Attributing and projecting heatwaves is hard: We can do better

Attributing and projecting heatwaves is hard: we can do better

Geert Jan van Oldenborgh[1], Michael F. Wehner[2], Robert Vautard[3], Friederike E. L. Otto[4], Sonia I. Seneviratne[5], , Peter A. Stott[6], Gabriele C. Hegerl[7], Sjoukje Y. Philip[1], Sarah F. Kew[1]

[1]KNMI, De Bilt, Netherlands
[2]Lawrence Berkeley National Laboratory, Berkeley, Ca, U.S.A.
[3]LSCE/IPSL, laboratoire CEA/CNRS/UVSQ, 91191 Gif sur Yvette Cedex, France
[4]Environmental Change Institute, University of Oxford, Oxford, U.K.
[5]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
[6]University of Exeter and Met Office Hadley Centre, Exeter, UK
[7]Geosciences, University of Edinburgh, Edinburgh, UK

Key Points:

- The IPCC AR6 WG1 states the "frequency and intensity of hot extremes have increased".
- The IPCC notes that the effect of increased greenhouse gas on high temperatures is moderated or amplified at local scales by other factors.
- Confident quantitative attribution statements of the human influence on heatwaves are limited by our understanding of these local processes.

Corresponding author: Michael F. Wehner, mfwehner@lbl.gov

Abstract

It sounds straightforward. As the Earth warms due to the increased concentration of greenhouse gases in the atmosphere, global temperatures rise and so heatwaves become warmer as well. This means that a fixed temperature threshold is passed more often: the probability of extreme heat increases. However, land use changes, vegetation change, irrigation, air pollution and other changes also drive local and regional trends in heatwaves. Sometimes they enhance heatwave intensity, but they can also counteract the effects of climate change, and in some regions the mechanisms that impact on trends in heatwaves have not yet been fully identified. Climate models simulate heatwaves and the increased intensity and probability of extreme heat reasonably well on large scales. However, changes in annual daily maximum temperatures do not follow global warming over some regions, including the Eastern US and parts of Asia, reflecting the influence of local drivers as well as natural variability. Also, temperature variability is unrealistic in many models, and can fail standard quality checks. Therefore, reliable attribution and projection of change in heatwaves remains a major scientific challenge in many regions, particularly where the moisture budget is not well simulated, and where land surface changes, changes in short lived forcers and soil moisture interactions are important.

Plain Language Summary

Heatwaves are arguably the most deadly weather phenomena. As the earth warms due to higher concentrations of greenhouse gases, one would expect heatwaves to become worse as well, killing even more people unless they are better protected against the heat. However, it turns out that the world is not so simple and that many other factors also influence heatwaves. Land use changes, irrigation, air pollution and other changes also drive trends in heatwaves. Some of these cause much larger trends while some have counteracted the climate change driven trends up to now. In some regions the causes of high trends have not yet been identified. Current generation climate models often do not simulate all these mechanisms correctly so will have to be improved before we can more confidently trust their description of past trends and projections of future trends in heatwaves.

1 Introduction

Extreme heat is one of the deadliest natural hazards (Harrington & Otto, 2020) and also is one where climate change really is a game changer. For instance, European heatwaves were diagnosed as the deadliest disaster of 2019 (Vautard et al., 2020). The recently released IPCC report concluded that “It is virtually certain that hot extremes (including heatwaves) have become more frequent and more intense across most land regions since the 1950s, with high confidence that human-induced climate change is the main driver of these changes. Some recent hot extremes observed over the past decade would have been extremely unlikely to occur without human influence on the climate system” (Seneviratne et al., 2021). However, challenges arise in the observed trends at the local to regional scales that matter for planning and adaptation. Although changes in heatwaves are widely thought to be simpler to attribute to anthropogenic climate change than precipitation events, there remain significant challenges. Recent work of the World Weather Attribution initiative (WWA) has highlighted the general issues in attributing regional changes in extreme weather (van Oldenborgh, van der Wiel, et al., 2021; Philip et al., 2020). This paper focuses on the specific problems that can be encountered when attributing the human influence on heatwaves.

First, the observed trends in heatwave frequencies are not always positive, or at least do not closely follow global warming. Figure 1a shows the trends in the maximum temperature of the hottest day of the year (TXx) for the last century of GHCN-D v2 stations as a regression on smoothed global mean temperature. Apart from individual station records showing breaks or spurious trends, there are coherent areas with negative or zero trends. In the Central Plains of the United States, the highest temperatures were observed during the Dust Bowl of the 1930s (Cook et al., 2011; Donat et al., 2016; Cowan, Hegerl, et al., 2020), not in recent years, while in the Central-Eastern US, hot extremes are also not steadily increasing with global mean temperature during recent decades and daytime maxima show different trends from minima (Portmann et al., 2009a). In India, extreme high temperatures have no or very small trends since the 1970s (van Oldenborgh et al., 2018a) (Fig. 1b). This highlights that drivers other than anthropogenic GHG emissions also play an important role in heatwaves. Second, climate models do simulate increasing frequencies and intensities in heatwaves on large geographical scales but their skill in simulating the observed trends on smaller scales is collectively poor across the world, notably in regions with good observational data and huge modelling efforts, e.g. Europe and North America. In an attribution study of the 2019 European record heatwaves, we found that the highest temperatures of the year in that region have increased in observations much more than in 44 analysed global and regional climate simulations in this region (Vautard et al., 2020). Similar problems were already reported in previous generation regional climate models (Min et al., 2013), suggesting that model development has not addressed these deficiencies. In other regions, notably eastern North America and India (Donat et al., 2017; van Oldenborgh et al., 2018a; Cowan, Hegerl, et al., 2020), the problem is reversed with models considerably overestimating the observed trends. In addition, there is a lack of consistency in simulating the magnitude of trends in heat extremes in different model ensembles (regional EURO-CORDEX vs. global CMIP5) and model generations (CMIP5 vs. CMIP6) (Coppola et al., 2020). While there is little difference between the CMIP5 and CMIP6 ensembles in global skill metrics of their simulation quality of average TXx, many models fail to adequately simulate long period return values of extreme heat (Wehner et al., 2020). Comprehensive analyses relating such metrics to model performance in simulating trends have yet to be conducted.

While there is no doubt that at very large spatial and temporal scales heatwaves are increasing and models do represent this, the change in daily maximum temperature, i.e., extreme heat, is very different on the scales where people live and decisions on preparedness are made. With the current generation of climate models we are unable to quantify this change reliably, which affects our ability to reliably attribute changes in probabilities of hot extremes on relevant spatial scales. Given these discrepancies in representing the past, confidence in quantitative projections of heat extremes remains low. In the remainder of the paper we illustrate the problem, and discuss and test reasons for these discrepancies that have been suggested in the literature, ending with a set of priorities for future research.

