sample, which is bootstrapped 2000 times.

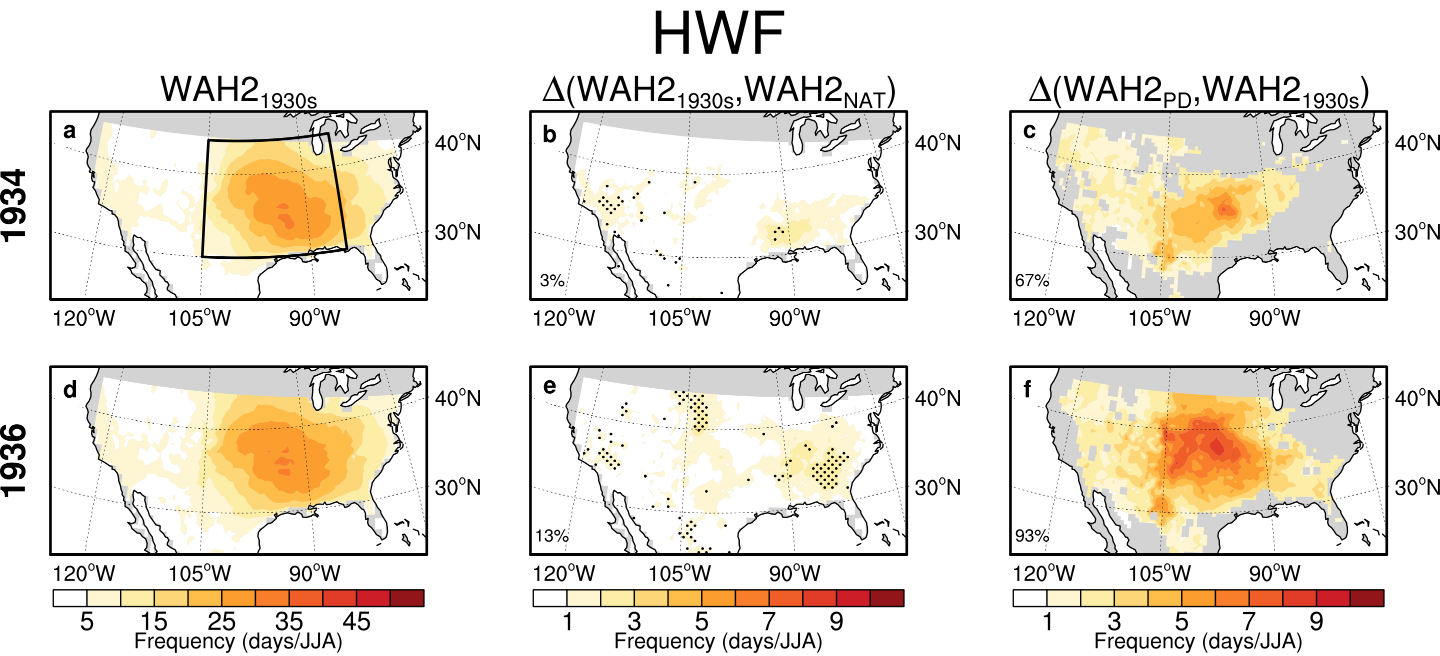
Extended Data Figure 1: Observed Dust Bowl heatwave conditions in 1934. A comparison between observations from (left) Global Historical Climatology Network-Daily (GHCN-D) stations, and (right) Berkley Earth Surface Temperature (BEST) for summer heatwave conditions averaged over 1934. These include **a,b** heatwave frequency (HWF), **c,d**, heatwave duration (HWD), and **e,f,** heatwave amplitude (HWA). The heatwave metrics are calculated against a 1920-2012 reference period. The outlined GHCN-D stations are those where 1934 was the year with the most heatwave days, and the longest and hottest events.