2 Heatwave characteristics

We start our investigation by laying out basic properties of heatwaves, not all of which are well-known. Any assessment on changes in heatwaves depends strongly on how these are defined. Definitions frequently employed include continent-averaged seasonal mean temperature (e.g., Stott et al. (2004)), a quantity climate models are able to simulate well, and that is strongly correlated to external forcing. Such a definition also maximises the signal to noise ratio as natural variability is averaged out more than the anomaly corresponding to the event itself (Angélil et al., 2018). At the other end of the spectrum is the local instantaneous highest single day temperature in a year (often denoted by TXx). This definition corresponds to a broad understanding of heatwaves in the general public as the media usually reports daily records. It also corresponds to health impacts in places where the most vulnerable population is working outdoors, for instance outdoor labourers in India (Nag et al., 2009) or in Central California (Castillo et al., 2021). In Europe, a few days' average of daily mean or maximum temperature describes the impacts of extreme heat on the population better, accounting for accumulation of the ef-

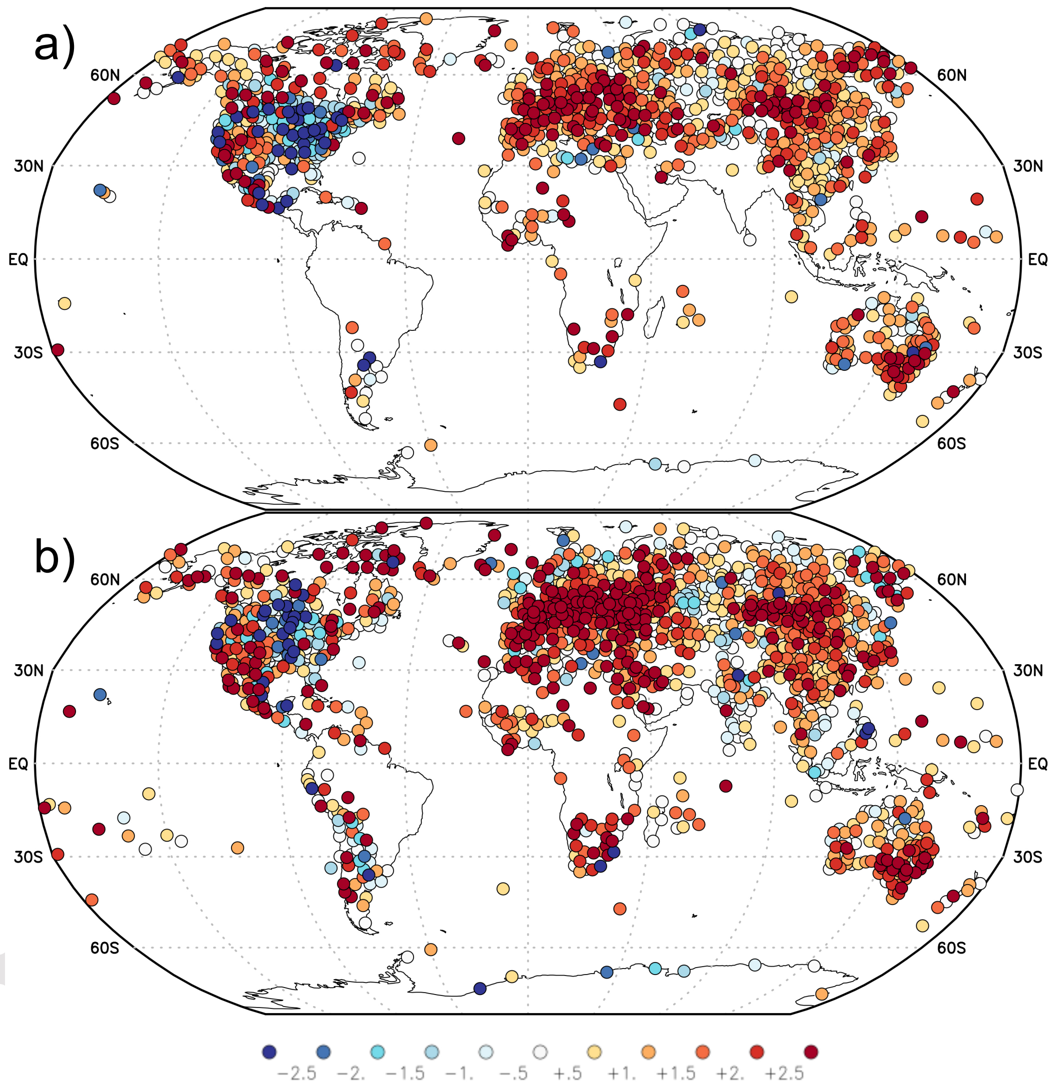

Figure 1. a) Trend in the highest maximum temperature of the year (TXx) as a regression on 4-yr smoothed global mean surface temperature (GMST). GHCN-D v2 stations with a minimum radial separation of 2°, and at least 50 years of data in 1900–2019 are shown. b) The same for 1970–2019 and at least 30 years of data. Units: °C per °C global warming.

fect of heat with the most vulnerable population indoors (D'Ippoliti et al., 2010; Heaviside et al., 2017). Here we consider annual maximum daily maximum temperature, TXx, generally the hottest summer afternoon each year, as it is very widely used (Hartmann et al., 2013) and can be well compared with station observations. We have previously found that using the maximum of the three-day running mean of daily mean temperature (TG3x) is more appropriate for health impacts in Europe and gives very similar results (Kew et al., 2019; Vautard et al., 2020).

In general, the distribution of the temperature of the hottest afternoon of the year is described well by a general extreme value distribution (GEV), in agreement with extreme value theory (Coles, 2001). The distribution is not stationary but changes with global warming and other drivers of local temperature trends. An efficient and often realistic way to describe these changes is to assume the whole distribution shifts up with an indicator of climate change, for which the smoothed global mean surface temperature (GMST) is an often used metric (Philip et al., 2020). This variable is well-estimated and updated in real time. The scale and shape parameters describing the variability and tail shape are thus assumed constant. As an example the observations and fit are shown in Fig. 2 for De Bilt in the Netherlands, which has a long homogenised record. The low-pass filtered time series resembles the well-known GMST increase, indicating that global warming is the dominant driver of the non-stationarity, thus justifying its use as a covariate. A secondary driver might be local aerosols intercepting incoming solar radiation, but this averages out in the fit as the effects of dimming from the 1960s to 1980s and the subsequent brightening up to the 2000s cancel if the analysis period includes both.

The curves in Fig. 2b denote the GEV for two values of the GMST, in a 1.2 °C cooler world (blue, early-industrial) and 2019 (red, the current climate during a recent extreme). For comparison the observations are shown for the early-industrial and current climates, shifted with the fitted trend from the actual smoothed GMST.

The shape parameter of the GEV distribution is almost always found to be negative in heatwave analyses, resulting in the distribution having an upper bound (Wehner et al., 2018). This shape of the tail implies that the probability of an event to occur decreases rapidly as the upper bound is approached and is zero above it. We are not aware of a rigorous derivation of the origin of the upper bound in the literature. We think it could be a consequence of the non-linearities in the surface energy balance and its interaction with the water balance, plus convection as a moderating effect. Both the sensible and latent heat fluxes increase rapidly with temperature. The assumption of constant scale and shape parameters in the distribution implies that the upper bound shifts with the rest of the distribution, which is found in observations as well as historical model simulations (Vautard et al., 2020).

This procedure of fitting a GEV shifting with GMST to the observed annual maxima allows us to answer the questions how much hotter and more likely extreme heat is now than it was a century ago. Applying this method to the TXx at De Bilt observed on 27 July 2019, 37.5 °C, denoted by the purple line in Fig. 2b, we find that the record observed in 2019 would have been virtually impossible in the climate of 1900 (the purple line is above the blue central curve representing the best fit). Taking the upper bound of the 95% confidence interval (obtained from bootstrapping (Philip et al., 2020)) gives a return time of *at least* 15,000 yr in the climate of 1900. In the warmer climate of today the return period of that event is about 30 yr, with a lower bound of 13 yr (intersections with the red curves), while the magnitude of a temperature extreme of this rarity is about 4.0 ± 1.1 °C (2σ bounds) higher than it would have been in the early-industrial climate. Similar analyses have been done for areas where heatwaves have not increased at all in temperature (van Oldenborgh et al., 2018a). However, these observational analyses only detect a trend or its absence, they cannot attribute the causes of it.

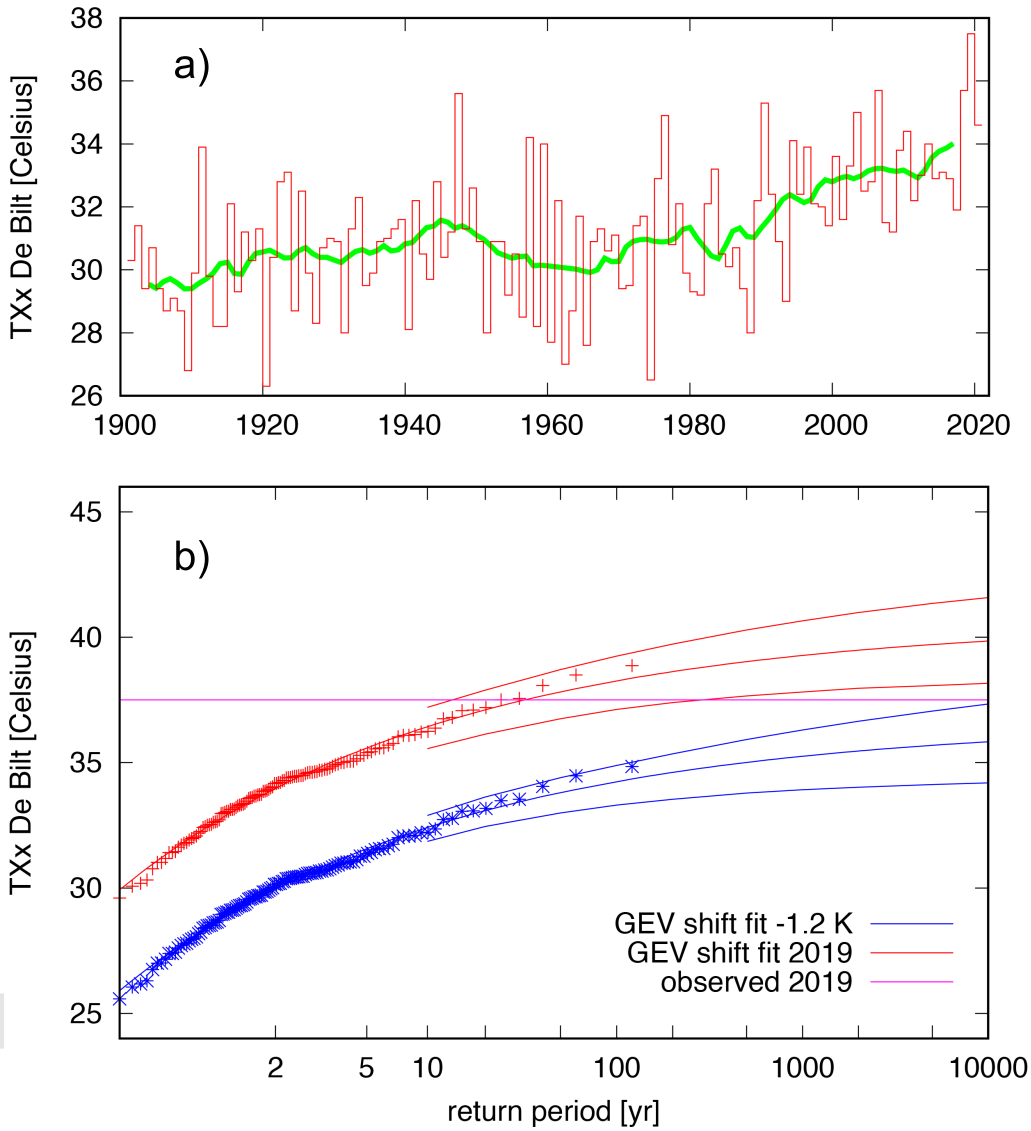

Figure 2. a) Highest maximum temperature of the year (TXx) at De Bilt, the Netherlands. The green curve denotes a ten-year running mean. b) Gumbel (return time) plot of a GEV fit of TXx shifted with the smoothed GMST. The red lines indicate the fit and the 95% confidence intervals in the current climate (2019), the blue lines in the early-industrial climate (1.2 °C lower GMST). The observations are shown twice: once shifted to the early-industrial climate using the fitted trend (blue stars), once shifted to the climate of 2019 (red pluses). The purple line denotes the value observed in 2019, which is included in the fit.

3 Potential causes of heatwave trends

Long-term changes in heatwaves are influenced not only by globally well mixed greenhouse gases but also by more localised influences, including aerosol trends (Péré et al., 2011), land use changes (Cowan, Hegerl, et al., 2020), vegetation and soil moisture changes (Donat et al., 2017), irrigation (Thiery et al., 2017) and urbanisation effects (Heaviside et al., 2017). Furthermore, the meteorological conditions conducive to heatwaves could change regionally by potential changes in mean atmospheric circulation or in the frequency of specific weather patterns leading to extreme heat (Horton et al., 2015).

Local circumstances such as the thermometer screen and its immediate surroundings also influence TXx in observations disproportionately and must be homogenised before observations can be compared to models. The De Bilt series has been homogenised for the change in screen and displacement to a less sheltered location in 1951 (Brandsma, 2016). Urban heat effects could in theory also affect the highest temperatures, but in De Bilt they are small as record temperatures are always attained during southerly or south easterly wind directions with no urban areas within 10 km upstream. However, for other stations it might not be as straightforward to identify whether such local effects are small as, for instance many inner-city stations are in city parks. The anomalously large or small trends observed in these stations (e.g., Madrid Retiro and Dublin Phoenix Park) give the suspicion that they could be influenced more by changes in the lawn sprinkling schedules than global warming, so we avoid using these observations in our attribution analyses and would recommend not using them for model/observation comparisons. These issues call for a detailed investigation of station temperature homogeneity and a massive effort in homogenisation in many places of the world.