Extended Data Figure 2: Comparison of simulated heatwave frequency between 1931 and 2015. a-c, Average over top 200 ranked experiments that simulate the most summer heatwave days over the central US in 1931 for **a,** WAH2$_{1930s}$, **b,** WAH2$_{PD}$; compared to **c,** WAH2$_{2015}$. **d-f,** Average over the bottom ranked experiments for **d,** WAH2$_{1930s}$, **e,** WAH2$_{PD}$; compared to **f,** WAH2$_{2015.}$

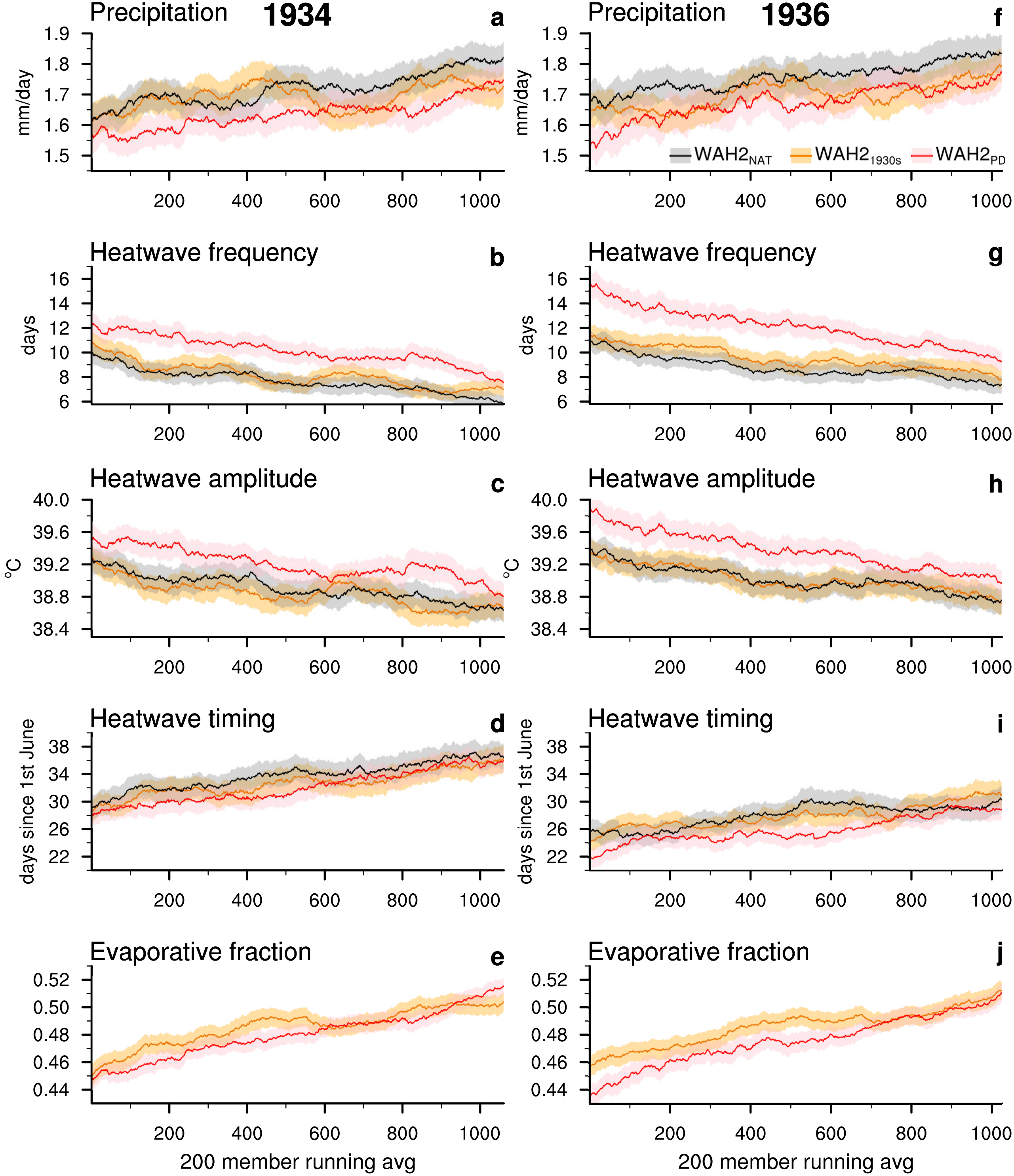
Extended Data Figure 3: Spatial maps of return period of the observed 1934 and 1936 HWF. Return period of summer HWF for (**a-c**) 1934 and (**d-f**) 1936, for **a,d,** WAH2$_{1930s}$, **b,e,** WAH2$_{PD}$, and **c,f,** WAH2$_{2015}$.