4 Climate model ensembles can misrepresent local heatwave trends

To disentangle all these effects on heatwaves and isolate the change driven by anthropogenic climate change, we have to turn to climate models. Fig. 3 shows the simulated trend in TXx in the CMIP5 ensemble of opportunity (Sillmann et al., 2013) with a median resolution of about 200 km over roughly the same periods as Fig. 1. (We have excluded MIROC-ESM and MIROC-ESM-CHEM, as these have intermittent physically impossible high temperatures in the deserts.) The maps show less structure than the observed trends. This is partly due to the natural variability being averaged out and partly due to missing or misrepresented local forcings. Notably, in neither time period do the multi-model average represent the observed negative or neutral trends in TXx in central and eastern North America. The central great plains early heatwaves have been linked to rapid devegetation in the 1930s associated with the dustbowl drought, which led to record heatwaves at the time not yet superseded (Cowan, Hegerl, et al., 2020; Cowan, Undorf, et al., 2020). A factor in the negative trends in the eastern US may be downstream effects of increasing irrigation further west (DeAngelis et al., 2010) from the 1950s onwards, coinciding with a positive precipitation trend (Portmann et al., 2009b; Kirtman et al., 2013). Changes in agricultural practices leading to higher evaporation have also been implicated (Changnon et al., 2003). It has been speculated that revegetation after the decline of agriculture might also have been a factor (Portmann et al., 2009b). The CMIP5 models do not include cooling due to irrigation, which leads to biases in trends over the US, Iran, Pakistan, and India (Thiery et al., 2017; Mueller et al., 2016) although some specialized simulations do (Lobell & Bonfils, 2008; Lawston et al., 2020). They furthermore likely misrepresent the warming effect of black carbon and the cooling effect of sulfate aerosols over India (Padma Kumari et al., 2007) nor are they forced with rapid vegetation changes (Cowan, Undorf, et al., 2020).

While Fig. 3 does show a stronger warming trend over Europe than in other parts of the world, the multi-model average does not accurately represent the much higher observed trends in western Europe (Min et al., 2013) or southeastern Australia (van Old-

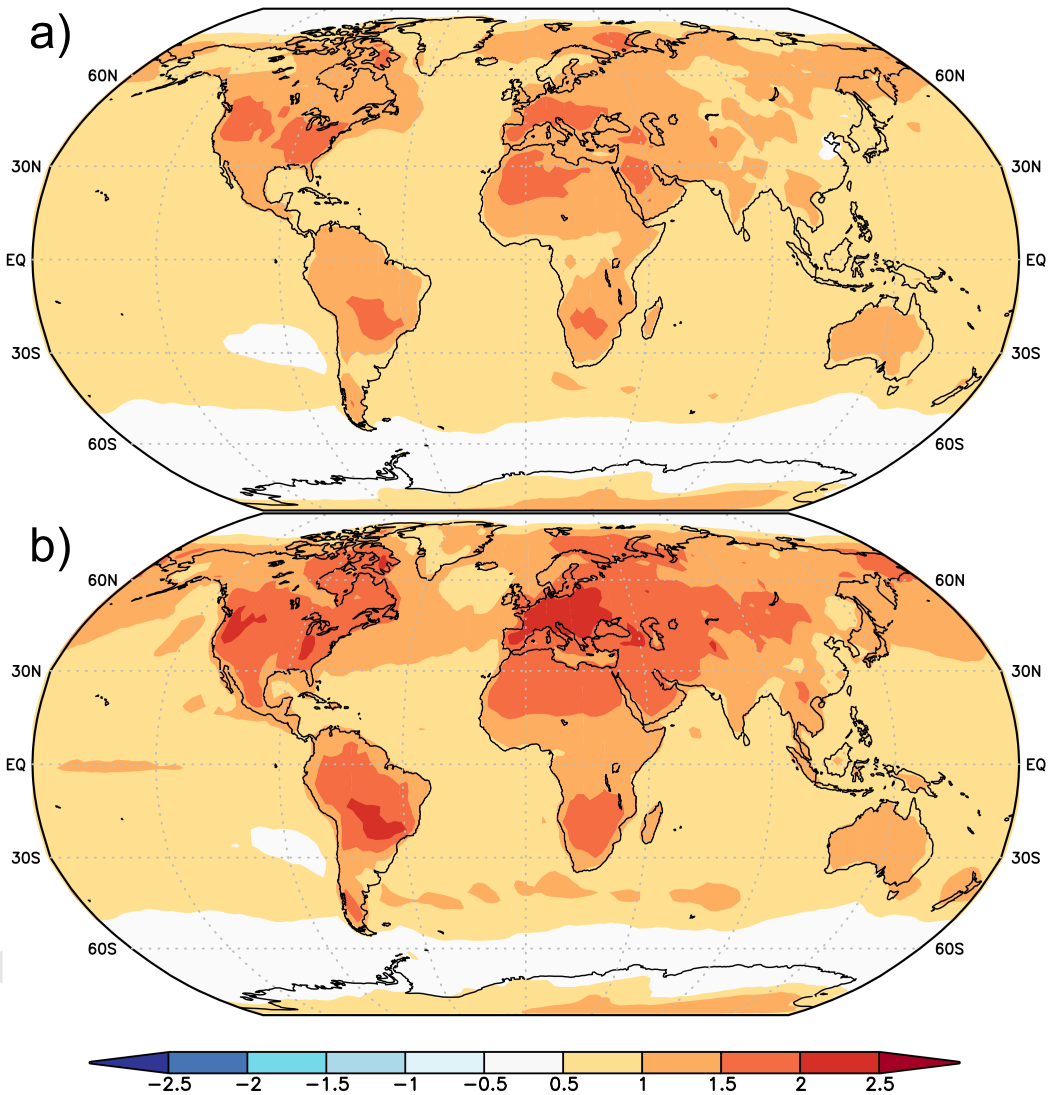

Figure 3. Trend in the highest maximum temperature of the year (TXx) as a regression on 4-yr smoothed global mean surface temperature (GMST) as in Fig. 1a but the historical/RCP4.5 CMIP5 ensemble (Sillmann et al., 2013) for a) 1900–2019 and b) 1950–2019 using the ensemble mean global mean temperature as covariate. Units: °C per °C global warming.