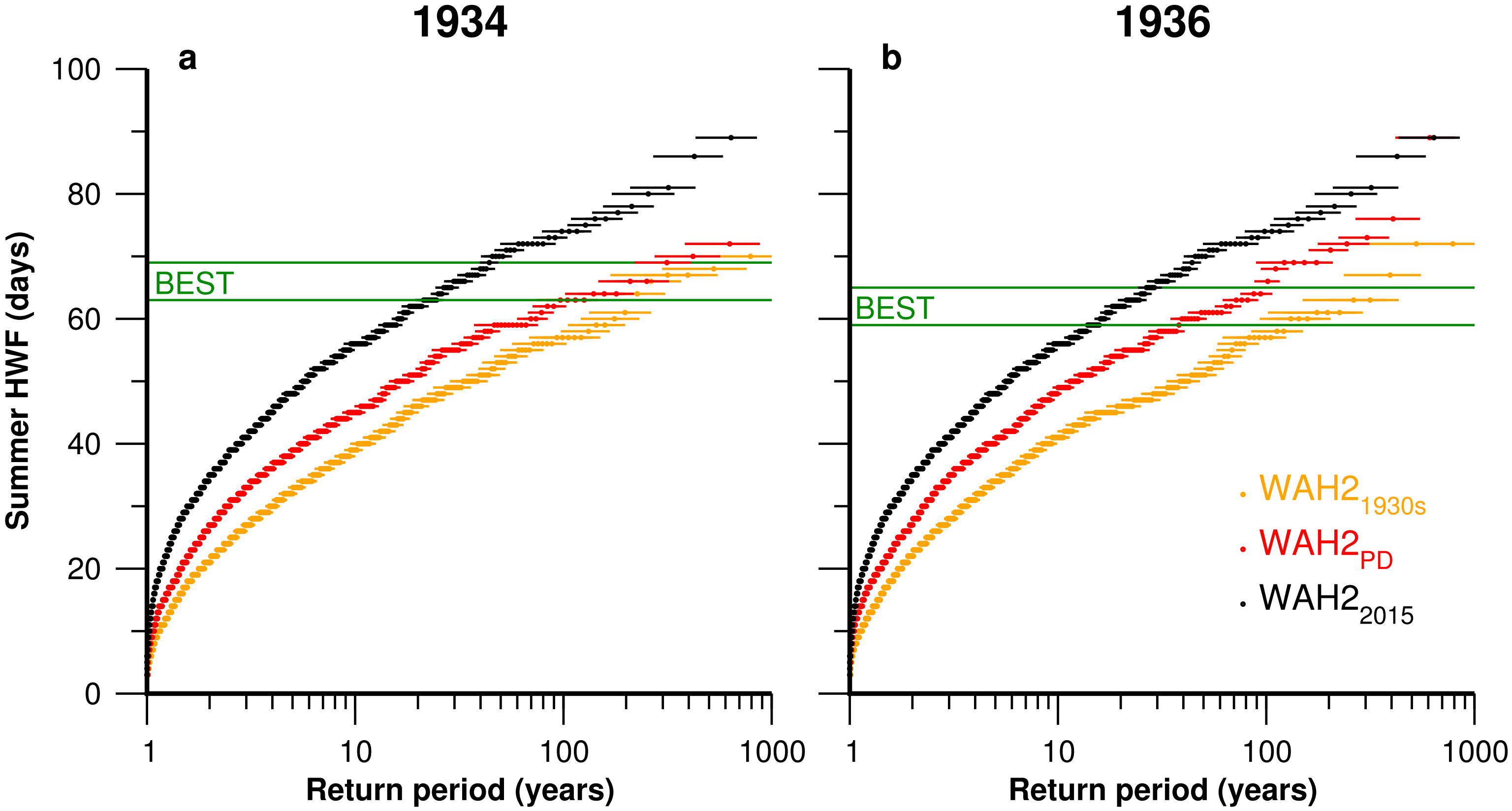
GHCN-D
BEST
HWF
HWD
HWA
a
b
c
d
e
f
40°N
30°N
120°W
105°W
90°W
75°W
5 10 15 20 25 30 35 40 45 50
days/summer
2 4 6 8 10 12 14 16 18 20
days
35 36 37 38 39 40 41 42 43 44
°C



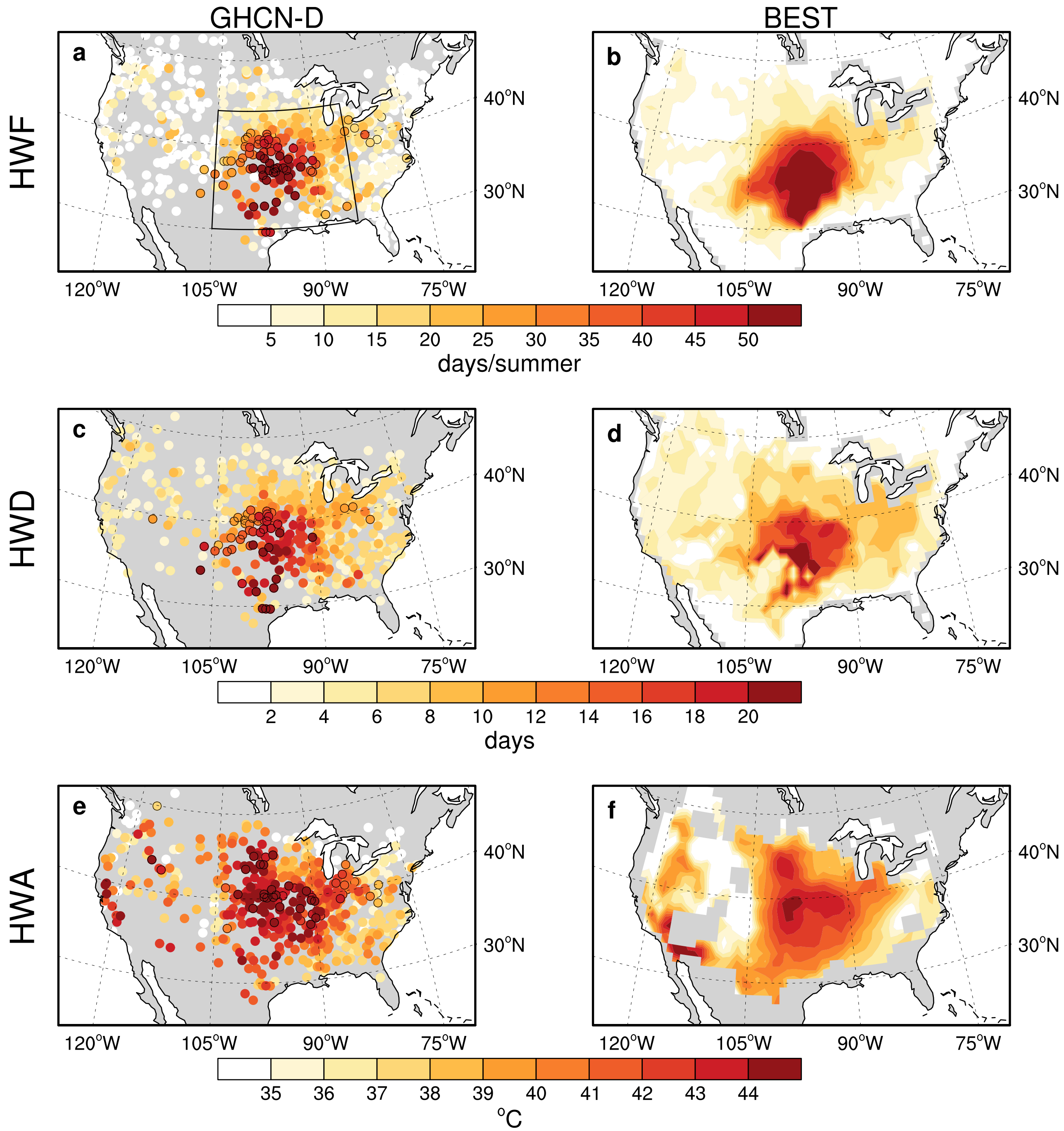
HWF
WAH2 1930s
Δ(WAH2 1930s, WAH2 NAT)
Δ(WAH2 PD, WAH2 1930s)
1934
1936
a
b
c
d
e
f
3%
67%
13%
93%
40°N
30°N
120°W
105°W
90°W
5
15
25
35
45
1