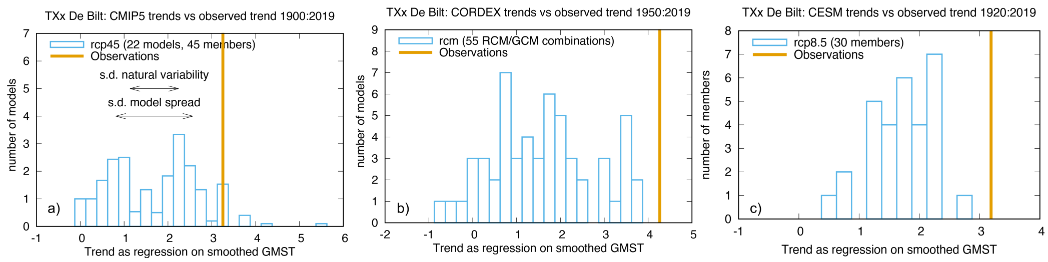

Figure 4. a) Histogram of the TXx trends at De Bilt in the CMIP5 (Taylor et al., 2011) ensemble compared to the observed trend 1900–2019, both expressed as a regression on the (modelled/observed) smoothed GMST. The standard deviation (s.d.) of natural variability is estimated from models with three or more ensemble members. b) The same for 55 CORDEX RCM/GCM combinations and observations using 1950-2019 or 1970-2019 depending on data availability, and using observed GMST (1950-2019) as reference (Coppola et al., 2020). c) The same for 30 realizations of the CESM Large Ensemble over the period 1920-2019 (Deser et al., 2020). Units: °C per °C global warming.

enborgh, Krikken, et al., 2021). So far there has been little progress determining whether these discrepancies are due to missing or misrepresented local forcings (aerosols, land use, vegetation, irrigation), overly strong land surface drying in historical heatwaves, or due to natural variability, misrepresented feedbacks or changes in the observational methods or their local surroundings. Two decades ago systematic errors in blocking frequency and persistence were a major source of biases in weather and climate models (Palmer et al., 1990), but in modern models these are realistic in the European summer (Vautard et al., 2020; Krikken et al., 2019; Iles et al., 2020) and thus not a reason for the persisting model deficiencies there.

With respect to natural variability, we find in many locations that the discrepancies between observed and modelled trends are much larger than can be expected on the basis of natural variability and model spread alone. We use again De Bilt as an example, but the results are similar across Western Europe. Fig. 4a shows the histogram of trends in the CMIP5 models at the location of De Bilt over 1900-2019, with models with N runs each entered with weight $1/N$ so that all models have equal weight. Model results show the grid cell enclosing the observation station. If that is an ocean cell, then the nearest cell to the east or west is used (van Oldenborgh, van der Wiel, et al., 2021). Only 6 of the 10 CSIRO ensemble members have trends higher than the observed one, all other models have lower trends. However, the CSIRO model places the Mediterranean warming trend too far north and underestimates the observed global warming trend by 30%. Both these factors give a high trend in heatwaves relative to the global mean temperature rise, but for the wrong reasons: the climate of the Netherlands is not Mediterranean and the CSIRO global mean temperature rise is unrealistically low. The CMIP5 ensemble thus fails to reproduce the observed trends, even though it includes all relevant natural variability, including possible low-frequency effects from the subpolar gyre (Haarsma et al., 2015) and the model spread as proxy for model uncertainty.

Fig. 4b shows the same for 55 RCM/GCM combinations at 11 km resolution of the CORDEX ensemble (Coppola et al., 2020; Vautard et al., 2021) for Europe, over 1951–2019 or 1971–2019 (depending on the models' data availability). In this comparison the observed trend is larger than all modelled trends. This again implies that in this case natural variability is unlikely to be the driver behind the strong increase in heatwaves that are at the moment not correctly represented in climate models. These models have much higher resolution than the CMIP5 models (median resolution 200km) so the problem is not simply solved by going to higher resolution, but also need other improvements.

The recent development of large single model ensembles (Deser et al., 2020) provides an opportunity to better quantify natural variability of extreme temperatures and place observed trends in that context (Tebaldi et al., 2021),. Fig. 4c shows that a histogram of the 1920-2019 De Bilt temperature trends from the Community Earth System Model Large Ensemble (CESM1 LENS) fails to include the observed trend. As for the CMIP5 and CORDEX multi-model ensembles, the modeled trends are lower than observed. We note that variability in mean temperature trends can be overestimated in large ensembles (McKinnon et al., 2017) further suggesting that important processes are missing. Nonetheless, we encourage the modeling community to expand large ensemble simulations to the DAMIP single forcing scenarios (Gillett et al., 2016) to aid in attribution studies.

For other regions where ensembles of models do not reproduce the observed trends in heatwaves, either too high or too low, similar conclusions that factors in addition to natural variability must be considered, as for instance shown qualitatively in India (van Oldenborgh et al., 2018a). For eastern North America, the agreement of models including realistic land use changes with observations (Cowan, Hegerl, et al., 2020) suggests a limited role of natural decadal variability of the atmosphere, which is in agreement with the low correlation between the natural decadal variability and longer term trends in temperature extremes (van Oldenborgh et al., 2012). While the trends shown in figure 4 show that the actual probability of extreme heat in De Bilt is higher than the CMIP5, CORDEX and the CESM1 LENS ensembles can produce, often the opposite is the case. As Knutson (2017) discusses, statements of "attribution without detected changes" in observations can still be useful, albeit with lower confidence than when observed and simulated trends are mutually consistent.

5 Biases in variability

These biases in the trend are not the only problem. In particular, to accurately attribute change in probability of extreme events to anthropogenic climate change the variability of the extremes in the model is as important as the trend. In the case of heatwaves the upper limit of the probability of an event in the current climate divided by its probability in a climate without global warming (commonly referred to as the probability ratio, PR) rises with increasing variability (Philip et al., 2020). Almost all climate models analysed have unrealistically high variability, with factors of 1.5 to 6 higher scale parameters in GEV fits of the high tail in Europe (Leach et al., 2020) and this bias is also apparent in subtropical and tropical regions (Freychet et al., 2021). Overestimation of the variability in extreme temperatures by climate models undermines confidence in our understanding of heatwave trends (Kew et al., 2019; Vautard et al., 2020; van Oldenborgh, Krikken, et al., 2021) as none of the models pass a frequently employed model evaluation test (Philip et al., 2020), demanding that the model GEV fits are compatible (within the sampling uncertainty) with the observed GEV fit of the observations. In such cases, best estimate attribution statements should not be made. However, this inconsistency between models and observations does not preclude placing conservative lower bounds on the human influence of heatwaves. The overestimation in the variability of extreme temperatures remains unexplained and could result from several processes, e.g., excessive land-atmosphere-cloud-precipitation feedbacks (Miralles et al., 2019).

6 Conclusions

While large scale changes in mean temperature are well understood, changes in local and regional heatwaves, particularly, daytime maxima, are much harder to simulate and hence attribute. This failure to understand today's observed trends and the discrepancies between the modelled and observed trends and variability also hinders confidence in projections of the future trends. The extrapolation of the observed trends shown in Fig. 1a is very different from the simulated trends from climate model output over the same period in Fig. 3.

Heatwaves on the scales people experience them are strongly influenced by the local energy budget that determines the use of energy between evaporation and heating, set by the land surface, vegetation, irrigation and urbanisation. Other factors such as circulation changes or aerosols may also be important and feedbacks may well be misrepresented in climate models during these extreme circumstances. Many of these drivers and feedbacks are not well-simulated in current climate models as evidenced by striking discrepancies between observed and modelled trends and variability in certain regions of the globe. We have shown above that the discrepancies cannot always be explained by natural variability and in some cases are well outside the range of CMIP historical simulations even in well understood regions (Cowan, Undorf, et al., 2020; van Oldenborgh et al., 2018b). Diffenbaugh et al. (2017) use four performance metrics to compare observed and simulated trends in TXx finding broad regions, mostly in Eurasia, where models are deemed adequate. But they also reject the parts of North America we discussed in section 1. (India was not included in their analyses of TXx). We have also shown that the failure of climate models to represent trends in heatwaves does not change with the resolution of the model nor can a previously observed failure in climate models to represent atmospheric blocking be identified in the current generation of models.

On the other hand, process studies have indicated a strong role of surface conditions such as vegetation and moisture availability, suggesting that adapting local vegetation and moisture conditions may be able to moderate, to an extent, extreme local heat (e.g., Stone Jr et al. (2014); Heaviside et al. (2017)). We have further highlighted that recent studies indicate that the overestimation of trends in regions like North America and India could be due to a misrepresentation of local irrigation and aerosol effects. Given no corresponding trends in either of these drivers in regions where models considerably underestimate trends in extreme heat this cannot be the explanation for all deficiencies. Similarly, while changes in measurement technique and location can explain some discrepancies, they are very unlikely to explain the systematic and widespread discrepancies.

This leaves still an uncomfortably large list of potential reasons for our current lack of understanding of the drivers of extreme heat, including land use changes and soil moisture, aerosol effects and atmosphere feedbacks as well as circulation effects other than blocking. Until we simulate realistic effects of all relevant drivers and feedbacks such that these properties agree within the uncertainties of natural variability of the weather in our climate model simulations, we cannot give confident estimates for the change in frequency and intensity in heatwaves due to anthropogenic global warming up to today in those areas where these missing processes are important, but only lower or upper bounds. Nor can we confidently trust the projections of future heatwaves there. There remain three possible broader reasons for divergence between observed and simulated heat extremes at the scales that affect people. First, the possibility that the models are right, but are being given incomplete local information such as missing land surface feedbacks and use changes, etc. Second, the possibility that the models are truly incorrect and would not have captured observed trends even if these regionally-specific matters were fully incorporated in the modeling framework. Third, that natural variability at local scales predominates over anthropogenic forcing and that the models either do not simulate internal variability correctly or our ensembles are not large enough to capture it. Careful simulation and evaluation of historic events in the context of natural variability helps distinguish these contributing factors. In our view it is thus an immensely important priority for climate model development studies to focus on extreme heat, the deadliest and most immediate effect of human-induced climate change.

Eulogy

Geert Jan van Oldenborgh passed away on October 12, 2021 before he could respond to the reviewers' comments. While we mourn the loss of our friend and colleague, we celebrate his life in this, his final scientific paper. His contributions to the science of extreme weather event attribution were immense and will continue to influence us and many others as we continue to understand the effects of global warming on extreme weather.

Acknowledgments

The research and pre-operational settings underpinning this work was supported by the EUPHEME and SERV_FORFIRE projects, which are part of ERA4CS, an ERA-NET initiated by JPI Climate and co-funded by the European Union (Grant #690462). The work was also supported by the French Ministry for solidarity and ecological transition through the 'climate services convention' between the Ministry and the Centre National de la Recherche Scientifique. MFW acknowledges support from the Regional & Global Model Analysis (RGMA) program area of the DOE's Office of Science under contract number DE-AC02-05CH11231 and GH acknowledges funding from the EMERGENCE project, NERC NE/S004661/1. This paper builds on results from the Copernicus COP 039 and C3S 62 contracts lead by KNMI, which focused on the development of a prototype for extreme events and attribution service within the context of the Copernicus Climate Change Service (C3S) implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) and funded by the European Union.

All data used this work are publicly available:

- GHCN-D v2 is available at https://climexp.knmi.nl/selectdailyseries.cgi
- De Bilt TXx was constructed from the daily time series of The Bilt at `https://climexp.knmi.nl/getdutchstations.cgi?TYPE=tx`
- The CMIP5 TXx values can be found at `https://climexp.knmi.nl/selectfield_cmip5_annual.cgi`
- The 55 RCM/GCM combinations at 11 km resolution of the CORDEX ensemble are available via the ESGF as linked from `https://cordex.org`
- The Community Earth System Model Large Ensemble (CESM1 LENS) is available at `https://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html`

References

Angélil, O., Stone, D. A., Perkins, S. E., Alexander, L. V., Wehner, M. F., Shiogama, H., … Christidis, N. (2018). On the nonlinearity of spatial scales in extreme weather attribution statements. *Clim. Dyn.*, *50*, 2739. doi: 10.1007/s00382-017-3768-9

Brandsma, T. (2016). *Homogenization of daily temperature data of the five principal stations in the Netherlands (version 1.0)* (Technical report No. 356). De Bilt, Netherlands: KNMI. Retrieved from `http://bibliotheek.knmi.nl/knmipubTR/TR356.pdf`

Castillo, F., Sanchez, A., Wehner, M., & Gilless, J. K. (2021). The impact of heat waves on agricultural productivity and output. In F. Castillo, M. Wehner, & D. Stone (Eds.), *Extreme events and climate change: A multidisciplinary approach* (pp. 11–20). Wiley and Sons.

Changnon, D., Sandstrom, M., & Schaffer, C. (2003). Relating changes in agricultural practices to increasing dew points in extreme Chicago heat waves. *Climate Research*, *24*(3), 243–254. doi: 10.3354/cr024243

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* London, UK: Springer Series in Statistics.

Cook, B. I., Seager, R., & Miller, R. L. (2011). Atmospheric circulation anomalies during two persistent north american droughts: 1932–1939 and 1948–1957. *Climate Dynamics*, *36*(11), 2339–2355. doi: 10.1007/s00382-010-0807-1

Coppola, E., Nogherotto, R., Ciarlò, J. M., Giorgi, F., van Meijgaard, E., Kadygrov, N., . . . Teichmann, C. (2020). Assessment of the european climate projections as simulated by the large EURO-CORDEX regional climate model ensemble. *J. Geophys. Res.*, *submitted*.