3
5
7
9
Frequency (days/JJA)



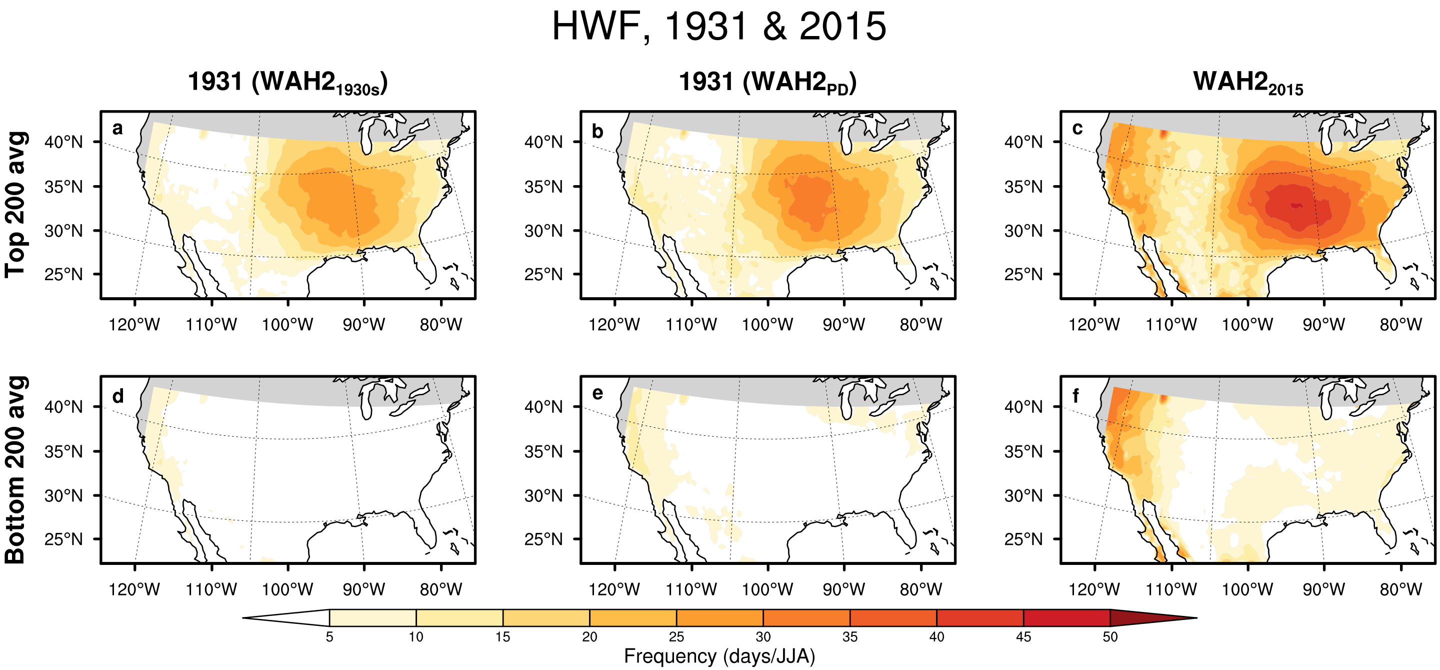
Precipitation
1934
a
mm/day
1.9
1.8
1.7
1.6
1.5
200
400
600
800
1000
Precipitation
1936
f
WAH2$_{NAT}$
WAH2$_{1930s}$
WAH2$_{PD}$
Heatwave frequency
b
days
16
14
12
10
8
6
Heatwave frequency
g
Heatwave amplitude
c
°C
40.0
39.6
39.2
38.8
38.4
Heatwave amplitude
h
Heatwave timing
d
days since 1st June
38
34
30
26
22
Heatwave timing