Cowan, T., Hegerl, G. C., Schurer, A., Tett, S. F. B., Vautard, R., Yiou, P., . . . Ng, B. (2020). Ocean and land forcing of the record-breaking Dust Bowl heatwaves across central United States. *Nature Communications*, *11*(1), 2870. doi: 10.1038/s41467-020-16676-w

Cowan, T., Undorf, S., Hegerl, G. C., Harrington, L. J., & Otto, F. E. L. (2020). Present-day greenhouse gases could cause more frequent and longer Dust Bowl heatwaves. *Nature Climate Change*, *10*(6), 505–510. doi: 10.1038/s41558-020-0771-7

DeAngelis, A., Dominguez, F., Fan, Y., Robock, A., Kustu, M. D., & Robinson, D. (2010). Evidence of enhanced precipitation due to irrigation over the great plains of the united states. *J. Geophys. Res. Atmos.*, *115*(D15). doi: 10.1029/2010JD013892

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., . . . Ting, M. (2020). Insights from earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. doi: 10.1038/s41558-020-0731-2

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., . . . Rajaratnam, B. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences*, *114*(19), 4881–4886. Retrieved from `https://www.pnas.org/content/114/19/4881` doi: 10.1073/pnas.1618082114

D'Ippoliti, D., Michelozzi, P., Marino, C., de'Donato, F., Menne, B., Katsouyanni, K., . . . Perucci, C. A. (2010, Jul 16). The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project. *Environmental Health*, *9*(1), 37. doi: 10.1186/1476-069X-9-37

Donat, M. G., King, A. D., Overpeck, J. T., Alexander, L. V., Durre, I., & Karoly, D. J. (2016). Extraordinary heat during the 1930s us dust bowl and associated large-scale conditions. *Climate Dynamics*, *46*(1), 413–426. doi: 10.1007/s00382-015-2590-5

Donat, M. G., Pitman, A. J., & Seneviratne, S. I. (2017). Regional warming of hot extremes accelerated by surface energy fluxes. *Geophys. Res. Lett.*, *44*(13), 7011–7019. doi: 10.1002/2017GL073733

Freychet, N., Hegerl, G., Mitchell, D., & Collins, M. (2021). Future changes in the frequency of temperature extremes may be underestimated in tropical and subtropical regions. *Communications Earth & Environment*, *2*(1), 28. doi: 10.1038/s43247-021-00094-x

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., . . . Tebaldi, C. (2016). The detection and attribution model intercomparison project (damip v1.0) contribution to cmip6. *Geoscientific Model Development*, *9*(10), 3685–3697. Retrieved from `https://gmd.copernicus.org/articles/9/3685/2016/` doi: 10.5194/gmd-9-3685-2016

Haarsma, R. J., Selten, F. M., & Drijfhout, S. S. (2015). Decelerating Atlantic meridional overturning circulation main cause of future west european summer atmospheric circulation changes. *Environ. Res. Lett.*, *10*(9), 094007. doi: 10.1088/1748-9326/10/9/094007

Harrington, L. J., & Otto, F. E. L. (2020). Reconciling theory with the reality of african heatwaves. *Nature Climate Change*. doi: 10.1038/s41558-020-0851-8

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., et al. (2013). Observations:

Atmosphere and surface. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis* (pp. 159–254). Cambridge, U.K. and New York, U.S.A.: Cambridge University Press.

Heaviside, C., Macintyre, H., & Vardoulakis, S. (2017). The urban heat island: Implications for health in a changing environment. *Current Environmental Health Reports*, *4*(3), 296–305. doi: 10.1007/s40572-017-0150-3

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., & Diffenbaugh, N. S. (2015). Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature*, *522*(7557), 465–469. doi: 10.1038/nature14550

Iles, C., Vautard, R., Strachan, J., Joussaume, S., Eggen, B., & Hewitt, C. (2020). The benefits of increasing resolution in global and regional climate simulations for European climate extremes. *Geosci. Mod. Dev.*, *submitted*.

Kew, S. F., Philip, S. Y., van Oldenborgh, G. J., Otto, F. E., Vautard, R., & van der Schrier, G. (2019). The exceptional summer heatwave in southern Europe 2017. *Bull. Amer. Met. Soc.*, *100*(1), S2–S5. doi: 10.1175/BAMS-D-18-0109.1

Kirtman, B., Power, S. B., et al. (2013). Near-term climate change: Projections and predictability. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis* (pp. 953–1028). Cambridge, U.K. and New York, U.S.A.: Cambridge University Press.

Knutson, T. (2017). Detection and attribution methodologies overview. In D. Wuebbles, D. Fahey, K. Hibbard, D. Dokken, B. Stewart, & T. Maycock (Eds.), *Climate science special report: Fourth national climate assessment, volume i* (pp. 443–451). Washington, DC, USA: U.S. Global Change Research Program. doi: 10.7930/J0319T2J

Krikken, F., Lehner, F., Haustein, K., Drobyshev, I., & van Oldenborgh, G. J. (2019). Attribution of the role of climate change in the forest fires in Sweden 2018. *Natural Hazards and Earth System Sciences Discussions*, *2019*, 1–24. doi: 10.5194/nhess-2019-206

Lawston, P. M., Santanello, J. A., Hanson, B., & Arsensault, K. (2020). Impacts of irrigation on summertime temperatures in the pacific northwest. *Earth Interactions*, *24*(1), 1 - 26. Retrieved from `https://journals.ametsoc.org/view/journals/eint/24/1/EI-D-19-0015.1.xml` doi: 10.1175/EI-D-19-0015.1

Leach, N. J., Li, S., Sparrow, S., van Oldenborgh, G. J., Lott, F. C., Weisheimer, A., & Allen, M. R. (2020). Anthropogenic influence on the 2018 summer warm spell in Europe: The impact of different spatio-temporal scales. *Bull. Amer. Met. Soc.*, *101*(1), S41–S46. doi: 10.1175/BAMS-D-19-0201.1

Lobell, D. B., & Bonfils, C. (2008). The effect of irrigation on regional temperatures: A spatial and temporal analysis of trends in california, 1934-2002. *Journal of Climate*, *21*(10), 2063 - 2071. Retrieved from `https://journals.ametsoc.org/view/journals/clim/21/10/2007jcli1755.1.xml` doi: 10.1175/2007JCLI1755.1

McKinnon, K. A., Poppick, A., Dunn-Sigouin, E., & Deser, C. (2017). An observational large ensemble to compare observed and modeled temperature trend uncertainty due to internal variability. *Journal of Climate*, *30*(19), 7585 - 7598. Retrieved from `https://journals.ametsoc.org/view/journals/clim/30/19/jcli-d-16-0905.1.xml` doi: 10.1175/JCLI-D-16-0905.1

Min, E., Hazeleger, W., van Oldenborgh, G. J., & Sterl, A. (2013). Evaluation of trends in high temperature extremes in north-western Europe in regional climate models. *Environ. Res. Lett.*, *8*(1), 014011. doi: 10.1088/1748-9326/8/1/014011

Miralles, D. G., Gentine, P., Seneviratne, S. I., & Teuling, A. J. (2019, 01). Land-atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Annals of the New York Academy of Sciences*, *1436*(1), 19–35. doi: 10.1111/nyas.13912

Mueller, N. D., Butler, E. E., McKinnon, K. A., Rhines, A., Tingley, M., Holbrook, N. M., & Huybers, P. (2016). Cooling of US Midwest summer temperature extremes from cropland intensification. *Nature Climate Change*, *6*(3), 317–322. Retrieved from `https://doi.org/10.1038/nclimate2825` doi: 10.1038/nclimate2825