i
Evaporative fraction
e
0.52
0.50
0.48
0.46
0.44
Evaporative fraction
j
200 member running avg



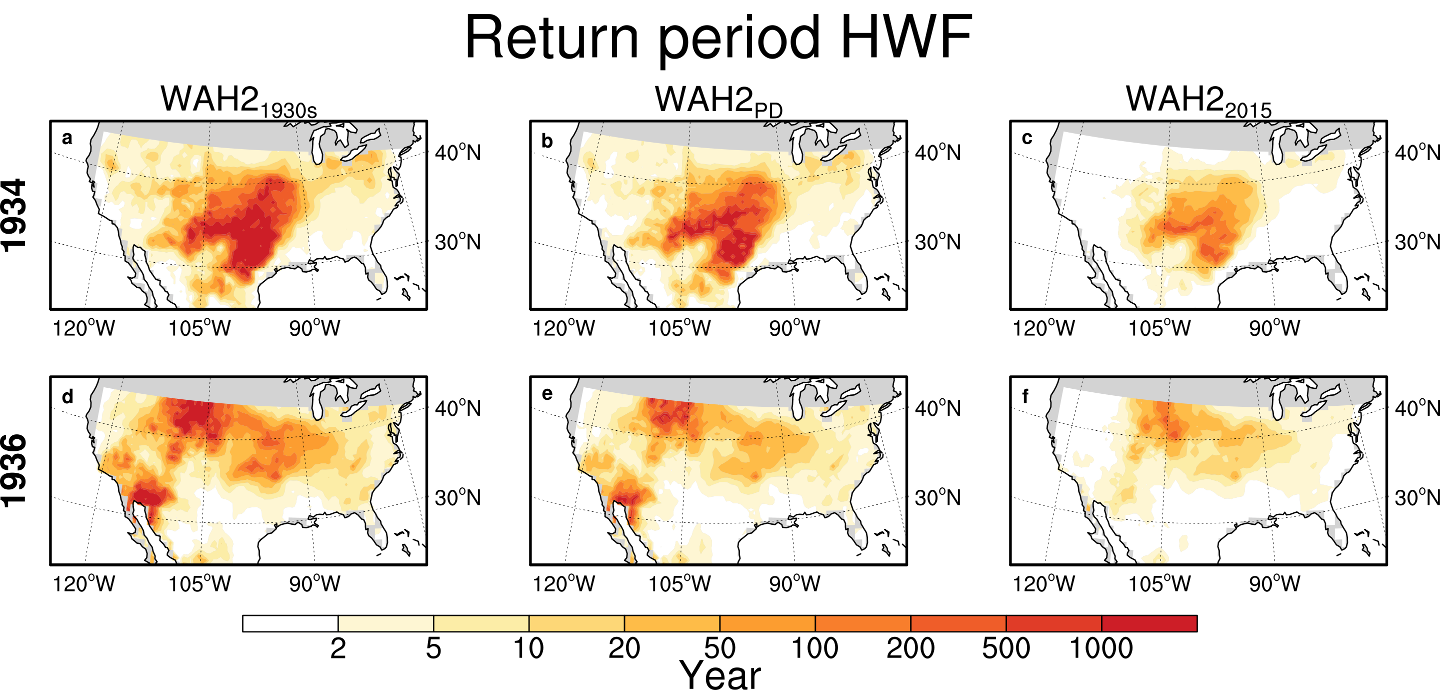
1934
1936
a
b
100
80
60
40
20
0
Summer HWF (days)
BEST
BEST
1
10
100
1000
Return period (years)
Return period (years)
WAH2$_{1930s}$
WAH2$_{PD}$
WAH2$_{2015}$



GHCN-D
BEST
HWF
HWD
HWA
a
b
c
d
e
f
40°N
30°N
120°W
105°W
90°W
75°W
5 10 15 20 25 30 35 40 45 50
days/summer
2 4 6 8 10 12 14 16 18 20
days
35 36 37 38 39 40 41 42 43 44
°C



HWF, 1931 & 2015
1931 (WAH2 1930s)
1931 (WAH2 PD)
WAH2 2015
Top 200 avg
Bottom 200 avg
a
b
c
d
e
f
40°N
35°N
30°N
25°N
120°W
110°W
100°W
90°W
80°W
5
10
15
20
25
30
35
40
45
50
Frequency (days/JJA)


Return period HWF
WAH2 1930s
WAH2 PD
WAH2 2015
1934
1936
a
b
c
d
e
f
40°N
30°N
120°W
105°W
90°W
2
5
10
20
50
100
200
500
1000
Year