Nag, P. K., Nag, A., Sekhar, P., & Pandt, S. (2009). *Vulnerability to heat stress: Scenario in Western India* (WHO APW No. SO 08 AMS 6157206). Ahmedabad: National Institute of Occupational Health. Retrieved from `https://www.who.int/docs/default-source/searo/india/health-topic-pdf/occupational-health-vulnerability-to-heat-stress-scenario-of-western-india.pdf?sfvrsn=9568d2b6\_2`

Padma Kumari, B., Londhe, A. L., Daniel, S., & Jadhav, D. B. (2007). Observational evidence of solar dimming: Offsetting surface warming over India. *Geophys. Res. Lett.*, *34*(21), L21810. doi: 10.1029/2007GL031133

Palmer, T., Brankovic, C., Molteni, F., Tibaldi, S., Ferranti, L., Hollingsworth, A., & Cubasch, E., U.and Klinker. (1990). The European Centre for Medium-Range Weather Forecasts (ECMWF) program on extended-range prediction. *Bulletin of the American Meteorological Society*, *71*(9), 1317–1330. doi: 10.1175/1520-0477(1990)071⟨1317:TECFMR⟩2.0.CO;2

Péré, J. C., Mallet, M., Pont, V., & Bessagnet, B. (2011). Impact of aerosol direct radiative forcing on the radiative budget, surface heat fluxes, and atmospheric dynamics during the heat wave of summer 2003 over western europe: A modeling study. *Journal of Geophysical Research: Atmospheres*, *116*(D23), D23119 (12pp). doi: 10.1029/2011JD016240

Philip, S. Y., Kew, S. F., van Oldenborgh, G. J., Otto, F. E. L., Vautard, R., van der Wiel, K., ... van Aalst, M. K. (2020). A protocol for probabilistic extreme event attribution analyses. *Advances in Statistical Climatology, Meteorology and Oceanography*, *6*, 177–203. doi: 10.5194/ascmo-6-177-2020

Portmann, R. W., Solomon, S., & Hegerl, G. C. (2009a). Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States. *Proceedings of the National Academy of Sciences*, *106*(18), 7324–7329. doi: 10.1073/pnas.0808533106

Portmann, R. W., Solomon, S., & Hegerl, G. C. (2009b). Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States. *Proceedings of the National Academy of Sciences*, *106*(18), 7324–7329. doi: 10.1073/pnas.0808533106

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., ... Zhou, B. (2021). Weather and Climate Extreme Events in a Changing Climate. In V. Masson-Delmotte et al. (Eds.), *Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change* (chap. 11). Cambridge University Press. Retrieved from `https://www.ipcc.ch/`

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.*, *118*(4), 1716–1733. doi: 10.1002/jgrd.50203

Stone Jr, B., Vargo, J., Liu, P., Habeeb, D., DeLucia, A., Trail, M., ... Russell, A. (2014, 06). Avoided heat-related mortality through climate adaptation strategies in three Us cities. *PLOS ONE*, *9*(6), 1–8. doi: 10.1371/journal.pone.0100852

Stott, P. A., Stone, D. A., & Allen, M. R. (2004, 12 02). Human contribution to the European heatwave of 2003. *Nature*, *432*(7017), 610–614. doi: 10.1038/nature03089

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2011). An overview of CMIP5 and the experiment design. *Bull. Amer. Met. Soc.*, *93*, 485–498. doi: 10.1175/BAMS-D-11-00094.1

Tebaldi, C., Dorheim, K., Wehner, M., & Leung, R. (2021). Extreme metrics from large ensembles: investigating the effects of ensemble size on their estimates. *Earth System Dynamics*, *12*(4), 1427–1501. Retrieved from https://esd.copernicus.org/articles/12/1427/2021/ doi: 10.5194/esd-12-1427-2021

Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., & Seneviratne, S. I. (2017). Present-day irrigation mitigates heat extremes. *Journal of Geophysical Research: Atmospheres*, *122*(3), 1403–1422. doi: 10.1002/2016JD025740

van Oldenborgh, G. J., Doblas-Reyes, F. J., Wouters, B., & Hazeleger, W. (2012). Decadal prediction skill in a multi-model ensemble. *Clim. Dyn.*, *38*, 1263-1280. doi: 10.1007/s00382-012-1313-4

van Oldenborgh, G. J., Krikken, F., Lewis, S., Leach, N. J., Lehner, F., Saunders, K. R., ... Otto, F. E. L. (2021). Attribution of the Australian bushfire risk to anthropogenic climate change. *Natural Hazards and Earth System Sciences*, *21*, 941–960. doi: 10.5194/nhess-21-941-2021

van Oldenborgh, G. J., Philip, S., Kew, S., van Weele, M., Uhe, P., Otto, F., ... AchutaRao, K. (2018b). Extreme heat in india and anthropogenic climate change. *Natural Hazards and Earth System Sciences*, *18*(1), 365–381. Retrieved from https://nhess.copernicus.org/articles/18/365/2018/ doi: 10.5194/nhess-18-365-2018

van Oldenborgh, G. J., Philip, S. Y., Kew, S. F., van Weele, M., Uhe, P., Otto, F. E. L., ... AchutaRao, K. (2018a). Extreme heat in India and anthropogenic climate change. *Natural Hazards and Earth System Sciences*, *18*(1), 365–381. doi: 10.5194/nhess-18-365-2018

van Oldenborgh, G. J., van der Wiel, K., Kew, S., Philip, S., Otto, F., Vautard, R., ... van Aalst, M. (2021). Pathways and pitfalls in extreme event attribution. *Climatic Change*, *166*(1), 13. doi: 10.1007/s10584-021-03071-7

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., ... Wulfmeyer, V. (2021). Evaluation of the large EURO-CORDEX regional climate model ensemble. *Journal of Geophysical Research: Atmospheres*. (e2019JD032344 2019JD032344) doi: 10.1029/2019JD032344

Vautard, R., van Aalst, M. K., Boucher, O., Drouin, A., Haustein, K., Kreienkamp, F., ... Wehner, M. F. (2020). Human contribution to the record-breaking June and July 2019 heat waves in Western Europe. *Environ. Res. Lett.*, *15*, 094077. doi: 10.1088/1748-9326/aba3d4

Wehner, M., Gleckler, P., & Lee, J. (2020). Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 1, model evaluation. *Weather and Climate Extremes*, 100283. Retrieved from http://www.sciencedirect.com/science/article/pii/S2212094719302440 doi: https://doi.org/10.1016/j.wace.2020.100283

Wehner, M., Stone, D., Shiogama, H., Wolski, P., Ciavarella, A., Christidis, N., & Krishnan, H. (2018). Early 21st century anthropogenic changes in extremely hot days as simulated by the C20C+ detection and attribution multi-model ensemble. *Weather and Climate Extremes*, *20*, 1 - 8. doi: 10.1016/j.wace.2018.03.001